How do the metabolites, GTP and (p)ppGpp, simultaneously control the occurrence of translational errors and resource allocation in bacteria?

Claire Baudier

université PARIS-SACLAY

NNT : 2018SACLS202

How do the metabolites, GTP and (p)ppGpp, simultaneously control the occurrence of translational errors and resource allocation in bacteria?

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

École doctorale n°577 structure et dynamique des systèmes vivants (SDSV)
Spécialité de doctorat: Sciences de la vie et de la santé

Thèse présentée et soutenue à Jouy-en-Josas, le 2 juillet 2018, par

Claire Baudier

Composition du Jury :

Philippe Bouloc Directeur de recherche, CNRS (– UMR 9198)	Président
Emmanuelle Bouveret Directrice de recherche, CNRS (– UMR 7255)	Rapportrice
Ivan Matic Directeur de recherche, Inserm (– UMR 1001)	Rapporteur
Isabelle Martin-Verstraete Professeure, Institut Pasteur	Examinatrice
Philippe Glaser Directeur de recherche, Institut Pasteur	Examinateur
Stéphane Aymerich Directeur de recherche, INRA (– UMR 1319)	Directeur de thèse
Vincent Fromion Directeur de recherche, INRA (– UR1404)	Co-Directeur de thèse
Matthieu Jules Professeur, AgroParisTech (– UMR 1319)	Co-Directeur de thèse

Thèse de doctorat

Remerciements

Cette thèse a été réalisée au sein du groupe Systems Biology for bacterial Engineering and Redesign (SyBER) qui fait partie de l'équipe Genetic Controls in Bacterial Systems (GCBS) de l'unité Microbiologie de l'Alimentation au service de la Santé (MICALIS) ainsi qu'au sein de l'équipe BioSys de l'unité Mathématiques et Informatique Appliquées du Génome à l'Environnement (MaIAGE) de l'INRA de Jouy-en-Josas. Le contrat de thèse a lui été financé par l'Initiative Doctorale Interdisciplinaire (IDI) 2014 de l'Université Paris Saclay ainsi que par la Fondation pour la Recherche Médicale (FRM : FDT20170437429).

Je tiens à remercier chaleureusement mon directeur de thèse Stéphane Aymerich de m'avoir acceptée au sein de son équipe GCBS et d'être resté disponible malgré un emploi du temps chargé.

Je souhaite remercier très chaleureusement mes deux co-directeurs de thèse Vincent Fromion et Matthieu Jules pour m'avoir acceptée au sein de leurs équipes respectives (BioSys et SyBER) mais surtout pour m'avoir encadrée tout au long de cette thèse. Je leur suis également reconnaissante pour le temps qu'ils m'ont accordée, leurs qualités pédagogiques et scientifiques (chacun à leur maniére), et leur grande sympathie. J'ai beaucoup appris à leurs côtés tout en ayant ma part d'autonomie et je leur adresse ainsi toute ma gratitude.

Je voudrais remercier les rapporteurs de cette thèse Emmanuelle Bouveret et Ivan Matic pour l'intérêt qu'ils ont porté à mon travail. J'associe à ces remerciements Philippe Bouloc, Philippe Glaser et Isabelle Martin-Verstraete, pour avoir accepté d'examiner mon travail.

Je remercie également mon tuteur de thèse Olivier Martin et une fois de plus Philippe Glaser pour avoir fait partie de mon comité de thèse. Ce travail de thèse a ainsi pu profiter de leur expertise et de leurs remarques pertinentes.

Je tiens à chaleureusement remercier tous les membres et ex-membres de l'équipe RG2B qui m'ont aidée, soutenue et donnée le sourire pendant toute cette thèse: Anne-Gaëlle (qui va me taquiner maintenant?!), Aurélie (merci pour ta bonne humeur et tes conseils!), Dominique, Etienne (un énorme merci pour tes conseils de grand sage qui ont assuré le succès de cette thèse!), Lydia, Magali (encore merci pour avoir gérer mes manips de fin de thèse!), Narimane, Michael (Vive les pity party du week-end!), Michelle, Olivier (Leçon retenue: utiliser le Lab collector et manger bio!), Paola, Philippe, Rama, Steven (je dois l'avouer, tes blagues d'un goût douteux vont me manquer!), Teddy (je glisse Mathieu P. dans la foulée et j'en profite pour réclamer mon poster d'Alexandra Ledherman!), Transito et Vincent (malgré le temps passé à discuter foot à la cantine, je ne suis toujours pas une passionnée du ballon rond!). Je tiens a remercier tout particulièrement mes codoctorantes Cécile, Gabriella et Marwa avec qui j'ai partagé les hauts et les bas du doctorat mes surtout de très bons moments.

Je tiens également à remercier les membres de l'équipe ProCeD dont l'aide m'a été précieuse à différents moments de ma thèse. Et de manière générale, je tiens à remercier tous les membres de MICALIS et en particulier la prépa qui ont contribué d'une manière ou d'une autre à la réussite de cette thèse. Je remercie également les doctorants et post-doctorants du centre INRA de Jouy-en-josas pour les moments passés à oublier un peu la thèse autour de jeux, barbecues et autre activités festives.

Je tiens également à remercier chaleureusement Pascale et Pascal pour m'avoir accueillie chez eux à Jouy-en-josas pendant ma derniére année de thèse ainsi que leurs trois enfants Alice, Marion et Vincent qui m'ont faite me sentir comme en famille.

Je souhaite également remercier mes amis de longue et de courte date qui m'ont suivie pendant toute cette thèse. La liste est trop longue pour tous les citer mais ils se reconnaîtront.

Je remercie toute ma famille pour leurs encouragements, leur affection et surtout leur patience pendant toute cette thèse.

Et je tiens à remercier mon amoureux Al'hossein pour son soutien indeffectible, sa (grande) patience et bien sûr ses bons petits plats.

How do the metabolites, GTP and (p)ppGpp, simultaneously control the occurrence of translational errors and resource allocation in bacteria?

Abstract

Even though diverse mechanisms cooperate to prevent protein synthesis errors in bacteria, missense and Translational Frameshift Errors (TFEs) can occur. In particular, TFEs were detected at low levels in the exponential growth phase and at higher levels in the stationary phase in both *Escherichia coli* and *Bacillus subtilis*. This observation led researchers to revisit the role of the "stringent response" in the occurrence of TFEs since it is the key mechanism involved in the bacterial adaptation to nutritional downshifts. It relies on the interaction between the RelA/SpoT proteins and the translating ribosomes, which leads to the detection of uncharged transfer RNAs (tRNAs) and to the production of an alarmone called (p)ppGpp. In a *relA* mutant strains unable to synthesize (p)ppGpp, translational errors are highly increased.

In this context, the main goal of our work was to revisit the role of the stringent response in the translational error control and to clarify the role of the two key, antagonistic metabolites GTP and (p)ppGpp. Indeed, while GTP enhances translation initiation by targeting the Initiation Factor (IF) IF2, (p)ppGpp inhibits GTP biosynthesis and translation initiation (competing with GTP on IF2).

For this purpose, we used the Gram positive model bacterium *B. subtilis*, designed three distinct reporter systems to detect TFEs and built a strain unable to synthesize (p)ppGpp (called "(p)ppGpp^0"). We observed that during growth in poor media TFEs were increased in the absence of (p)ppGpp in the exponential phase (*i.e.* steady-state growth) and that by contrast to the wild type, the (p)ppGpp^0 strain exhibited a TFE burst during the transition in rich medium to the stationary phase. By controlling intracellular levels of GTP in the (p)ppGpp^0 strain, we showed that GTP abundance is the trigger factor of TFEs occurrence. Nevertheless, upon a "weak" induction of GTP biosynthesis leading to sub-optimal growth rates, the TFEs rate still peaked during the transition to the stationary phase, which demonstrated that the mode of action of (p)ppGpp to prevent TFEs occurrence did not only rely on its inhibition of GTP biosynthesis. We then focused on the (p)ppGpp inhibitory effect on IF2 and mimicked its action by injecting drugs known to inhibit translation initiation. Hence, we demonstrated that by reducing translation initiation (injecting drugs) upon aminoacyl-tRNAs depletion (p)ppGpp^0 strain is able to control the rate of TFEs in the transition to the stationary phase.

In a second part, we studied how transcription and translation are affected by variations in GTP and (p)ppGpp abundances. We observed that genes possessing a Transcription Start Site (TSS) made of two guanines were more importantly transcribed at higher growth rates than genes possessing a

TSS made of two adenines. This difference was even more pronounced for (p)ppGpp^0 strains grown in rich medium upon guanosine addition (leading to a high level of GTP). Moreover, the ribosomal RNA (rrn) for which the TSS is a guanine synthesis level seemed to be positively correlated to GTP levels during exponential growth in poor and rich media as observed by the modulation of GTP biosynthesis.

In conclusion, we demonstrated that (p)ppGpp controls the occurrence of translational errors during steady-state growth by decreasing GTP levels and during a nutritional downshift by specifically inhibiting translation initiation ensuring a parsimonious resource allocation.

Keywords: Translational error ; Stringent response ; GTP and (p)ppGpp; Bacteria; Cell adaptation; Resource allocation

Comprendre comment les métabolites, GTP et (p)ppGpp, contrôlent simultanément l'apparition d'erreurs traductionnelles et l'allocation des ressources chez les bactéries.

Résumé

Bien que divers mécanismes coopèrent pour empêcher les erreurs lors de la synthèse des protéines chez les bactéries, des erreurs traductionnelles de type "frameshift" (ETFs) ou "faux-sens" peuvent avoir lieu. En particulier, les ETFs ont été détectées à de faibles niveaux lors de la phase de croissance exponentielle et à des niveaux plus élevés durant la phase de croissance stationnaire chez *Escherichia coli* et *Bacillus subtilis*. Ces observations ont conduit les chercheurs à revoir le rôle de la "réponse stringente" dans la survenue des ETFs, qui constitue l'un des mécanismes clé de l'adaptation bactérienne aux changements nutritionnels. Elle découle de l'interaction entre un ribosome en cours de traduction et la protéines RelA/SpoT ce qui permet de détecter les ARNs de transfert (ARNts) non chargés et résulte en la production d'une molécule appelée (p)ppGpp . Dans une souche mutante *relA* incapable de synthétiser le (p)ppGpp, les ETFs sont fortement augmentées.

Dans ce contexte, notre objectif principal a été de revisiter le rôle de la réponse stringente dans le contrôle des erreurs traductionnelles et de clarifier le rôle des deux métabolites antagonistes GTP et (p)ppGpp. Par exemple, le GTP stimule l'initiation de la traduction (en ciblant le facteur d'initiation IF2) alors que le (p)ppGpp inhibe l'initiation de la traduction (en rentrant en concurrence avec le GTP pour se fixer sur IF2).

A cette fin, nous avons utilisé le modèle des bactéries à Gram positif *B. subtilis*, conçu trois systèmes rapporteurs distincts pour détecter les ETFs et construit une souche incapable de synthétiser du (p)ppGpp (appelée "(p)ppGpp^0"). Nous avons observé qu'au cours de la croissance dans des milieux pauvres, les ETFs augmentent en l'absence de (p)ppGpp durant la phase exponentielle et que, contrairement à la souche sauvage, la souche (p)ppGpp^0 présente un pic d'ETFs en milieu riche pendant la transition à la phase stationnaire. En contrôlant les niveaux intracellulaires de GTP dans la souche (p)ppGpp^0, nous avons montré que l'abondance de GTP est le facteur qui déclenche l'apparition des ETFs. Néanmoins, après une "faible" induction de la biosynthèse du GTP conduisant à des taux de croissance sous-optimaux, le niveau d'ETFs forme toujours un pic lors de la transition vers la phase stationnaire, ce qui montre que le mode d'action du (p)ppGpp pour prévenir l'apparition des ETFs ne repose pas uniquement sur son action inhibitrice de la biosynthèse du GTP. Nous nous sommes alors concentrès sur l'effet inhibiteur du (p)ppGpp sur IF2 et avons mimé son action en injectant des drogues connues pour inhiber l'initiation de la traduction. Nous avons ainsi démontré qu'en réduisant l'initiation de la traduction lors de l'épuisement des aminoacyl-ARNts, la souche "(p)ppGpp^0" est capable de contrôler de façon optimale le taux d'ETFs lors de la transition vers la phase stationnaire.

Dans une deuxième partie, nous avons étudié comment la transcription et la traduction sont affectées par les variations du niveau de GTP et de (p)ppGpp. Nous avons observé que les gènes possédant un "+1" de transcription (TSS, " transcription start site ") composé de deux guanines (gènes artificiels et ARNs ribosomaux) ont vu leur taux de transcription positivement corrélés au taux de croissance à l'inverse des gènes possédant un TSS composé de deux adénines. Cette différence est encore plus prononcée pour la souche (p)ppGpp^0 cultivée en milieu riche lors de l'ajout de guanosine (ce qui conduit à un niveau élevé de GTP).

En conclusion, nous avons démontré que le (p)ppGpp contrôle le niveau d'erreurs traductionnelles lors de la croissance en régime permanent en abaissant les niveaux de GTP et lors d'un changement nutritionnel en inhibant spécifiquement l'initiation de la traduction, assurant une allocation parcimonieuse des ressources au sein de la bactérie.

Mots clés: Erreur traductionnelle; Réponse stringente, GTP et (p)ppGpp, Bactéries, Adaptation cellulaire, Allocation des ressources

Synthèse en français

Introduction

Le (p)ppGpp est connu depuis des décennies comme un métabolite qui aide les bactéries à s'adapter aux changements nutritionnels. De plus, des travaux menés chez *Escherichia coli* suggèrent qu'il est impliqué dans le contrôle des erreurs de synthèse des protéines. En effet, il est crucial pour la bactérie de maintenir son taux d'erreur de synthèse protéique à un certain niveau : assez faible pour éviter des dommages irréversibles mais relativement élevé pour conférer des avantages adaptatifs à la cellule. Les deux laboratoires impliqués ont mené des recherches sur la façon dont les bactéries s'adaptent aux changements environnementaux au niveau de la réplication, de la transcription et de la traduction. Ils se sont particulièrement intéressés à la façon dont ces processus cellulaires sont affectés pendant la croissance en régime permanent, et comment cela peut se traduire en termes d'allocation des ressources. Dans ce travail, nous défendons le point de vue selon lequel l'allocation des ressources est orchestrée par le GTP et le (p)ppGpp lorsque le milieu de culture change (*i.e.* enrichissement ou appauvrissement en nutriments) mais aussi en régime permanent de croissance. En effet, les régulations achevées par le GTP et le (p)ppGpp peuvent être vues comme des boucles de rétroaction: le GTP favorise la transcription des gènes impliqués dans les mécanismes de traduction et stimule sa propre production (boucle de rétroaction positive); tandis que le (p)ppGpp inhibe l'initiation de la traduction (boucle de rétroaction négative rapide) et la production de GTP, ce qui à terme "freine" la traduction (boucle de rétroaction négative lente).

Lorsque l'enzyme RelA rencontre un ARN de transfert (ARNt) non chargé, il produit du (p)ppGpp à partir du GTP, ce qui inhibe directement la traduction et indirectement la transcription chez *B. subtilis*. En l'absence de RelA, et par conséquent de (p)ppGpp intracellulaire, les ARNt non chargés ne sont pas détectés. En effet, après que des souches d'*E. coli* aient été cultivées dans un milieu privé d'un acide aminé particulier, les réductions des niveaux de chargement de l'ARNt correspondant étaient importantes à la fois pour les souches relA$^+$ et relA$^-$ (*i.e.* souche ne possédant pas l'enzyme RelA suite à la suppression du gène qui la code). Cependant, suite à la déplétion en cet acide aminé, la charge résiduelle des ARNt dans la souche relA$^-$ était inférieure à celle de la souche relA$^+$ [Sørensen, 2001].

Des observations similaires sont attendues pour les souches de *B. subtilis* relA$^+$ (souche qui possède l'enzyme RelA mais pas les enzymes secondaires RelP et RelQ qui ont une activité de synthèse du (p)ppGpp) et (p)ppGpp^0 (souche incapable de produire du (p)ppGpp, obtenues en supprimant les trois gènes codant les (p)ppGpp synthétases RelA, RelP et RelQ). Sorensen (2001) a suggéré que l'observation d'un niveau de charge plus faible des ARNt au sein de la souche relA$^-$ par rapport à la souche relA$^+$ pourrait être responsable de la "misincorporation" (c.- à-d. lorsqu'un acide aminé ne correspondant pas au codon présent dans le site ribosomal A est incorporé au sein de la chaine peptidique pendant la traduction) plus élevée que l'on trouve dans la souche relA$^-$ par rapport à la souche relA$^+$ [O'Farrell, 1978]. Sorensen (2001) explique que le site ribosomal A reste vide pendant une plus longue

période, ce qui augmente la probabilité que le ribosome accepte des ARNt "quasi-correspondants" (*i.e.* dont l'anticodon est proche de celui correspondant à l'ARNt codant pour le triplet présent dans le site A) chargés, ce qui se traduit par un niveau de misincoporation plus élevé dans la souche relA$^-$ que dans la souche relA$^+$.

Le même raisonnement peut également s'appliquer aux erreurs "frameshift" traductionnelles qui correspondent généralement à une translocation de 2 ou 4 paires de bases (c'est- à-dire que le ribosome déplace le cadre de lecture d'une paire de bases en amont ou d'une paire de bases en aval). En effet, plus le ribosome est bloqué longtemps sur un site A vide, plus il est probable qu'une erreur traductionnelle "frameshift" puisse se produire [Urbonavičius et al., 2001][Caliskan et al., 2017]. La "misincorporation" d'un ARNt "quasi-correspondant" dans le site A stimule également l'erreur traductionnelle "frameshift" [Jäger et al., 2013] et les niveaux de "misincorporation" sont plus élevés dans une souche relA$^-$ d'*E. coli* (et donc supposément dans une souche (p)ppGpp^0) [O'Farrell, 1978]. En résumé, on s'attend à ce que les erreurs de traduction ("misincorporation" et erreurs "frameshift") soient plus élevées lorsque les niveaux de chargement des ARNts sont plus faibles et inversement, ce qui signifie que l'observation du niveau d'erreurs de traduction fournirait également des informations sur les niveaux de chargement des ARNts.

Ainsi, l'objectif de ce projet de recherche est de mieux comprendre comment les erreurs traductionnelles, la transcription et la traduction sont affectées par les boucles de régulations GTP/(p)ppGpp lors de changements nutritionnels.

L'alarmone (p)ppGpp permet d'éviter les erreurs traductionnelles "frameshift" par l'intermédiaire de son contrôle rétroactif sur l'initation de la traduction chez *B. subtilis*

Nous avons étudié l'apparition d'erreurs traductionnelles "frameshift" (ETF) pendant la croissance au sein de différents milieux de cultures en perturbant les boucles de régulations GTP/(p)ppGpp. Dans ce contexte, nous avons construit trois systèmes différents qui rapportent les ETFs basé sur la détection de fluorescence et nous les avons inséré au sein des souches sauvages (WT), RelA$^+$ et (p)ppGpp^0; puis nous avons détecté l'évolution des niveaux de fluorescence en utilisant un lecteur de microplaques. Nous avons d'abord observé que les ETFs sont augmentées en l'absence de (p)ppGpp dans les cellules *B. subtilis* lors d'une croissance exponentielle. Puis nous avons modulé le niveau intracellulaire de GTP en insérant le gène codant l'enzyme GuaB (impliquée dans la synthèse d'une précurseur du GTP) sous le contrôle du promoteur hyperspank inductible à l'isopropyl-β-D-1-thiogalactopyranoside (IPTG) (obtention des souches (p)ppGpp^0 *guaB*^{inducible} et RelA$^+$ *guaB*^{inducible}). Le taux des ETFs augmente avec l'induction de la biosynthèse du GTP lors de la croissance en régime permanent de cellules (p)ppGpp^0, mais pas pour les cellules RelA$^+$.

Lors de la croissance en milieu riche CH, le taux des ETFs atteint son maximum lors de la transition vers la phase stationnaire en l'absence des enzymes relP et relQ, un phénomène amplifié de manière significative en l'absence de (p)ppGpp (figure 1). En modulant les niveaux intracellulaires en GTP, nous avons observé que l'abondance du GTP est le facteur déclencheur de l'apparition des ETFs lors de la phase de croissance exponentielle et durant la transition vers la phase stationnaire. Cependant, contenir les ETFs seulement par le biais du contrôle de la biosynthèse du GTP au sein des cellules (p)ppGpp^0 entraîne également une croissance sous-optimale.

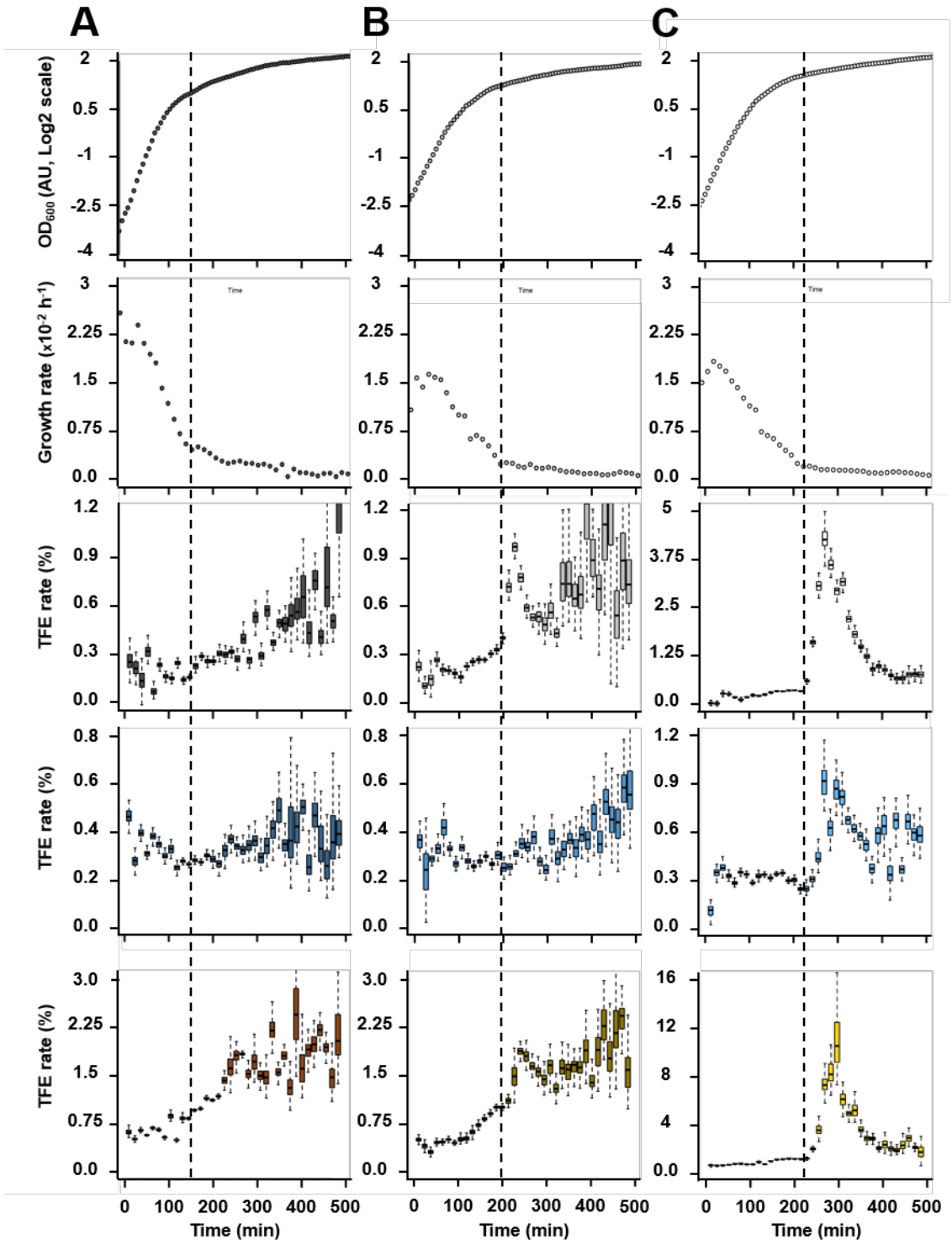

A
B
C
OD600 (AU, Log2 scale)
Growth rate (x10-2 h-1)
TFE rate (%)
TFE rate (%)
TFE rate (%)
Time (min)
Time (min)
Time (min)

Figure 1: Evolution du taux d'ETFs au cours des différentes phases de croissance dans un milieu riche. A, B et C. Le premier panel représente la courbe de croissance ; le second représente le taux de croissance instantané à chaque point dans le temps ; et le troisième, le quatrième et le cinquième panel représentent respectivement les taux d'ETFs de trois systèmes rapporteurs différents. Il faut noter que l'échelle de l'axe des ordonnées est différente entre les trois derniers panels. A. Ces cinq panels représentent la croissance, le taux de croissance instantanée et les taux d'ETFs de la souche WT cultivée en CH. B. Ces cinq panels représentent la croissance, le taux de croissance instantanée et les taux d'ETFs de la souche RelA$^+$ cultivée en CH. C. Ces cinq panels représentent la croissance, le taux de croissance instantanée et les taux d'ETFs de la souche (p)ppGpp^0 cultivée en CH.

Pour finir, nous avons imité le rôle du (p)ppGpp sur le facteur d'initiation de la traduction IF2 en utilisant des antibiotiques connus pour inhiber l'initiation de la traduction et nous les avons injectés avant que les cellules n'entrent dans la phase stationnaire. L'inhibition active de l'initiation de la traduction dans les cellules (p)ppGpp^0 est suffisante pour empêcher de manière optimale l'apparition d'un pic d'ETFs lors de la transition vers la phase stationnaire (figure 2).

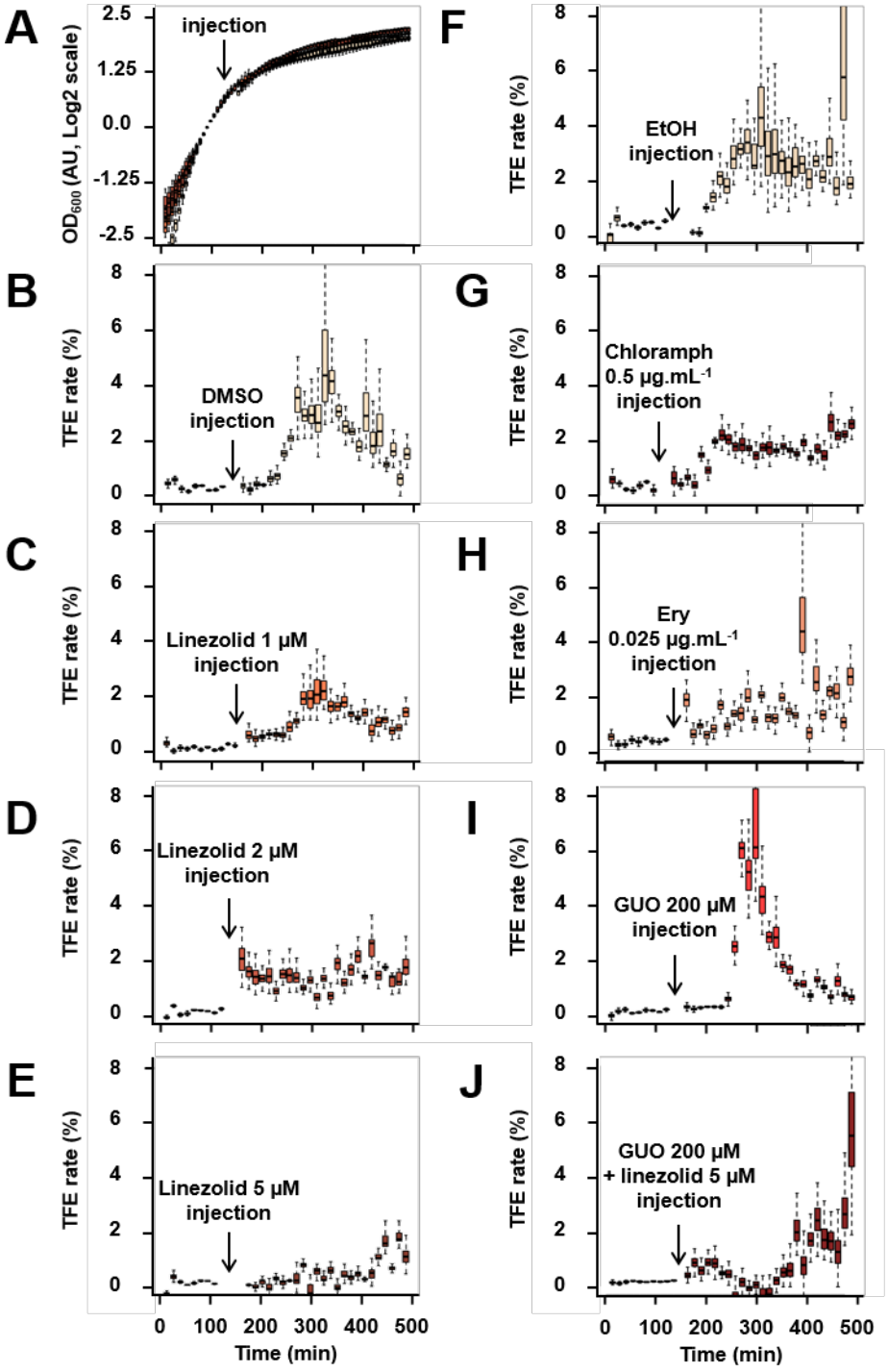

A
B
C
D
E
F
G
H
I
J
injection
DMSO injection
Linezolid 1 µM injection
Linezolid 2 µM injection
Linezolid 5 µM injection
EtOH injection
Chloramph 0.5 µg.mL-1 injection
Ery 0.025 µg.mL-1 injection
GUO 200 µM injection
GUO 200 µM + linezolid 5 µM injection
OD600 (AU, Log2 scale)
TFE rate (%)
Time (min)

Figure 2: Prévention de l'apparition d'un pic d'ETFs suite à l'inhibition de l'initiation de la traduction. A. Courbes de croissance de la souche (p)ppGpp^0 suite à des injections de DMSO ou de linézolide à différentes concentrations finales (1, 2 et 5 μM). B, C, D et E. Taux d'ETFs en fonction du temps lorsque les souches (p)ppGpp^0 et dérivées ont été cultivées en CH où du DMSO a été injecté (B) ou bien du linézolide à des concentrations finales de 1 μM (C), 2 μM (D) et 5 μM (E). F, G et H. Taux d'ETFs en fonction du temps lorsque les souches (p)ppGpp^0 et dérivées ont été cultivées en CH où de l'éthanol a été injecté (F) ; ou bien du chloramphénicol à une concentration finale de 0,5 μg.mL^{-1} (G) ; ou bien de l'érythromycine à une concentration finale de 0,25 μg.mL $^{-1}$(H) I. Taux d'ETFs en fonction du temps où les souches (p)ppGpp^0 et dérivées ont été cultivées en CH où de la guanosine a été injectée à une concentration finale de 200 μM. J. Taux d'ETFs en fonction du temps où les souches (p)ppGpp^0 et dérivées ont été cultivées en CH lors de l'injection simultanée de guanosine (200 μM en concentration finale) et de linézolide (5 μM en concentration finale).

Effets du GTP et du (p)ppGpp sur le processus de transcription et sur la synthèse des ribosomes

Dans une deuxième partie, nous avons étudié comment la perturbation des boucles de régulation GTP/(p)ppGpp affecte le processus de transcription ainsi que la synthèse des ribosomes. Nous avons modulé le niveau intracellulaire de GTP, comme décrit précédemment, afin d'observer les effets des variations de la concentration en GTP sur l'initiation de la transcription de constructions génétiques synthétiques avec des "+1" de transcription (TSS, "transcription start site") de composition différente en nucléotides (guanine versus adénine). Tout d'abord nous avons montré que le TSS influence l'évolution de l'abondance en ARN messagers (ARNm) en fonction du taux de croissance, et qu'en l'absence de (p)ppGpp l'abondance en ARNm est plus élevée pour les gènes possédant des guanines comme TSS (figure 3 A). Ensuite, nous avons observé l'impact du contrôle du niveau intracellulaire en GTP sur l'expression des gènes en fonction de la composition du TSS en milieu pauvre ainsi qu'en milieu riche (figure 3 B).

Ensuite, nous nous sommes intéressés à la façon dont la production de ribosomes est affectée par la dérégulation des boucles GTP/(p)ppGpp en surveillant la synthèse des ARNs ribosomaux (*rrns*). Tout d'abord, le niveau de transcription des *rrns* est affecté par l'absence de (p)ppGpp en milieux pauvres et riches ; et il est particulièrement perturbé par un excès de GTP intracellulaire. De plus, en l'absence de (p)ppGpp intracellulaire, l'expression des *rrns* continue d'augmenter lors d'un appauvrissement nutritionnel du milieu (figure 4). En particulier, pendant la transition vers la phase stationnaire, une abondance élevée en GTP se traduit par des niveaux d'expression des *rrns* élevés, qui varient également en fonction de la nature du promoteur *rrn*.

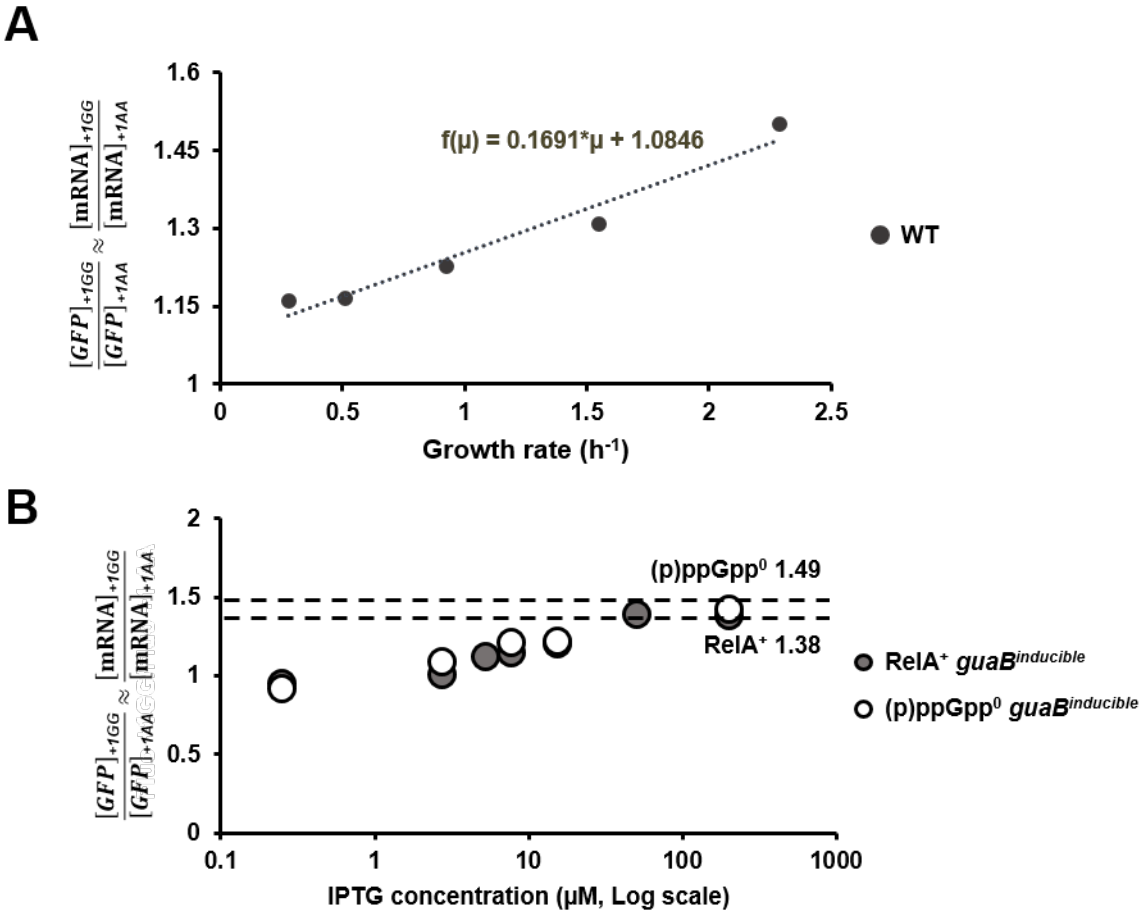


Figure 3: Effets du TSS sur l'expression des ARNm en fonction du taux de croissance et du niveau de synthèse en GTP. A. Ratio d'abondance en GFP $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ de la souche WT en fonction du taux de croissance. La ligne en pointillés correspond à la courbe de tendance du ratio d'abondance en GFP $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ avec le taux de croissance ; et l'équation à côté correspond à la fonction linéaire de cette courbe de tendance où μ désigne le taux de croissance. B. Ratio d'abondance en GFP $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ des souches RelA$^+$ *guaB*^{inducible} et (p)ppGpp^0 *guaB*^{inducible} en fonction de la concentration IPTG. Les lignes pointillées noires correspondent à la valeur de ce rapport pour les souches RelA$^+$ et (p)ppGpp^0 cultivées en CH.

Conclusion

Ces résultats montrent que *B. subtilis* utilise les TSS des promoteurs comme un moyen d'orienter l'ARN polymérase vers les promoteurs des gènes impliqués au sein de processus cellulaires stratégiques et ce en fonction du ratio GTP/ATP, qui est dépendant du taux de croissance. De plus, nous avons observé que la transcription des *rrns* et *in fine* la synthèse des ribosomes sont positivement corrélées au ratio GTP/ATP. Dans l'ensemble, ces résultats suggèrent que perturber les boucles de régulations GTP/(p)ppGpp et ainsi augmenter le ratio GTP/ATP conduit à une réaffectation des ressources vers la machinerie traductionnelle, ce qui est en adéquation avec un taux d'ETFs plus élevé lors de la phase de croissance en régime permanent ainsi que durant la transition vers la phase stationnaire. Cela

démontre également que la production de (p)ppGpp par l'enzyme RelA n'a pas seulement lieu lors de fluctuations environnementales (appauvrissement en nutriments ou autres stress cellulaires) comme le supposait la littérature [Potrykus and Cashel, 2008][Hauryliuk et al., 2015] jusqu' à récemment, mais aussi pendant la croissance en régime permanent (c.- à-d. croissance exponentielle).

De plus, nos résultats suggèrent fortement que : (i) l'action du (p)ppGpp sur IF2 agit comme une boucle rapide de rétroaction négative afin d'adapter l'activité globale de l'appareil de traduction au niveau des ARNt chargés et ainsi prévenir l'apparition d'un pic d'ETFs; et que (ii) l'action inhibitrice du (p)ppGpp sur l'activité des enzymes impliquées dans la biosynthèse du GTP servant à abaisser les niveaux en GTP constitue une boucle de rétroaction lente dont le but est de diminuer la concentration en ribosomes; et ce afin qu'elle corresponde au niveau de chargement en ARNts lors de la phase stationnaire et ainsi maintenir le taux d'ETFs en dessous d'une certaine valeur.

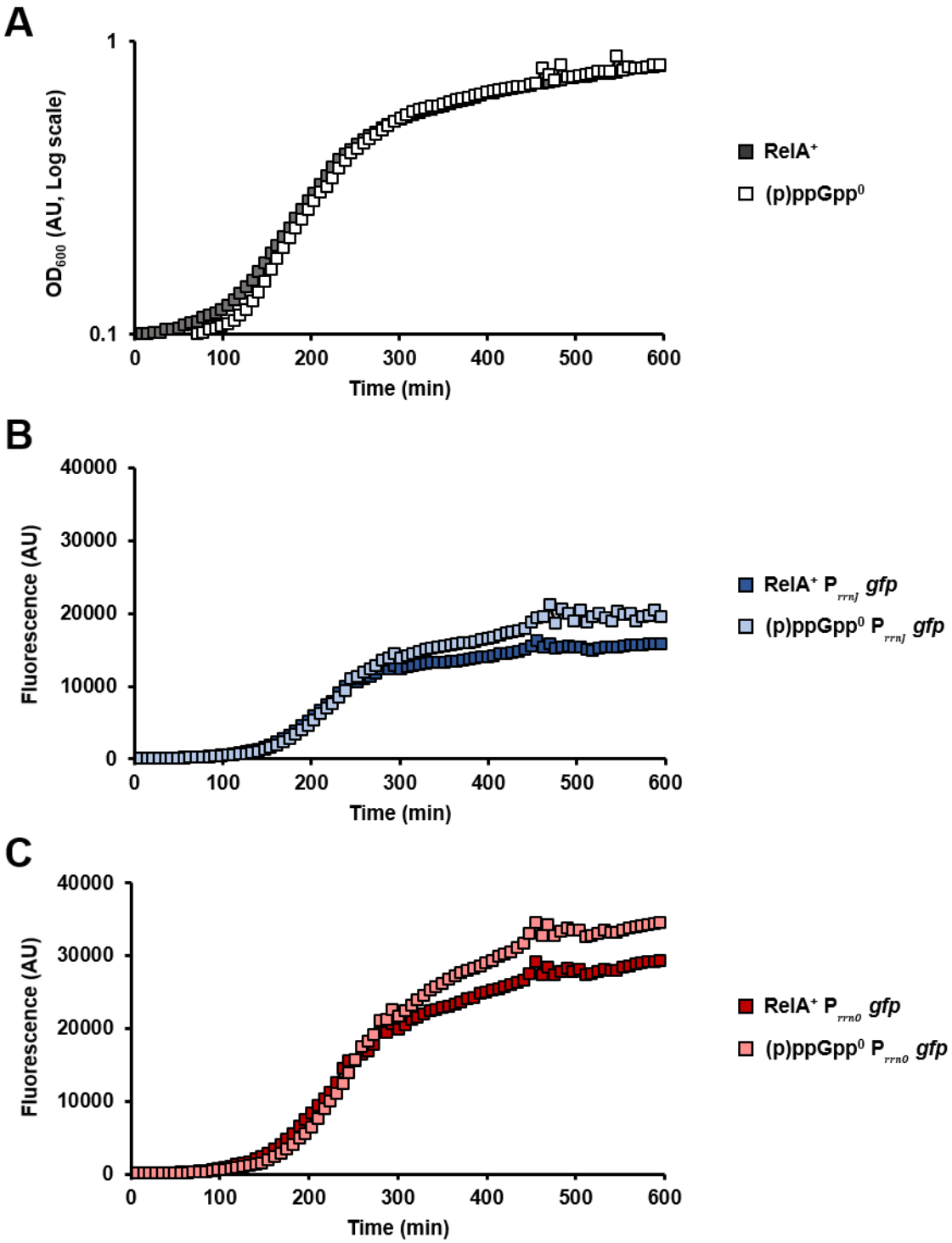


Figure 4: Effets de l'absence en (p)ppGpp intracellulaire sur l'expression des *rrns* lors d'une appauvrissement en nutriments du milieu de croissance. A. Densité optique à 600 nm (OD_{600}) en fonction du temps pour les souches $RelA^+$ et $(p)ppGpp^0$ cultivées en CH. B. Fluorescence en fonction du temps pour les souches $RelA^+$ et $(p)ppGpp^0$ cultivées en CH et contenant le système rapporteur de l'expression de *rrnJ* (P_{rrnJ} *gfp*). C. Fluorescence en fonction du temps pour les souches $RelA^+$ et $(p)ppGpp^0$ cultivées en CH et contenant le système rapporteur de l'expression de *rrnO* (P_{rrnO} *gfp*).

Contents

List of Figures

List of Tables

Abbreviations

aa-tRNA aminoacyl-tRNA

ADF aa-tRNA depletion-stimulated frameshifting

ASL Anticodon Stem and Loop

asRNAs antisense RNAs

BCAAs Branched-Chain Amino Acids

CDS coding sequence

CTD C-terminal domain

CTP cytosine 5'-triphosphate

DNA Deoxyribonucleic Acid

DSL D-arm of the peptidyl-tRNA

EC Elongation Complex

ECF extracytoplasmic function

EFs Elongation Factors

FBA Flux Balance Analysis

GDP guanosine 5'-diphosphate

GFP Green Fluorescent Protein

GMP guanosine 5'-monophosphate

Gpt xanthine-guanine phosphoribosyltransferase

GRAS Generally Regarded As Safe

Gsk Guanosine kinase

GTP Guanosine triphosphate

GUO guanosine

HprT hypoxanthine phosphoribosyltransferase

IF Initiation Factor

IMP inosine 5'-phosphate

IPTG isopropyl-β-D-1-thiogalactopyranoside

LB Lysogeny Broth

LIC Ligation-Independent Cloning

MTF Methionyl tRNA Transformylase

Ndk NDP kinase

NTD N-terminal domain

NTP Nucleoside TriPhosphate

OD Optical Density

PA Promoter Activity

PCR Polymerase Chain Reaction

polyP polyphosphate

ppGpp guanosine tetraphosphate

pppGpp guanosine pentaphosphate

PRF Programmed Translational Frameshift

PRPP 5-phospho-α-D-ribose 1-diphosphate

Prs phosphoribosylpyrophosphate synthetase

PTC peptidyltransferase center

RBA Resource Balance Analysis

RBS Ribosome Binding Site

RFs Release Factors

RNAP RNA Polymerase

RNAs Ribonucleic Acids

RRF Ribosome Recycling Factor

rrn ribosomal RNA

RSH RelA/SpoT Homology

S Svedberg

SAS Small Alarmone Synthetases

SD Shine-Dalgarno

SRL Sarcin-Ricin Loop

TFEs Translational Frameshift Errors

TFs Transcription Factors

TIR Translation Initiation Region

tRNAs transfer RNAs

TSS Transcription Start Site

UTR untranslated region

WT wild-type

XMP xanthosine 5'-phosphate

Part I

Introduction

Chapter 1

Biological background on bacteria

Background

1.1 *Bacillus subtilis*

1.1.1 General description

Bacillus subtilis belongs to the *Firmicute* phylum. The main classes of this phylum are the *Clostridia*, *Bacilli* (*Bacillus*, *Listeria*, *Enterococcus*, *Staphylococcus* and *Streptococcus*) and *Mollicute* [Wolf et al., 2004]. Most species from this phylum exhibit a low G+C content in their Deoxyribonucleic Acid (DNA) and are Gram-positive bacteria (except for *Negativicutes*). While the Gram-negative bacteria possess an inner and outer membrane flanking their cell wall within the periplasmic space, the Gram-positive bacteria possess only a single membrane at the internal face of the peptidoglycan and thus lack a true periplasmic space (except for mollicutes bacteria) [Sonenshein et al., 2002].

Gram-positive and Gram-negative bacteria diverged about 2 billion years ago but, the two respective model organisms, *B. subtilis* and *Escherichia coli*, exhibit similarities in their genome sequences [Feng et al., 1997]. About 1000 *B. subtilis* genes have clear orthologous counterparts in *E. coli* (one-quarter of the genome) [Kunst et al., 1997] [Lawrence and Ochman, 1998]. Operons that code for the core of the translation, transcription machineries and for the major integrated functions such as ATP synthesis (*atp* operon) and electron transfer machinery (*cta* and *qox* operons) are well conserved [Kunst et al., 1997].

1.1.2 *B. subtilis* lifestyle

B. subtilis is a ground living bacterium which can grow between 10°C and 55°C with an optimal growth rate in a range of temperature between 35°C and 45°C [Ratkowsky et al., 1983]. *B. subtilis* is referred to a rhizosphere bacterium which benefits from components excreted by plant roots, such as amino acids, organic acids, sugars, phenolics acids and other secondary metabolites [Earl et al.,

2008]. In turn, *B. subtilis* contributes to the plant defense against pathogenic bacteria [Bais et al., 2006] [Timmusk and Wagner, 1999].

In such ecosystem, *B. subtilis* exhibits several strategies of growth adaptation to changing conditions. At 37°C, in rich medium, such as Lysogeny Broth (LB) medium [Bertani, 2004], it exhibits a growth rate of 1.8 h^{-1} (*i.e.* a doubling time of $\simeq$ 23 min). At 37°C, in poor conditions, such as a minimal medium supplemented with pyruvate as carbon source, the growth rate can decrease down to 0.2 h^{-1} (*i.e.* a doubling time of $\simeq$ 200 min) [Kleijn et al., 2010].

When an essential nutrient becomes limiting, different adaptation processes can occur:

(i) *B. subtilis* can sporulate by forming a small, tough, protective and metabolically dormant endospore [Stephens, 1998]. Spores can survive years in the absence of nutrients, and are resistant to desiccation, very high and low temperatures and even radiations. The spore formation also activates intracellular mechanisms that make sister cells lysing and thus allow the sporulating bacterium to feed on the nutrients thereby released [González-Pastor et al., 2003]. This phenomenon is referred to as 'cannibalism'.

(ii) *B. subtilis* cells can also become genetically competent which allows cells to scavenge the DNA present in the environment [Anagnostopoulos and Spizizen, 1961]. Hence exogenous DNA can be uptaken and used as a nutrient source or, when possible, integrated into the chromosome by recombination to gain new functions.

(iii) *B. subtilis* is also able to form biofilms that are resistant structures against environmental stresses and otherwise bactericidal compounds such as antibiotics and disinfectants [Russell, 2004].

(iv) Amino-acid starvation triggers the stringent response in *B. subtilis* (as in *E. coli*) which entails pleiotropic responses leading to adaptation of cell growth to nutrient depletion. This stress response is activated via the guanosine tetraphosphate (ppGpp) or guanosine pentaphosphate (pppGpp) alarmone (together referred as (p)ppGpp) and the ribosome-associated protein RelA (described in details in part 2.3.1) [Wendrich and Marahiel, 1997].

1.1.3 *B. subtilis*: an important model organism

Along with *E. coli* and *Salmonella typhimurium*, *B. subtilis* is one of the most studied bacteria. *B. subtilis* was used as a model organism to study bacterial physiology [Monod, 1949], to analyze metabolism, gene regulation, differentiation, sporulation and protein secretion in bacteria. It is a Generally Regarded As Safe (GRAS) organism. The tryptophan auxotroph *B. subtilis* strain 168, used by Spizizen (1958) to identify and optimize the conditions for efficient transformation in the laboratory [Spizizen, 1958], was entirely sequenced in 1997 [Kunst et al., 1997] and again in 2009 [Barbe et al., 2009]. Over 4000 genes were annotated and around 260 were found to be essential for growth [Kobayashi et al., 2003]. Two recent studies found respectively that 261 genes coding for 259 proteins and two functional Ribonucleic Acids (RNAs) [Commichau et al., 2013] or 257 genes [Koo et al., 2017] are essential for growth. In both studies, the largest group of essential proteins correspond to proteins involved in protein synthesis (*i.e.* ribosome synthesis, translation,...), secretion and protein

quality control [Commichau et al., 2013] [Koo et al., 2017]. An other important set of essential proteins concern lipid biosynthesis, cell wall metabolism, cell division, and DNA replication [Commichau et al., 2013][Koo et al., 2017]. A third group corresponds to proteins involved in protecting the cell against endogenous toxic proteins, metabolites, or other intermediate compounds [Commichau et al., 2013] [Koo et al., 2017].

Several molecular biology experimental tools have been optimized for *B. subtilis* and a lot of dedicated, reliable databases are available: genome databases (Subtilist[1], GenoList [Lechat et al., 2007]), a database of transcriptional regulation in *B. subtilis* DBTBS[2] [Nicolas et al., 2012]), proteome databases (KEGG[3] [Kanehisa et al., 2016], UniProt[4] [Consortium, 2016], PDB[5] [Burley et al., 2017]) and also a highly detailed database pulling all informations from the specific databases of genes, messenger RNAs and protein expression and regulation (SubtiWiki[6] [Zhu and Stülke, 2017]).

1.1.4 *B. subtilis*: a bacterium of industrial interest

B. subtilis and other *Bacilli* are of interest for food, pharmaceutical or chemical companies. The ability of *Bacilli* strains to secrete extracellular enzymes has placed *Bacilli* among the industrial enzyme producers for a long time. Industrial enzymes produced in large amounts by engineered *Bacilli* species are amylases and proteases [Schallmey et al., 2004]. Industrial enzymes of interest from *Bacilli* species are used in different industries: household cleaning (alkaline proteases, alkaline amylase), textile (amylases, glucose isomerase), baking (amylase) and beverage industries (amylase, glucanase) [van Dijl and Hecker, 2013]. *Bacilli* produce also various classes of antibiotics such as polymyxins (break up the bacterial cell membrane), subtilin (sporostatic activity), mycobacillin (antifungal cyclic peptide), bacitracin (inhibits cell wall synthesis) or butirosin (affects ribosome function) [van Dijl and Hecker, 2013]. Vitamins (e.g. riboflavin), flavor enhancers (e.g. purine nucleotides) and insecticides (such as endotoxins endogeneous to *Bacillus thuringiensis*) are also industrially produced using *Bacilli* species [van Dijl and Hecker, 2013]. Moreover, *B. subtilis* is traditionally used in Japan in food fermentation for the production of *natto*. The most important enzymes in the production of *natto* are proteases responsible of the soybean protein hydrolysis.

Overview: *B. subtilis* is an ubiquist, importantly studied gram-positive bacterium which, through evolution, developed several adaptive strategies under stress conditions. It has been studied for decades by a large community and is used for a broad range of research and industrial applications, which makes it an important model microorganism.

[1] http://genolist.pasteur.fr/SubtiList/
[2] dbtbs.hgc.jp/
[3] www.genome.jp/kegg/
[4] www.uniprot.org/
[5] https://www.rcsb.org/
[6] subtiwiki.uni-goettingen.de/

1.2 The fundamental cellular processes of gene expression in bacteria

Bacterial growth occurs via an iterative process of active synthesis of cellular biomolecules, increasing the size of the cell, followed by the division of the cell into two daughter cells. This process requires biosynthesis and equal transmission of DNA and cellular components between the two daughter cells. Most of the molecular knowledge about DNA replication, RNA synthesis and translation have been acquired in *E. coli*. This knowledge will be described in the following part. When they have been established, similarities and differences with *B. subtilis* will be highlighted.

1.2.1 The DNA replication

Most of bacteria replicate their entire DNA (a single circular chromosome) so they can then divide and pursue growth. The DNA replication, which is bidirectional, starts at the origin of replication (*oriC*) and ends at the terminus (*ter*) site (Figure 1.1). This cellular process is carried out by multiple proteins and consists in three steps: initiation, elongation and termination.

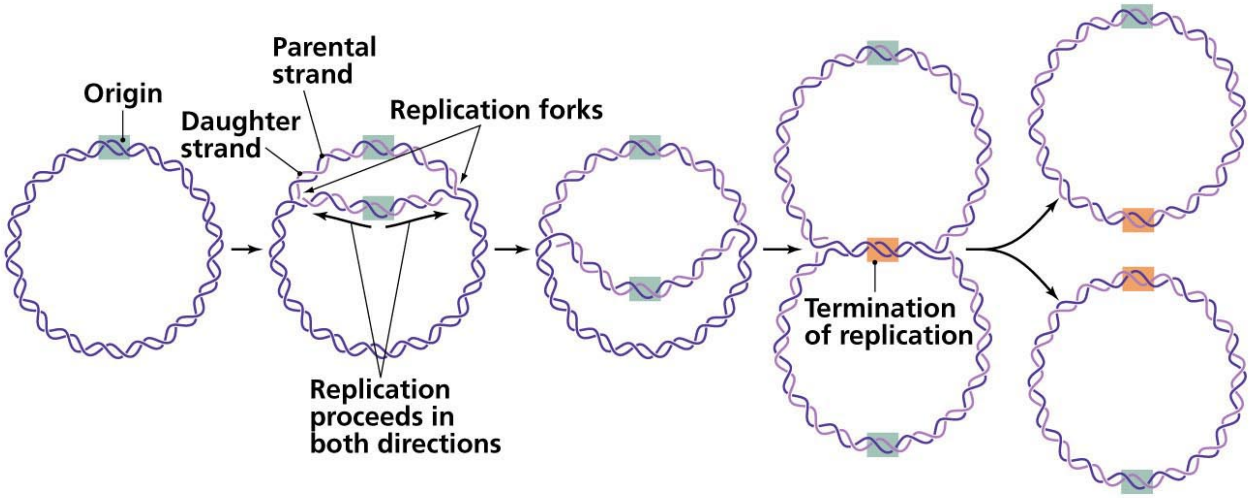


Figure 1.1: DNA replication in bacteria. (Adapted from [Tortora et al., 2004]) The DNA replication occurs in three steps: the initiation at the origin; the elongation that requires the formation of replication forks and which proceeds in both direction; and the termination which takes place approximately at the opposite side of the origin of replication. This leads to the formation of a second chromosome.

1.2.1.1 Initiation of DNA replication

DNA replication is initiated at the *oriC* region which is targeted by the initiator protein DnaA [Berg et al., 2002] [Fukuoka et al., 1990]. During DNA replication initiation, the initiator protein recruits other proteins onto the origin of replication *oriC* to form the pre-replication complex which unzips

the double-stranded DNA (Figure 1.1). The unzipping process continues to take place during the elongation process and requires the DNA helicase enzyme. These enzymes create a replication fork which is a structure that has two branching "prongs", each one made up of a single DNA strand.

1.2.1.2 Elongation of DNA replication

Once the two DNA strands are separated, the primase DnaG adds RNA primers to initiate replication [Jameson and Wilkinson, 2017] (Figure 1.2). The process is slightly different for each strand: there is the leading strand which receives only one RNA primer while the lagging strand receives several. The DNA polymerase DNA Pol. III then proceeds to replication on the leading strand while the lagging strand is extended discontinuously from each primer forming Okazaki fragments [Kornberg and Baker, 1992][Dervyn et al., 2001](Figure 1.2). Then, an RNase removes the RNA primers and another DNA polymerase DNA Pol. I enters to fill in the gaps previously filled in with RNA primers and the DNA ligase ensures the junction between the Okazaki fragments [Kornberg and Baker, 1992] (Figure 1.2). Other DNA polymerases are involved in prokaryotic DNA replications and possess like DNA Pol. I and III a proofreading activity to avoid misincorporation of nucleotides resulting in DNA mismatches [Kornberg and Baker, 1992]. The observed rate of spontaneous mutations in *E. coli* is approximately 10^{-10} mutations/base pair/chromosome duplication [Schaaper, 1993].

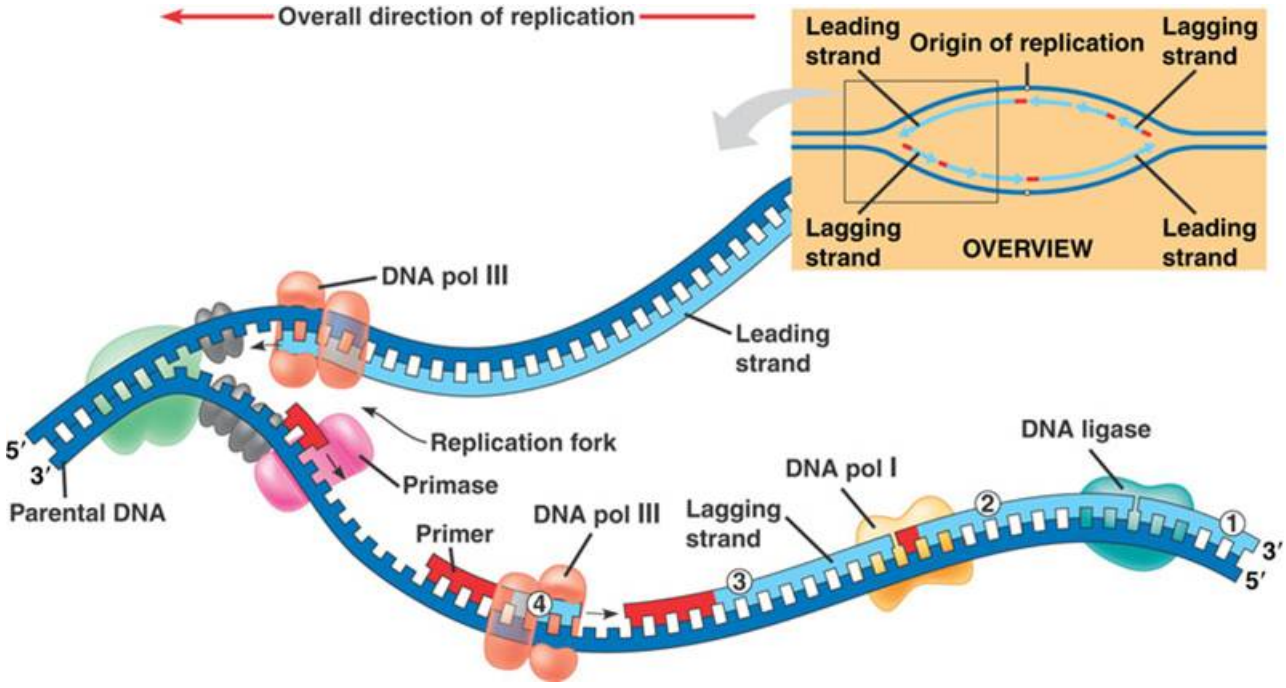


Figure 1.2: Elongation of DNA replication in bacteria. (Adapted from [Tortora et al., 2004]) The DNA replication involves the action of different proteins on the DNA strands and Okazaki fragments such as the primase, the DNA Pol I and III, the DNA ligase.

1.2.1.3 Termination of DNA replication

The termination of DNA replication occurs when the replication machineries that started from both sides of the *oriC* reach the *ter* sites, located approximately at the opposite of *oriC* (Figure 1.1). The termination is made possible by the action of the DNA replication terminus site-binding protein Ter. The binding of Ter at the *ter* sites will prevent the passage of replication forks. Then, the terminus utilization substance (Tus) protein will halt the DNA polymerase movement.

Overview: DNA replication in bacteria consists in three main steps (initiation, elongation and termination) which require the action of several proteins, in particular: the initiator protein DnaA and the DNA helicase during initiation; the primase DnaG, the DNA polymerases I and III and the DNA ligases during elongation; the Ter and the Tus protein during termination. DNA replication is a tightly controlled cellular process concerning the appearance of errors in the copied DNA fragments, leading to a low rate of misincorporated nucleotides ($\simeq 10^{-10}$ mutations/base pair/chromosome duplication).

1.2.2 Transcription

To produce proteins, the first step is to transcribe DNA fragments corresponding to the genes encoding the proteins that the cell requires. This step is carried out by the RNA Polymerase (RNAP) holoenzyme and consists in three steps: initiation, elongation and termination. The rate of transcription errors in *E. coli* is estimated to be in the range of 10^{-4}-10^{-5} mutations/nucleotide and seems to be similar among bacterial species [Traverse and Ochman, 2016].

1.2.2.1 The RNA polymerase: core subunit and σ-factors

The RNAP is the enzyme required for transcription and it is composed of three subunits that form the core subunit in both *E. coli* and *B. subtilis*. These subunits are the α, β and β' subunits encoded by *rpoA*, *rpoB* and *rpoC*, respectively [Burgess and Mach, 1971]. The complex exhibits a stoichiometry of two α subunits for one β subunit and one β' subunit (Figure 1.3) [Boor et al., 1995] [Suh et al., 1986]. The α subunit performs three functions: (i) it is the initiator for RNAP assembly (it acts as a scaffold for the β and β' subunits), (ii) it participates in promoter recognition by sequence-specific protein-DNA interaction and (iii) it is the target for transcriptional regulators [Ebright and Busby, 1995] [Vassylyeva et al., 2002]. The catalytic site for RNA synthesis is located on the β subunit [Glass et al., 1986][Murakami, 2015]. The β subunit is also involved in the recognition of the promoter sequence [Nomura et al., 1984]. Contacts between the β' subunit and the downstream DNA have been found to stabilize the elongating complex [Nudler et al., 1996].

Although it is catalytically active, the core enzyme is unable to specifically initiate transcription [Vassylyeva et al., 2002] in both *E. coli* and *B. subtilis*. Specific initiation is determined by sigma (σ) factors which recognize the promoter elements and initiate transcription at these sites (Figure

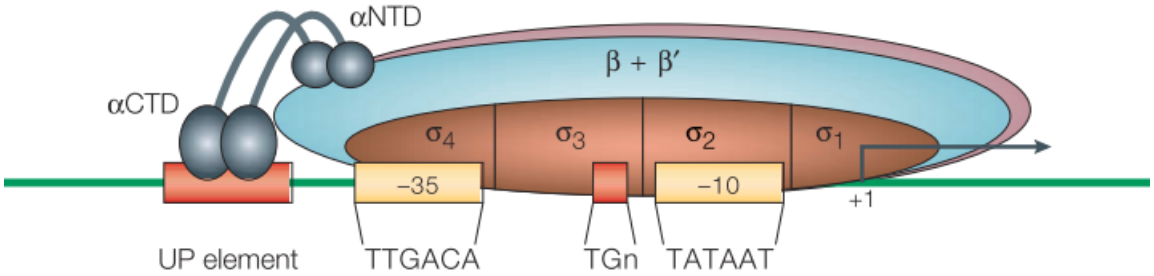


Figure 1.3: The RNA polymerase. (from [Browning and Busby, 2004]) The chromosome (or DNA) is in green, with the -10 and -35 boxes highlighted in yellow and the UP elements and TGn highlighted in red. The RNAP is shown with the β and β' subunits colored in light blue and pink, respectively. The α subunits (CTD and NTD) are colored in grey.

1.3). The complex composed of the core subunits and of a σ factor is called the holoenzyme (RNAP). Several σ factors have been identified in *E. coli* (the "housekeeping" σ^{70} and six alternative σ factors) [Cho et al., 2014] and in *B. subtilis* (the "housekeeping" σ^{A} and eighteen alternative σ factors) [Zhu and Stülke, 2017]. σ^{A} and σ^{70} factors are always expressed while the other σ factors are expressed and/or activated in response to specific environmental stimuli such as stress, entry into sporulation, etc. The primary "housekeeping" sigma factor is σ^{A} in *B. subtilis*, [Juang and Helmann, 1994] and σ^{70} in *E. coli*, and both share similar transcription specificities even though the holoenzymes are not identical between these two species [Haldenwang, 1995][Browning and Busby, 2016].
The interaction between the σ factor and the core subunit leads to a conformational change by facilitating the unwinding of the DNA duplex near the transcription start site. σ^{A} is composed of four domains joined by linkers (Figure 1.3). Domains 2, 3 and 4 are known to be involved in promoter recognition [Browning and Busby, 2004] and the domain 1 accelerates the formation of the open complex for certain promoters in *E. coli* [Murakami and Darst, 2003].
The housekeeping and alternative sigma factors bind to the same site on the surface of the RNAP core enzyme but, under most conditions, the housekeeping sigma factor is more abundant and thus able to outcompete alternative sigma factors [Browning and Busby, 2016]. However, when the abundance of an alternative sigma factor increases, it can then compete with and displace the housekeeping sigma factor to reprogramme a subset of RNAP molecules [Browning and Busby, 2016]. Nearly all alternative sigma factors are evolutionarily related to housekeeping sigma factors, consisting of two, three or four domains [Browning and Busby, 2016]. In *B. subtilis*, a recent work revisited the dynamic of σ factors on the RNAP and showed that they are not activated at constant levels but through repetitive pulsing [Park et al., 2018].

In *B. subtilis*, three accessory subunits are closely associated with the core subunits [Mukherjee et al., 1999]. The ω subunit, the δ subunit and the ε subunit are respectively encoded by *yloH*, *rpoE* and *ykzG* [Doherty et al., 2010]. The ω protein exhibits a structural role in the maintenance of the conformation of the β subunit and in the recruitment of the β' subunit [Mathew and Chatterji, 2006]. The δ protein

operates together with the σ factor as an initiation subunit of RNAP [Juang and Helmann, 1994] and the function of the ε subunit remains unclear. The UP element which is a sequence modulating the RNAP activity (Figure 1.3) is specific to *E. coli* and these sequences have not been yet identified in *B. subtilis*.

1.2.2.2 Transcription initiation

Transcription initiation can be divided into three reversible steps and one irreversible step. The first step is the reversible initial specific binding: the holoenzyme binds to the DNA and forms the "closed" complex (the two DNA strands remain hybridized). The second step is the reversible conformational changes of the "closed" complex: it forms the final binary open complex. The two DNA strands open around the transcription start site (TSS) also referred as +1 position. The complex is poised to bind the initiating nucleotide. During this slow process, the DNA sequence opens up along with a conformational change in RNAP. The third step is the reversible binding of initiating ribonucleotides: the first complementary Nucleoside TriPhosphate (NTP) binds at the +1 position and forms a ternary "initiated" complex. Alternatively, the "initiated" complex may revert by releasing the nascent RNA chain. This leads to an abortive initiation. The RNAP engages multiples abortive cycles of synthesis and releases short products at this step [Margeat et al., 2006]. The fourth and last step is the transition to elongation: the σ factor is released after the polymerization of 7 to 12 ribonucleotides. At this step, the RNAP does not interact anymore with the promoter sequence and elongation can proceed. Afterward, the binding between the enzyme and the DNA template is very stable.

The transcription initiation can be enhanced or repressed by specific DNA binding proteins, the Transcription Factors (TFs). There exists more than 200 in *B. subtilis* [Moreno-Campuzano et al., 2006]. Contrary to the σ factors, most of the TFs exhibit very few promoter targets. However, there exists a few "general" regulators in *B. subtilis* that have ten to more than one hundred targets, such as CodY, which repress the expression of genes induced during the transition from exponential growth to stationary phase and sporulation [Sonenshein, 2005]; Spo0A which regulates the expression of the genes involved in the initiation of sporulation [De Hoon et al., 2010]; ComK which regulates the transcription of genes required for the expression of the late competence genes [Hamoen et al., 2003] or CcpA which is the master regulator of the catabolite repression, which is the regulatory mechanism that allows the cells to choose among several available carbon sources [Henkin, 1996][Meyer et al., 2011].

1.2.2.3 Transcription elongation

In bacteria, elongation is composed of three steps [Vassylyev et al., 2007]: (i) the binding of the complementary NTP, (ii) the reaction of the RNA chain 3'-OH with the NTP (catalyzed by a pair of bound Mg^{2+}) and (iii) the translocation of the NTP assembly. Despite the high stability of the

complex, various interactions allow lateral mobility of DNA and RNA through the complex during translocation. Moreover, the RNAP is known to exhibit spontaneous backtracking along the DNA which could modify the elongation rate. Backtracking is reduced if more than one RNA polymerase molecule initiates from the same promoter [Epshtein and Nudler, 2003].

1.2.2.4 Transcription termination

The termination process is controlled by many factors and two main mechanisms have been identified, especially in *E. coli* and *B. subtilis*.

First, the termination can take place in presence of a terminator sequence which possesses two essential components: (i) a GC-rich inverted-repeat sequence (about 9 nucleotides upstream of the messenger RNA release site) and (ii) an adjacent U-rich segment [Peters et al., 2011]. The nucleotide composition of the stem-loop sequence affects the stability of the resulting RNA structure and thus the termination efficiency. The particular stem-loop conformation of the RNA triggers destabilization of the elongation complex and allows termination. This phase includes a pausing step (induced by the U-rich segment) which is required for efficient termination. Then the hairpin nucleation step is followed by the Elongation Complex (EC) disruption (Figure 1.4). The speed of formation of the hairpin loop can be as fast as microseconds. It depends on the sequence and on the formation of other RNA secondary structures which could compete with the terminator stem-loop formation.
In the first step of termination, a transcriptional pause stops nucleotide addition which allows the terminator hairpin to form [Gusarov and Nudler, 1999]. In the next step, the elongation complex is disrupted to favour dissociation. The termination can occur through three pathways: (i) the hairpin shearing, (ii) the hypertranslocation or (iii) the hairpin invasion. In the hairpin shearing, the hairpin pulls the RNA out of the elongation complex without forward translocation of the RNAP. In hypertranslocation, formation of a hairpin pushes the RNAP forward. In the hairpin invasion, the hairpin induces a conformational change in the RNAP which disturbs the elongation complex (Figure 1.4) [Wang and Greene, 2011].

The second mechanism requires the homohexameric ring protein Rho which is an RNA helicase [De Hoon et al., 2005] [Peters et al., 2011]. Rho binds to a C-rich segment of RNA, and then it dissociates RNAP from RNA which is coupled to ATP hydrolysis [Richardson, 2003]. Termination efficiency depends on the competition between Rho translocation and the RNA chain elongation. Pausing in the elongation phase increases the efficiency of Rho termination [Jin et al., 1992]. Unlike in *E. coli*, the Rho protein is dispensable in *B. subtilis*, suggesting a limited role for Rho-dependent termination in this organism [Ingham et al., 1999]. However, in a mutant strain lacking the termination factor Rho, the mRNA extensions reached up to 12 kb (the average is approximately of 2.8 kb). Without Rho, additional antisense RNAs (asRNAs) are formed by extension of a subset of transcription units. Many of these sRNAs have only partially efficient intrinsic terminators, indicating that

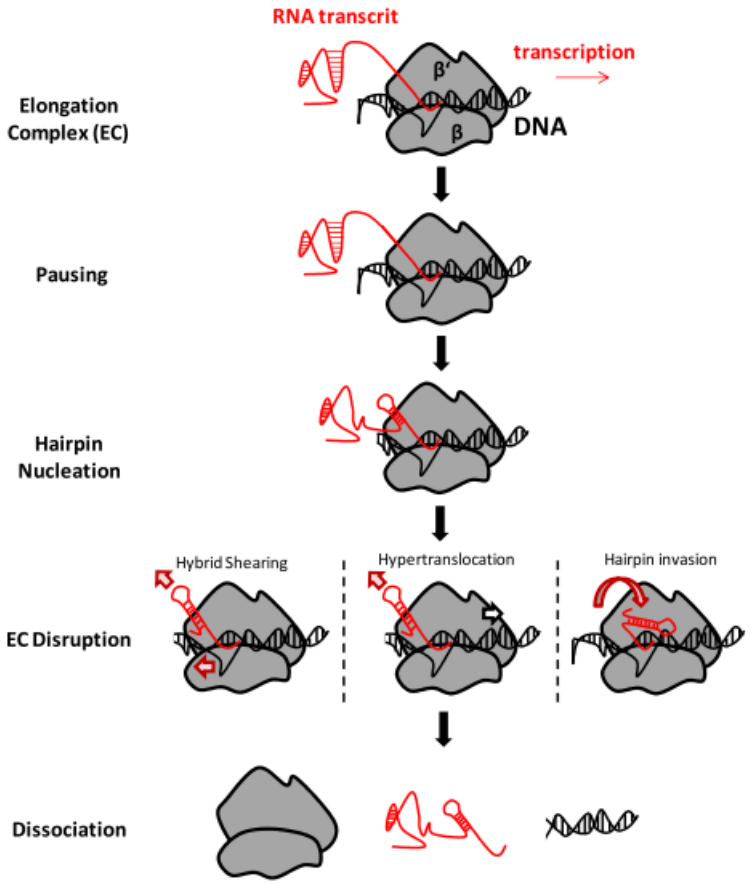


Figure 1.4: Model of the mechanisms of intrinsic termination. (modified from [Peters et al., 2011]) The elongation complex (EC) is composed of the β and β' subunits (grey), and the DNA (black). The mRNA is in red. Three alternative routes to elongation complex disruption by hairpin completion are depicted: the hybrid shearing, the hypertranslocation and the hairpin invasion.

Rho is a general inhibitor of antisense transcription [Nicolas et al., 2012]. Moreover, Rho is involved in bacterial decision making such as cell motility, biofilm formation, and sporulation [Bidnenko et al., 2017].

Overview: Transcription in bacteria consists in three main steps: initiation, elongation and termination. It is triggered by the action of the RNA polymerase (RNAP) holoenzyme which is composed of α and β subunits and a σ factor. The σ factor is mainly the "housekeeping" σ factor σ^{70} (*E. coli*) or σ^{A} (*B. subtilis*) but alternative σ factors exist and respond to different specific environmental stimuli. The transcription starts at a specific nucleotide, the transcription start site (TSS) also referred as +1 position. The transcription initiation can be regulated by TFs which bind to the promoter DNA sequence in order to enhance or repress the gene expression. Then the RNAP proceeds to the elongation step by adding the nucleoside triphosphate (NTP) complementary to the nucleotide of the DNA strand. Transcription termination occurs by two main mechanisms: one requires the presence of a terminator sequence which has a stem-loop conformation; the other one requires the homohexameric ring protein Rho which is linked to cell decision making.

1.2.3 Translation

Translation is the last step of the central dogma [Crick, 1970]. This is the step where the genetic information will lead to the building of a complete protein made of amino acids linked by peptide bonds. This step is carried out by the ribosome and consists in three steps: initiation, elongation and termination.

1.2.3.1 The Ribosome

Bacteria contain two ribosomal subunits sedimenting at 30 Svedberg (S) and at 50S. The 30S subunit is composed of 21 ribosomal proteins in *E. coli* [Henkin, 2002] [Weber, 1972] and 20 ribosomal proteins in *B. subtilis* [Roberts and Rabinowitz, 1989]. The 50S ribosomal subunit is composed of more than 30 proteins (34 proteins in *E. coli* [Henkin, 2002] and 33 proteins in *B. subtilis* [Barbe et al., 2009]). It consists of a rounded base with three protuberances called the L1 protuberance, the central protuberance, and the L7/L12 stalk. The 30S subunit contains a ribosomal RNA (rRNA) sedimenting at 16S and the 50S subunit contains two RNAs sedimenting at 5S and 23S [Laursen et al., 2005]. A tunnel starts at the peptidyltransferase center (PTC), where the formation of peptide bonds occurs. The genes coding for the multiple proteins and rRNA which compose the 30S and 50S subunits are organized into operons. *E. coli* and *B. subtilis* have respectively 7 and 10 rRNA operons which encode the 16S, 23S and 5S rRNAs [Boros et al., 1979] [Stewart et al., 1982]. In *E. coli*, there are 19 ribosomal protein operons which encode proteins present in the 30S and the 50S subunits [Kaczanowska and Rydén-Aulin, 2007]. In *B. subtilis*, 21 out of the 53 genes which encode ribosomal proteins are gathered in a single operon [Li et al., 1997b]. 13 other genes encoding ribosomal proteins are in the same region of the chromosome and the other ones are scattered around the chromosome [Henkin, 2002]. Most of the genes that encode ribosomal proteins are present as a single copy.

1.2.3.2 Translation initiation

Translation starts at a specific mRNA sequence which is referred to as the Translation Initiation Region (TIR) and corresponds to the sequence between the transcription start site (+1 position) and the start codon (Figure 1.5). AUG is the most common initiation codon and is present in 82,6% of *E. coli* genes [Blattner et al., 1997] and in 78 % of *B. subtilis* genes [Rocha et al., 1999]. The Shine-Dalgarno (SD) sequence or Ribosome Binding Site (RBS) is included in the TIR and it is usually located around 8 bases upstream of the start codon AUG (Figure 1.5)[Malys, 2012]. Its sequence base pairs with the anti-SD sequence GAUCACCUCCUUA localized in the 3' end of 16S rRNA [Malys, 2012]. TIRs that lead to the most efficient translation in *B. subtilis* exhibit a long U-rich sequence in the untranslated region (UTR) of the mRNA that acts as an enhancer of translation [Makrides, 1996].

Figure 1.5: Structure of a transcription unit. The gene sequence is generally composed of different parts: the promoter, the translation initiation region (TIR) and the ORF region. The promoter is required to allow transcription initiation. It is composed of the -35 box and the -10 or TATAAT box which are respectively 35 and 10 base pairs upstream of the TSS (approximately). The TIR is required for initiating translation and it starts from the TSS until the start codon (ATG usually). It also comprises the Shine-Delgarno sequence (or ribosome binding site) generally located around 8 bases upstream of the start codon. The TIR corresponds to the 5' untranslated region (5'-UTR) minus the start codon. The ORF corresponds to the sequence that will be translated into a protein by the ribosome.

Translation initiation occurs in three steps [Ramakrishnan, 2002] (Figure 1.6). First, the initiation complex binds on the mRNA. The ribosome binds to single-stranded regions of the mRNA, exploiting the base complementarity between the Shine-Dalgarno sequence of the mRNA and the anti-Shine-Dalgarno sequence present in the 16S rRNA which serves as a guide. Then, the ribosome accommodates onto the start codon. In a third step, the 50S subunit associates with the 30S subunit. The recognition of the start codon by the fMet-tRNA^{fMet} triggers conformational changes that convert the 30S to a functionally competent 30S initiation complex [Milón et al., 2012]. Then, the 30S subunit is joined by the large (50S) ribosomal subunit which is poised for translation of the selected mRNA. The first two steps are reversible but the association of the 50S subunit with the 30S subunit is considered to be irreversible, and results in the formation of the 70S complex [Milon et al., 2008].

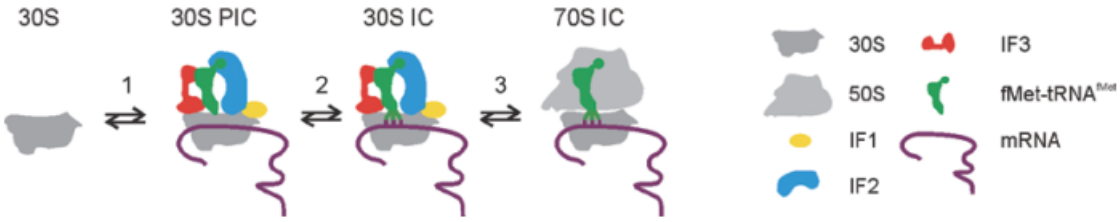


Figure 1.6: Recruitment of the initiation factors (IFs) and formation of the ribosomal complex. (modified from [Milón et al., 2012]) Phase 1: Assembly of the 30S PIC upon recruitment of initiation factors, mRNA, and fMet-tRNA^{fMet} to the 30S subunit. Phase 2: Conversion of 30S PIC to 30S IC after start codon recognition by fMet-tRNA^{fMet} . Phase 3: Formation of the 70S IC following 50S subunit joining and release of initiation factors.

Translation initiation involves the interaction between the initiation complex and the mRNA. It requires initiation factors (IFs that are IF1, IF2, IF3), the 30S subunit of the ribosome and the tRNA^{fMet}. In the first phase of initiation, IF1, IF2, IF3, mRNA and the initiator tRNA^{fMet} bind to the 30S ribosomal subunit, forming the 30S initiation complex (Figure 1.6). The order of interaction of the 30S subunit, the initiation factors and the tRNA^{fMet} has been revealed by Milon *et al.* [Milón et al., 2012]. IF2 and IF3 are the first two elements to bind to the 30S subunit, followed by IF1 (Figure 1.7). The tRNA^{fMet} is the last element recruited by the 30S•IFs complex. The mRNA recruitment is independent of the IFs and tRNA^{fMet} binding to the 30S subunit [Milón et al., 2012].

Figure 1.7: Schematic of the three main phases of translation initiation. (from [Milón et al., 2012]) Kinetically favored recruitment pathway for initiation factors and fMet-tRNA^{fMet}. IF3 and IF2 are the first to bind to the 30S subunit, followed by IF1.

The first amino acid of a polypeptide chain is always a methionine. Methionine is brought to the ribosome by the Thermo unstable Elongation Factor (EF-Tu). In bacteria, the methionine bound to the tRNA^{fMet} is N-formylated. It selectively excludes the fMet-tRNA^{fMet} from the elongation phase of translation [Laursen et al., 2005]. Alternative initiation codons related to AUG by a single base change are found in some genes. These codons are all decoded by the initiator fMet-tRNA^{fMet} and translated as formylmethionine. Methionine-isoaccepting initiator and elongator tRNAs are both aminoacylated by methionyl tRNA synthetase. Aminoacylated initiator tRNA is formylated by Methionyl tRNA Transformylase (MTF). Formylation of Met-tRNA^{fMet} is important for protein synthesis in *E. coli*. Indeed, mutants defective in formylation are extremely poor in initiation of protein synthesis. More-

over, a strain of *E. coli* carrying disruptions in the *fmt* gene encoding MTF has severe growth defects [Laursen et al., 2005].

The roles of the different initiation factors are not fully elucidated but some IFs functions have been widely studied in *E. coli* and *B. subtilis*.

i. In *B. subtilis*, IF1 is encoded by the *infA* gene. It stimulates the activity of IF2 and IF3 [Pon and Gualerzi, 1984]. Interaction between IF2 and the 30S ribosomal subunit is favored when IF1 is bound, and the release of IF2 is indirectly promoted when IF1 is ejected. IF1 cooperates with IF2 to ensure that only the initiator tRNA binds to the P-site and that it interacts with the initiation codon of the mRNA [Zucker and Hershey, 1986]. IF1 occludes the A-site until the 70S initiation complex has formed [Carter et al., 2001]. Release of IF1 consequently opens the A-site for incoming aminoacyl-tRNAs (aa-tRNAs).

ii. IF2 is encoded by the *infB* gene. Guanosine triphosphate (GTP) hydrolysis in translation initiation has been suggested to be important for the release of IF2 from the 70S initiation complex. It also seems to be important for the adjustment of the initiator tRNA (fMet-tRNA$^{\mathrm{fMet}}$ which will be further noted tRNA$^{\mathrm{fMet}}$) in the ribosomal P-site [Fabbretti et al., 2012] [Laursen et al., 2005]. The GTP-bound form of IF2 accelerates association of the ribosomal subunits. And it has been suggested that not only GTP but, to a lesser extent, also GDP drive the GTPase IF2 into its active conformation [Mitkevich et al., 2010]. On the 30S subunit, GTP -but not GDP- strongly drives IF2 into an "active" conformation with a high affinity for tRNA$^{\mathrm{fMet}}$ [Mitkevich et al., 2010]. The subsequent GTP hydrolysis accelerates the release of IF2 from the 70S ribosome. Three isoforms of IF2 exist in *E. coli*. *B. subtilis* also possesses more than one isoform of IF2 [Hubert et al., 1992]. These homologues have similar functions to those of IF2, including GTPase activity, promotion of ribosomal subunit association, and probably interaction with the tRNA$^{\mathrm{fMet}}$.

iii. IF3 is encoded by the *infC* gene. IF3 prevents the association of the ribosomal subunits by binding to the 30S subunit, thereby blocking binding of the 50S subunit [Ogle et al., 2001]. Initiation complexes with an incorrectly bound aminoacyl-tRNA (non-initiator tRNA) and complexes with triplets other than AUG, GUG, and UUG in the P-site are dissociated by IF3. IF3 stimulates the rapid formation of codon-anticodon interaction at the ribosomal P-site. IF3 is involved in the adjustment of the mRNA from the standby site to the decoding P-site of the 30S ribosomal subunit. IF3 enhances the dissociation of deacylated tRNAs from post termination complexes as well as the dissociation of 70S ribosomes into subunits [Laursen et al., 2005].

1.2.3.3 Translation elongation

As for the initiation factors, the roles of the different Elongation Factors (EFs) have been elucidated in *E. coli*. The elongation factor EF-Tu (also called EF1A) and elongation factor G (EF-G, also called EF2) from bacteria are multi-domain GTPases with essential functions in the elongation phases of translation. The general biochemical outline of the translation elongation cycle is well preserved in

bacteria [Andersen et al., 2003].
EF-Tu is activated upon GTP binding, and forms a ternary complex with aminoacylated elongator tRNAs. On the ribosome, the ternary complex of EF-Tu decodes the genetic information. It is realized via hydrogen bounds between the mRNA codon and the anticodon of a cognate tRNA. Such a decoding event triggers the ribosome to induce GTP hydrolysis. The GDP-bound EF-Tu is then released from the ribosome. EF-G participates in the translocation of both the tRNAs and the mRNA by exactly one codon on the ribosome.

The elongation phase involves the three sites of the 70S complex, the aminoacyl (A), the peptidyl (P) and the exit (E) sites. At the end of the initiation phase the P-site of the 70S complex is occupied by the $tRNA^{fMet}$ and is used to start the elongation cycle. The aminoacyl-tRNA (aa-tRNA) binds to the empty A-site of the 70S complex with the EF-Tu and a GTP molecule. If the anti-codon sequence pairs with the codon, the interaction results in a conformational change which stabilizes the aa-tRNA binding and activates the GTP hydrolysis by EF-Tu [Ogle et al., 2002] (Figure 1.8). Then, the aa-tRNA swings into the peptidyl transferase site (accommodation steps). The formation of the peptide bond is spontaneous [Ramakrishnan, 2002]. Translocation of the tRNAs interacting with the mRNA involves EF-G and its GTPase activity and results in an empty A-site that can receive another aa-tRNA corresponding to the codon present in the A-site.

1.2.3.4 Translation termination

Translation termination is catalysed by Release Factors (RFs) that recognize stop codons present in the A-site. In *E. coli*, three RFs are involved in ribosome release, RF1, RF2 and RF3. RF1 and RF2 are involved in the recognition of termination signals: RF1 recognizes codons UAG or UAA and RF2 recognizes codons UAA or UGA [Buckingham et al., 1997]. In response to a stop codon, RF1 or RF2 hydrolyses and releases the completed polypeptide from the peptidyl-tRNA (Figure 1.9). The bacterial translational GTPase RF3 promotes translation termination by recycling RF1 or RF2 [Kihira et al., 2012].
The Ribosome Recycling Factor (RRF) is a factor that catalyses the ribosome recycling when added along with EF-G and GTP (Figure 1.9). EF-G is required along with RRF to carry out ribosome recycling. It is supposed that RRF binds to the ribosomal A-site and that it is translocated by EF-G in a manner similar to tRNAs since it structurally resembles tRNAs. It was postulated that this translocation activity results in the release of tRNA, followed by the release of mRNA from the ribosome [Hirokawa et al., 2002]. However, the mechanistic details of such a 'dissassembly' process remain unclear.
Finally, RF3 replaces the deacylated tRNA onto the 30S subunit (Figure 1.9).

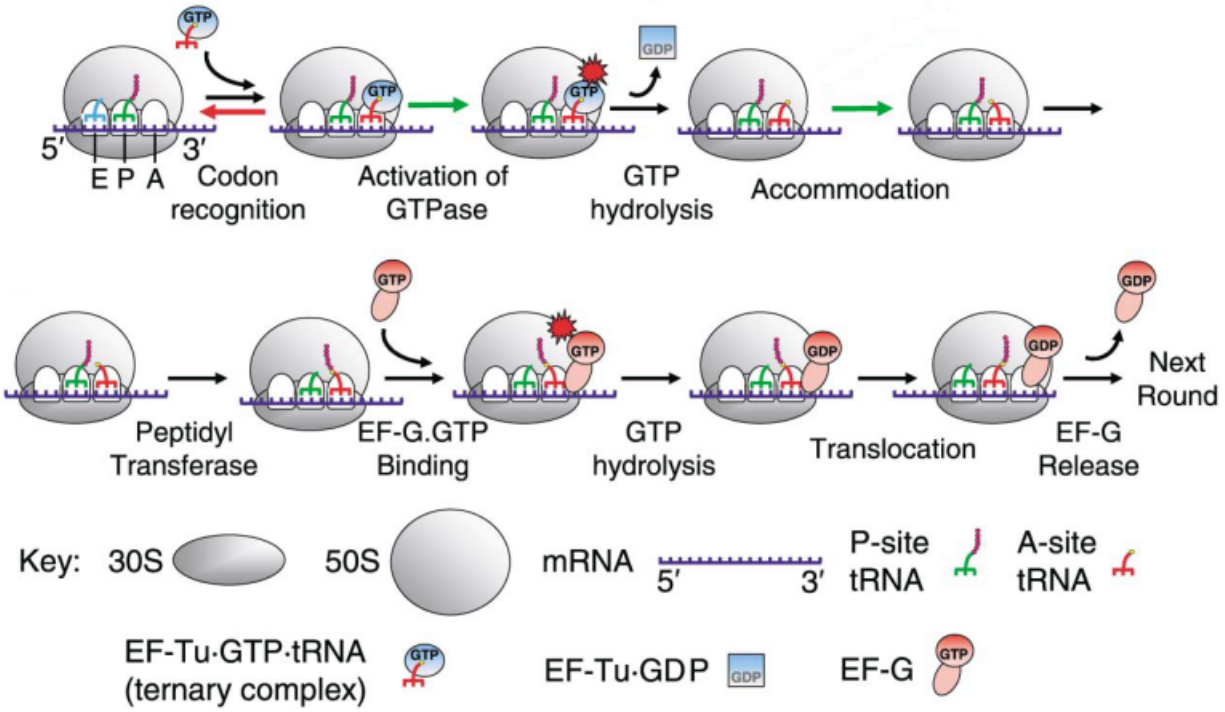


Figure 1.8: Overview of the elongation pathway (adapted from [Ramakrishnan, 2002]) After codon recognition in the A-site, the GTP is hydrolysed leading to accommodation of the amino acid onto the forming protein followed by translocation of the A-site and P-site tRNAs to the P-site and E-site respectively. There are two GTP hydrolysis steps in the elongation process, the first one catalysed by EF-Tu and the second one by EF-G. Then another round of elongation can start.

Overview: Translation in bacteria consists in three main steps: initiation, elongation and termination. It requires the action of the ribosome which consists in two subunits, the 30S and 50S, made of *rrns* and ribosomal proteins. The 30S will specifically recognize the translation initiation region (TIR) on the mRNA before initiating translation. This cellular process also requires the action of specific factors: initiation factors (IFs) during initiation; elongation factors (EFs) during elongation; and release factors (RFs) as well as the Ribosome Recycling Factor (RRF) during termination. IF1, IF2 and IF3 are required to initiate translation by binding to the 30S subunit, followed by the $tRNA^{fMet}$ incorporation. Then the 50S subunit is recruited to form the 70S while the IFs are released. Once the $tRNA^{fMet}$ is in the P-site, the elongation proceeds and for each new codon to translate, the same mechanism occurs. If the aa-tRNA present in the A-site corresponds to the codon also present there, EF-Tu hydrolysis GTP. Then, the tRNA is translocated onto the P-site while the empty tRNA already present in the P-site is translocated onto the E-site from which it will be ejected. This step takes place through the hydrolysis of GTP by EF-G. The translation process terminates when a stop codon is present in the A-site which is recognized by RFs that then trigger the release of the peptides chain.

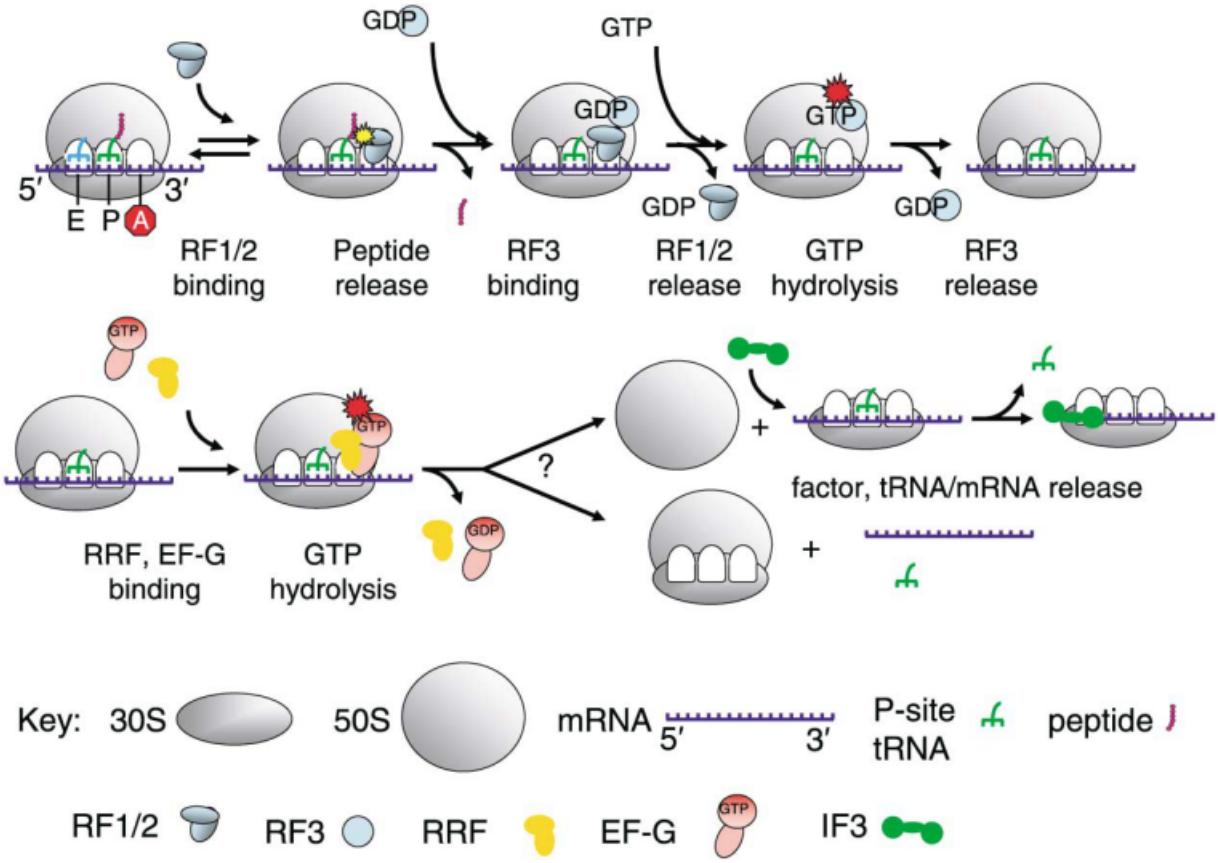


Figure 1.9: Termination of the translation (from [Ramakrishnan, 2002]) The question mark highlights the peptide release after stop-codon recognition and means that the mechanism of signal transduction is still unclear.

1.2.4 Mechanisms to prevent amino acid misincorporation during translation

Protein synthesis errors can be harmful to the cell if it concerns a too high percentage of them. Thus, during translation there exists different steps to prevent mistakes from happening.

1.2.4.1 Post-transcriptional modifications of tRNAs

tRNAs are adaptor molecules of typically 75 to 90 nucleotides in length which are needed by the ribosome to decode the mRNA sequence and build the desired protein [Dirheimer et al., 1995]. It possesses an anticodon that is complementary to the codon found in the A-site. The nucleosides that constitute it are found at position 34, 35 and 36 and the codon they form is noted $N_{34}N_{35}N_{36}$. Given that across the different species there exists 61 amino acids codons but far fewer tRNAs, Francis Crick published the Wobble hypothesis: the nucleoside at position 34 of the anticodon could form non-canonical hydrogen bonding. This means that the uridine (U) at this position can base pair with Guanosine (G) or Inosine (I) [Crick, 1966].

Moreover, there exists modified nucleosides in the Anticodon Stem and Loop (ASL) domain and they are usually found at positions 27, 28, 31, 39 and 40 (stem part) and at positions 32, 34, 35, 37 and 38

(loop part) [Agris et al., 2017](Figure 1.10). These nucleosides are not all modified in only one tRNA but a set of 3-5 specific nucleosides are found modified in a single tRNA. The modifications of these nucleosides lead to the creation of nucleosides different from the nucleotides found in the DNA/RNA code. Such modifications have different roles such as the expansion of codon recognition (*i.e.* different codons are decoded by the same tRNA) or ensuring the stability of the tRNA in the A-site [Agris et al., 2017].

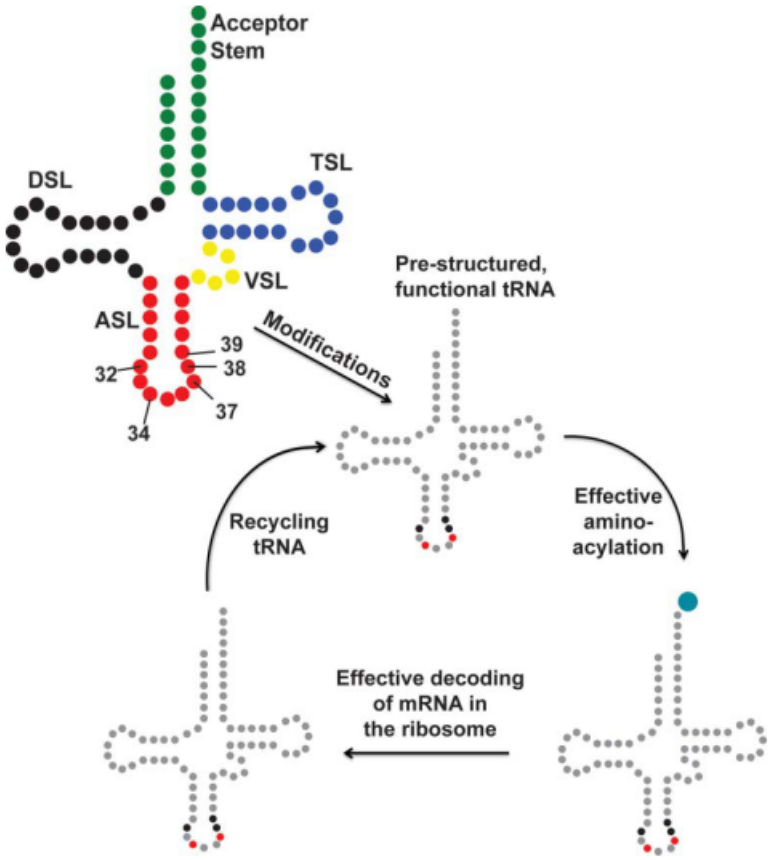


Figure 1.10: Structure of tRNA and its life cycle. (from [Agris et al., 2017]) The secondary structure of tRNA with its constituent domains marked in different colors (top left): Acceptor Stem (green); Dihydrouridine Stem and Loop, DSL (black); Anticodon Stem and Loop, ASL (red); Variable Stem and Loop, VSL (yellow); Thymidine Stem and Loop, TSL (blue). tRNA transcripts are processed by sizing and modification, some are spliced, before functioning in translation. Modification of tRNAs, particularly the anticodon stem and loop (ASL) domain at positions 32, 34, 37, 38 and 39, is an important step toward achieving functional chemistry and architecture. The wobble nucleoside, first of the anticodon, is position 34. Red and black highlights of mature tRNA after modification indicate the locations in the ASL where it is heavily modified.

Once the tRNA has been modified at specific nucleosides, it can be aminoacylated by dedicated aa-tRNA synthetases and then be used by the ribosome to build the desired protein (Figure 1.10). To be accepted by the ribosome, the aa-tRNA has to be cognate, meaning its anticodon corresponds to the codon present in the A-site. If the aa-tRNA anticodon is very different from the A-site codon

(only one nucleoside able to interact with one mRNA base out of three or none), it is rejected by the ribosome and referred to as non-cognate aa-tRNA. If the aa-tRNA anticodon is able to interact with two base pairs out of three of the mRNA codon, it is referred to as near-cognate. Such aa-tRNAs can be mistakenly accepted leading to the incorporation of a wrong amino acid into the protein as will be discussed further. Once the empty tRNA has left the ribosome E-site, it is recycled and then it can be aminoacylated, and the cycle starts again (Figure 1.10).

1.2.4.2 First selection criterion: Watson-Crick base pairing

Structural studies have shown that the cognate ASL binding induces conformational changes in the 30S subunit, referred to as "domain closure" [Rozov et al., 2016]. It has been suggested that most of the near-cognate aa-tRNAs present in the A-site will be rejected since they will not be able to form a stable codon-anticodon base pairing (*i.e.* canonical Watson-Crick base pairing), thus preventing the change of conformation leading to 30S-domain closure.

However, some near-cognate aa-tRNAs will be incorporated leading to misincorporation. This is due to the formation of a stable Watson-Crick-like base pair in some codon-anticodon combinations such as a G•U pair [Rozov et al., 2016](Figure 1.11). Indeed, the complex translation machineries involved in codon recognition are sensitive to the shapes of the base pairs but not to the number or types of hydrogen bonds formed between the codon and the anticodon of the A-site tRNA [Westhof et al., 2014]. This explains how Watson-Crick-like base pairs are able to trigger GTP hydrolysis. Such conformations can be obtained by tautomerism in one of the bases (*i.e.* tautomerization is a reversible chemical reaction between an isomer couple), by a non-natural non-polar residue complementary to a standard base, or by a mixture of both in some nucleotide analogues that can pair to standard bases [Westhof et al., 2014].

Moreover, the aa-tRNA nucleoside residue 34, the one interacting with the third nucleotide of the mRNA codon present in the A-site, can be modified as explained in the previous part. This leads to non-standard pairing which is still able to form a Watson-Crick-like geometry like, for instance, a G pairing with a U. Consequently this will trigger GTP hydrolysis and *in fine* a "wrong" amino acid will be accommodated to the synthesized protein. This gives a hint about why a near-cognate aa-tRNA can still be accepted.

1.2.4.3 GTPase activation in presence of a cognate aa-tRNA.

Loveland *et al.* recently deciphered how aa-tRNA recognition activates the GTPase center of EF-Tu, and how cognate and near-cognate aa-tRNAs are discriminated [Loveland et al., 2017]. There are three binding steps for the cognate and near-cognate ternary complexes and there exists differences between their pre-accommodation states, explaining why only the cognate aa-tRNA is incorporated (Figure 1.12).

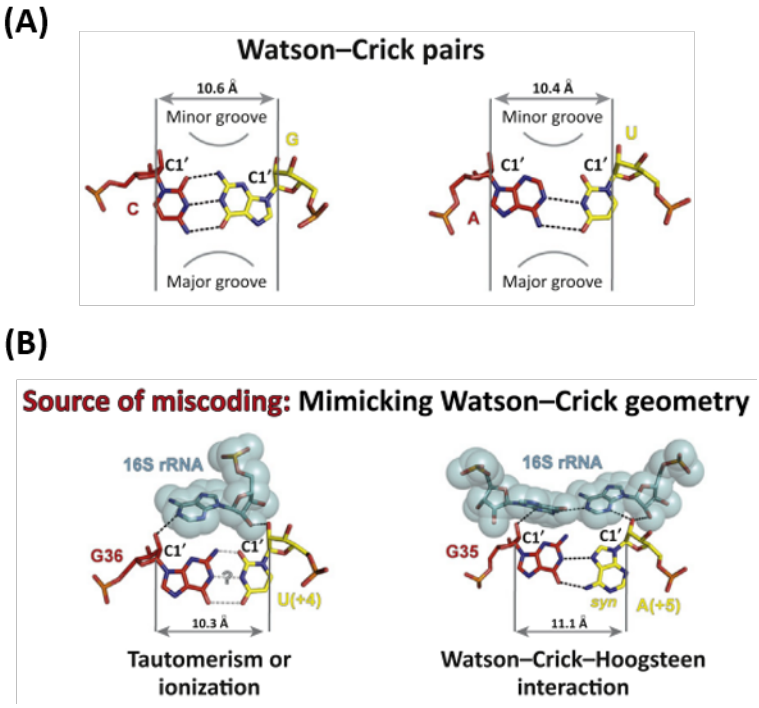


Figure 1.11: Watson-Crick pairs and pairs mimicking Watson-Crick geometry (adapted from [Westhof et al., 2014]) (A) The canonical Watson-Crick paires. (B) The base pairs that are capable of mimicking the canonical geometry may avoid discrimination and result in translational errors. For (A) and (B), 16S rRNA nucleotides are shown in teal, mRNA nucleotides in yellow and aa-tRNA nucleosides in red. C1'-C1' distances are indicated and average distances for the pairs are presented in parenthesis.

Concerning the cognate ternary complex, in structure I, it binds the 30S subunit, but the anticodon does not base-pair with the codon, nor does EF-Tu contact the 50S subunit. In structure II, the anticodon base pairs with the codon, while EF-Tu remains distant from the 50S subunit. In structure III, the anticodon base-pairs with the codon, and EF-Tu contacts the Sarcin-Ricin Loop (SRL), a highly conserved region of the 50S subunit. The conformation changes from structure I to structure III, which leads to 30S-domain closure through the nucleotide G530 at the tip of the 30S shoulder. Then, such conformation allows the GTPase domain of EF-Tu to bind the SRL. This binding prearranges the EF-Tu catalytic site for GTP hydrolysis. Once GTP is hydrolyzed into GDP, EF-Tu is released and elongation can continue. Thus the conformation of the nucleotides of the decoding center, especially G530, allows the 30S-domain closure leading *in fine* to GTP hydrolysis and EF-Tu release.

Concerning the near-cognate ternary complex, the three structures are similar to the three structures of the cognate one. For structure I and II, the difference lies in the decoding center, especially G530 structure which is less resolved. Indeed, since one couple of nucleotides does not match, the codon-anticodon interaction deviates from the Watson-Crick conformation. Thus this event shifts the

anti-codon and prevents the closure of the 30S-domain. However, concerning structure III, there is a tautomeric Watson-Crick-like conformation for the codon-anticodon interaction which has already been observed as discussed previously. Thus, the decoding-center nucleotides adopt a conformation that allows G530 to act as a latch and thus this leads to 30S-domain closure. As seen previously, this triggers GTP hydrolysis and EF-Tu is finally released. This explains why misincorporation can potentially occur when a near-cognate tRNA is present at the A-site.

1.2.4.4 Proofreading steps occuring after GTP hydrolysis

Two proofreading steps also occur after GTP hydrolysis (Figure 1.13) [Ieong et al., 2016]. Previous works have shown that non-matching aa-tRNAs can be discriminated through the GTP hydrolysis rate which is different according to the nature of the ternary complex: cognate, near-cognate or non-cognate [Maracci and Rodnina, 2016]. The step described previously that includes the 30S domain closure followed by GTP hydrolysis is the first one for aa-tRNA selection (called "Initial selection" in Figure 1.13).

First, after GTP hydrolysis, either the ternary complex EF-Tu•GDP•aa-tRNA or EF-Tu•GDP dissociates from the ribosome (Figure 1.13). If EF-Tu•GDP releases the aa-tRNA slowly, then there is a high probability that the whole EF-Tu•GDP•tRNA complex actually leaves the ribosome.

Then, a second proofreading step occurs after EF-Tu has left the ribosome complex. Indeed, at this stage, it is still possible that the ribosome ejects a non-matching aa-tRNA. This proofreading step (Figure 1.13) is based on kinetic discrimination: a cognate aa-tRNA in complex with EF-Tu•GTP will rapidly activate GTP hydrolysis and EF-Tu accommodation into another conformation [Maracci and Rodnina, 2016]. The GTP hydrolysis and the EF-Tu change of conformation are both rate-limiting steps, thus a higher kinetic for these two events will favour cognate aa-tRNAs over non or near-cognate ones. More precisely, if the kinetic is not fast enough due to the presence of a non or near-cognate aa-tRNA in the A-site, this will lead to the release of the tRNA even after EF-Tu has left the ribosomal complex; which should not be the case for a cognate tRNA.

Overview: To prevent the incorporation of a wrong amino acid in the peptide chain, the cell can rely on different molecular and kinetic properties during translation. The amino acid selection takes place in the A-site and is based on conformational and kinetic discriminations to prevent misincorporation. It is important to note that specific tRNAs'nucleosides undergo several post-transcriptional modifications that enhance aa-tRNA selectivity and stability in the A-site. An aa-tRNA will not be rejected if the nucleotides of the A-site mRNA codon and the nucleosides of the aa-tRNA anticodon form stable interactions (Watson-Crick or Wobble base-pairing). This should discriminate cognate aa-tRNAs from non-cognate or near-cognate aa-tRNAs. Moreover, after GTP hydrolysis by EF-Tu, two proof-reading steps take place to eject non-cognate or near-cognate aa-tRNAs which would still be in the A-site after GTP hydrolysis. Nevertheless, near-cognate aa-tRNAs can still go through these three proofreading steps by adopting favourable conformations which will lead to misincorporation.

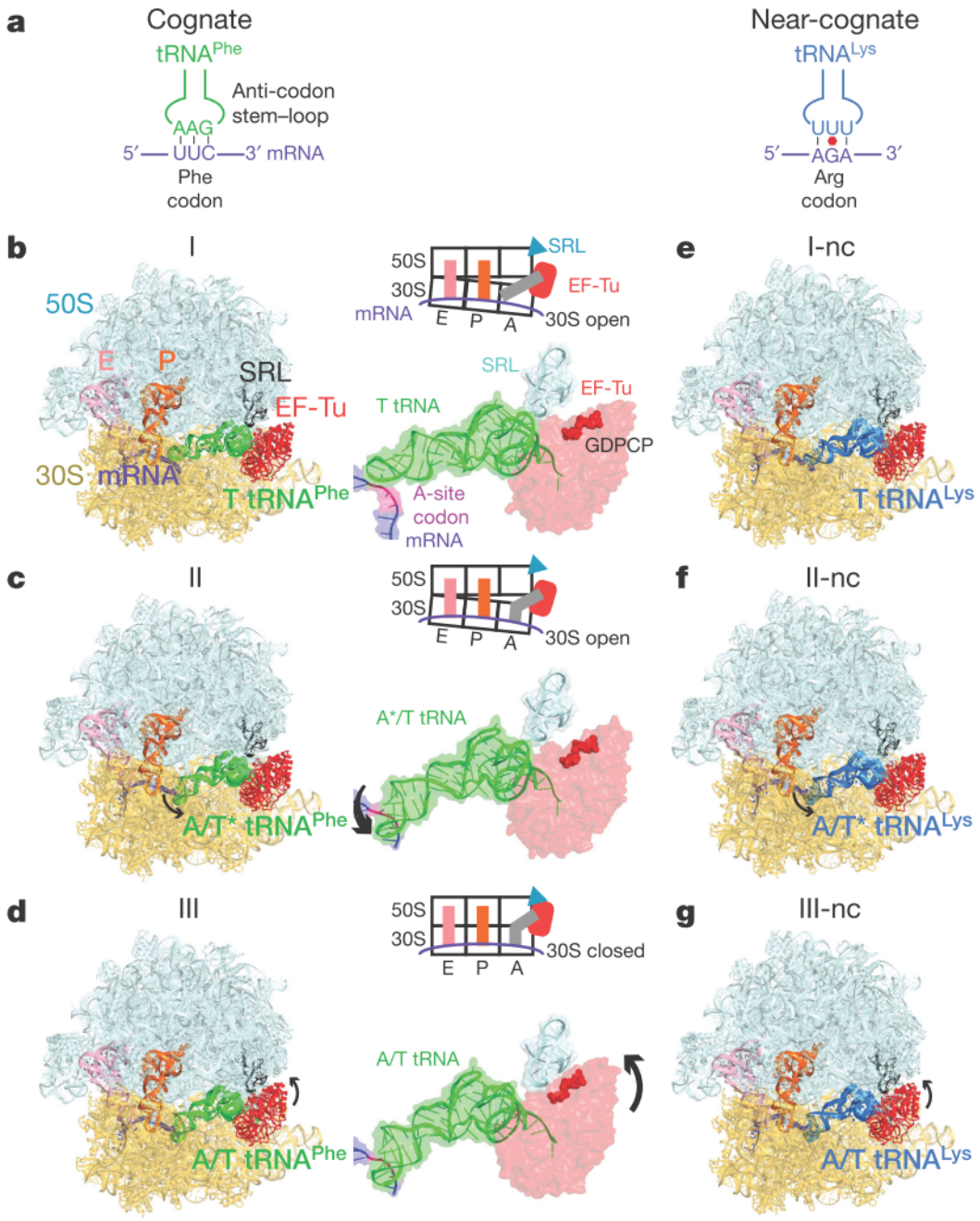


Figure 1.12: Comparison of the structures of a cognate and a near-cognate ternary complex on the 70S ribosome (from [Loveland et al., 2017]) (a) Schematic representation of the cognate (green) and near-cognate (blue) tRNA anticodons and A-site codons used in complexes. b, c and d correspond to the three structures (I, II and III respectively) that a cognate aa-tRNA can adopt when present in the A-site. e, f and g correspond to the three structures (I-nc, II-nc and III-nc respectively) that a near-cognate aa-tRNA can adopt when present in the A-site. In b, c, d, e, f and g the E-site tRNA is in pink, the P-site tRNA is in orange, EF-Tu is in red, the 50S subunit is in light blue and the 30S subunit is in camel. The A-site aa-tRNA is in green in b, c, d, and it is in dark blue in e, f, and g. The structures I-nc, II-nc and III-nc of the near-cognate aa-tRNA (represented in e, f, and g) globally resemble the three cognate structures I, II and III (respectively represented in b, c and d).

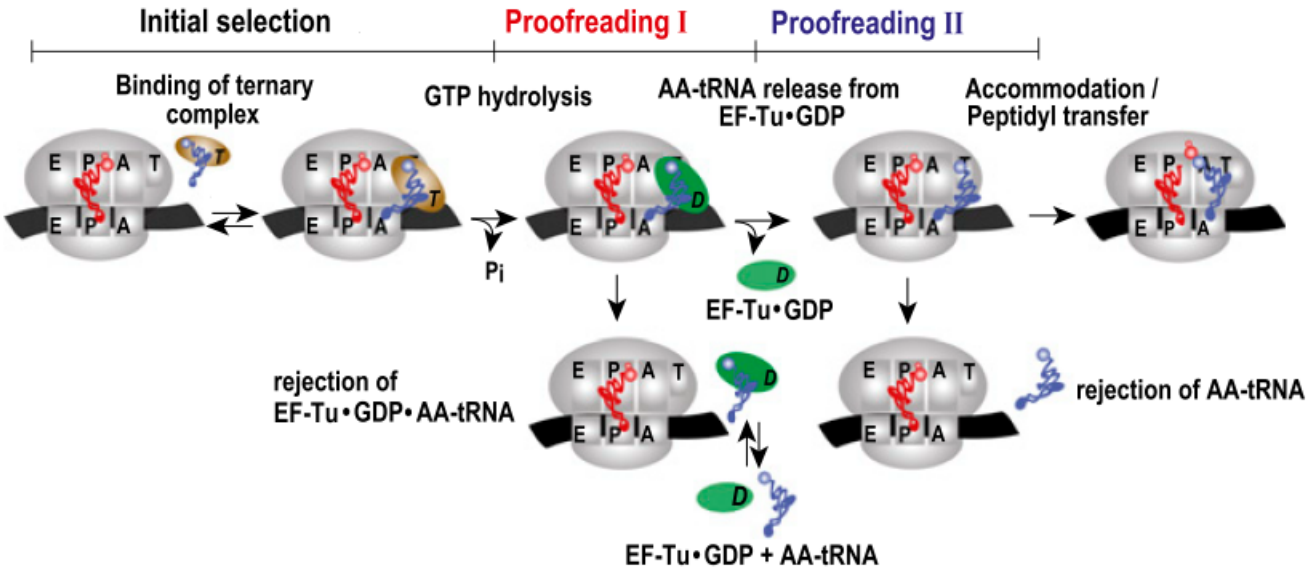


Figure 1.13: Two proofreading steps occur after GTP hydrolysis by EF-Tu during the initial selection step. (adapted from [Ieong et al., 2016]) During the Initial selection step, a ternary complex composed of aa-tRNA and EF-Tu•GTP binds to the pre-translocation ribosome. Then this ternary complex either dissociates or GTP is hydrolyzed by EF-Tu leading to ribosome-bound ternary complex EF-Tu•GDP•aa-tRNA. Then this complex either dissociates or EF-Tu•GDP from the complex leading to an aa-tRNA-bound preaccommodation state of the ribosome (step Proofreading I). From this state, the aa-tRNA either dissociates or accommodates into the A-site (step Proofreading II) and then translocation can take place. Near-cognate tRNA can be rejected during these three different steps: Initial selection, Proofreading I, and Proofreading II.

Chapter 2

Bacterial adaptation to nutritional changes

2.1 Background

Bacteria, as all organisms, have developed many strategies throughout evolution to adapt to environmental changes. For instance, their size will be modified according to the medium where they grow in (Figure 2.1). The larger cells have been grown in 'rich' medium while the smaller ones have been grown in 'poor' medium. Thus, the nutrient composition directly impacts the cell's physiology. The study of bacterial growth has led to the establishment of fundamental growth laws described in the next parts.

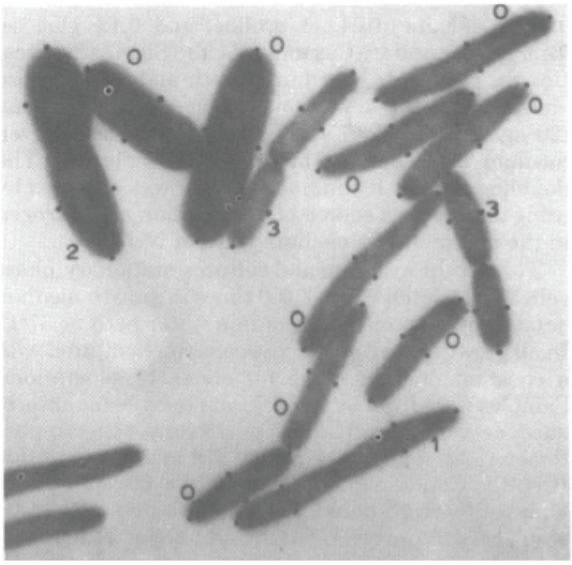


Figure 2.1: *E. coli* cell size is different under different growth conditions (adapted from [Trueba and Woldringh, 1980]). Electron microscopic picture of *E. coli* cells grown in different nutrient conditions.

The growth of bacteria in batch cultures is modelled with four different phases: the lag phase, the log

or exponential phase, the stationary phase and the death phase [Schaechter et al., 2006] (Figure 2.2). During the lag phase, bacteria adapt to the growth conditions according to the nutrients quantity and quality. During the exponential or steady-state phase, the cells are dividing and are characterized by a growth rate which is a measure of the number of divisions per cell per unit time. The characteristics of this phase are time-invariant which makes it a standard for microbial growth studies (see next section). Then the stationary phase occurs when the first non-dispensable nutrient becomes limiting. Once the required nutrients are completely depleted the cells start to die but certain bacterial species such as *B. subtilis* can sporulate and thus survive during many years. Cell adaptation to these nutritional changes requires rearrangements at the macromolecular level as well as re-tuning of several cellular processes. This is what will be discussed in the following parts.

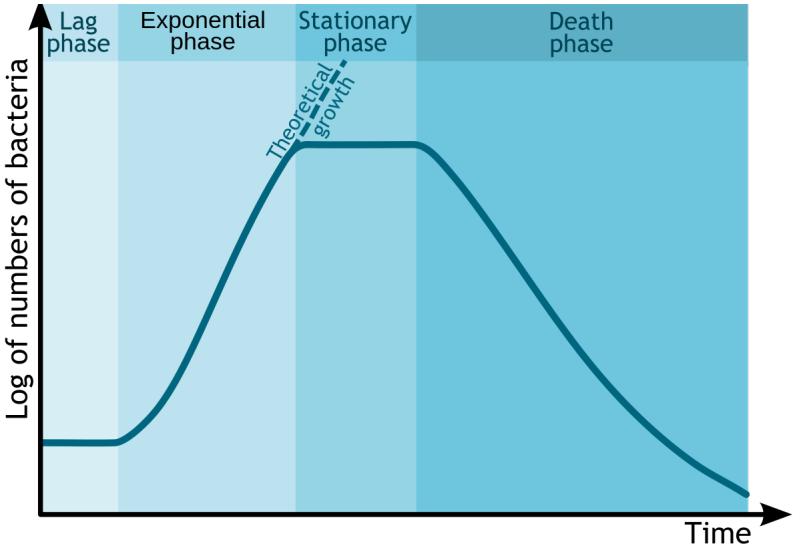


Figure 2.2: Bacterial growth curve (illustration by Michal Komorniczak). The growth of bacteria presents four different phases.

2.1.1 The growth laws of bacterial physiology

From experimental observations, two main growth laws have been stated to describe the link between cell size and growth rate as well as the relation between the cell macro-molecular content and its growth rate.

2.1.1.1 The first growth law

In his review of 1949 [Monod, 1949], Jacques Monod showed that the exponential growth rate μ has an hyperbolic dependence on the concentration of a growth-limiting substrate Su:

$$\mu = \mu^0_{max} * \frac{Su}{Su + K_D}$$

where the phenomenological parameters μ^0_{max} and K_D are properties of the bacterial strain and the growth-limiting nutrient Su (Figure 2.3).

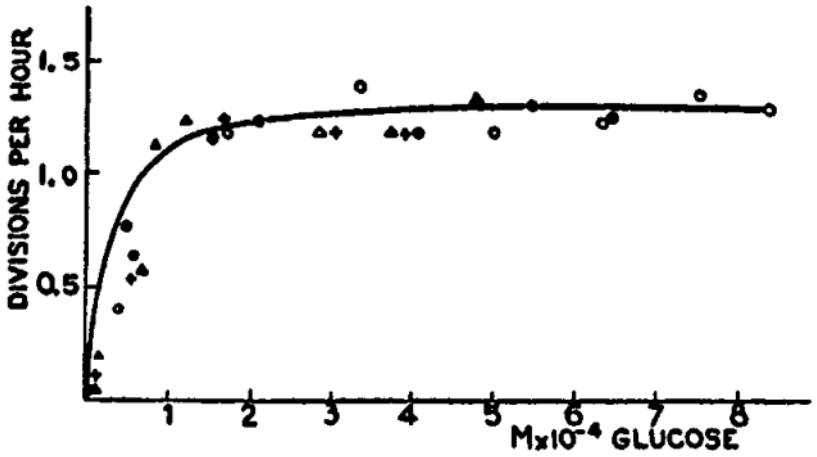


Figure 2.3: The first growth law (adapted from [Monod, 1949]). Representation of the growth rate of *E. coli* grown in synthetic medium at 37°C as function of glucose concentration (Su). The solid line is drawn based on Monod's equation with $\mu^0_{max} = 1.35$ divisions per hour, and $K_D = 0.22$ x10^{-4} M.

2.1.1.2 The second growth law

In 1958, Schaechter *et al.* observed a relationship between the average cell size and the nutrient-imposed growth rate (Figure 2.4). This foundational principle in bacterial physiology states that the average cell size (S) has an exponential dependence on the nutrient-imposed growth rate μ in steady-state growth such that:

$$S \propto e^{a\mu}$$

where a is a constant. This means that media leading to the same cell size confer to the cell the same macromolecular composition. This growth law has been recently confirmed at the population level in both *E. coli* and *B. subtilis* [Taheri-Araghi et al., 2015].
Nevertheless, this growth law does not apply at the single-cell level since individual cells exhibit intrinsic size variability even under constant growth conditions [Taheri-Araghi et al., 2015]. Cells are found to employ an "adder" principle: the size added between birth and division (Δ) is constant for given growth conditions [Taheri-Araghi et al., 2015]. Δ varies significantly between growth conditions and between individual cells but it is constant on average. This explains why the second growth law at the population level is valid even if it is not observed at the single-cell level.

Moreover, Maaløe and Kjeldgaard (1966), known as the 'Copenhagen School', were the first to experimentally determine the biomass composition of *E. coli.* From their results emerged another relationship derived from the second growth law: the macro-molecular constituents of the cell (*i.e.* DNA, RNA,

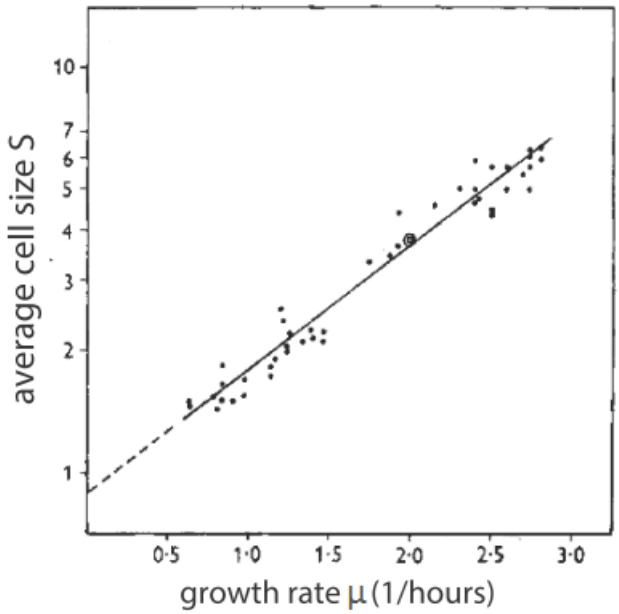


Figure 2.4: The second growth law: dependence of the average cell size on the growth rate. (adapted from [Schaechter et al., 1958] and the review of [Jun et al., 2018]). The nutrient growth law discovered in 1958 reveals a quantitative relationship between the average cell size and the nutrient-imposed growth rate.

protein contents as well as ribosomes and RNAPs) exhibit a functional dependence on the growth rate (Figure 2.5) [Maaløe and Kjeldgaard, 1966].

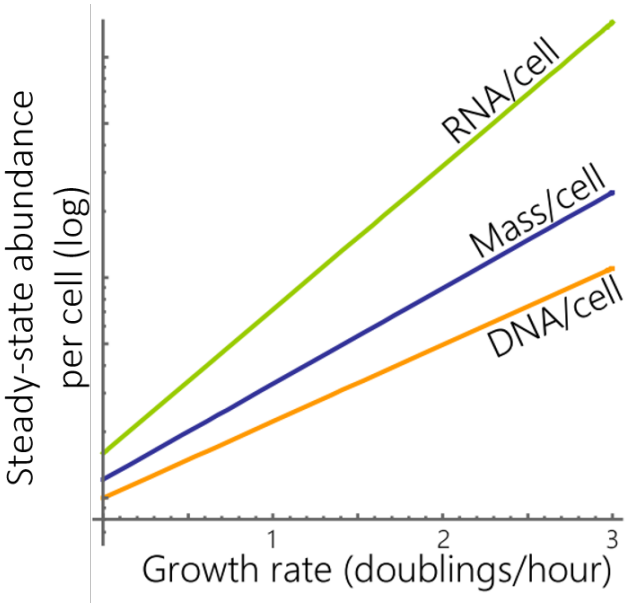


Figure 2.5: The second growth law: dependence of the cell macro-molecular constituents on the growth rate. (adapted from the review of [Jun et al., 2018]). When the growth rate is changed by the quality of the available nutrients, the per-cell abundance of RNA, Mass and DNA scale approximately exponentially with the growth rate μ.

2.1.1.3 The third growth law

The third growth law derives from observations made by Neidhardt and Magasanik [Neidhardt and Magasanik, 1960] who established that the ribosome plays a catalytic role in protein synthesis. They inferred this from experimental data which showed that the correlation between the RNA/protein ratio and growth rate was approximately linear for doubling rates above 0.6 doublings/hour (Figure 2.6); and also that the ribosomal RNA fraction of total RNA representing 86% was growth rate independent. Scott *et al.* (2010) mathematically formulated the third growth law:

$$r = r_0 + \frac{\mu}{\kappa_t}$$

where r is the RNA/protein ratio, r_0 is the vertical intercept and the parameter κ_t is the inverse of the slope predicted to be proportional to the rate of protein synthesis.

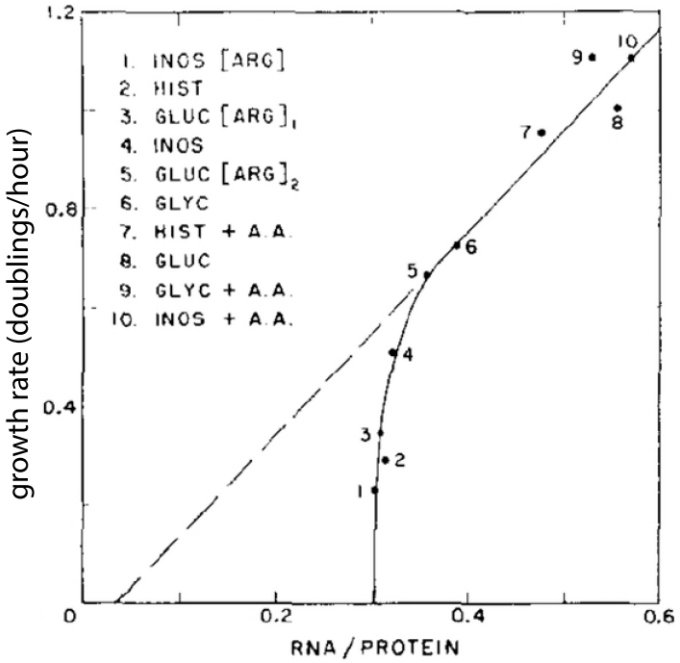


Figure 2.6: The third growth law: relation between RNA/protein ratio and growth rate (adapted from [Neidhardt and Magasanik, 1960]). The graphic represents the relation between the growth rate and the RNA content (RNA/protein) of *Aerobacter aerogenes* cells grown in different media.

Overview: The cell's physiology is impacted by its nutritional environment. In batch cultures, bacterial growth is modelled with four different phases: the lag phase, the exponential (or steady-state) phase, the stationary phase and the death phase. Experimental observations lead to the statement of two growth laws during steady-state phase. The first one states that the exponential growth rate μ has an hyperbolic dependence on the concentration of a growth-limiting substrate. The second one states that the average cell size has an exponential dependence on the nutrient-imposed growth rate μ. By extension, this law also states that the abundance of macro-molecular constituents of the cell (*i.e.* DNA, RNA, proteins, ribosomes and RNAPs) increases exponentially with the growth rate μ. The third growth law states that the ribosome plays a catalytic role in protein synthesis.

2.2 Bacterial physiology in Steady-State: from the fundamental cellular processes to resource allocation

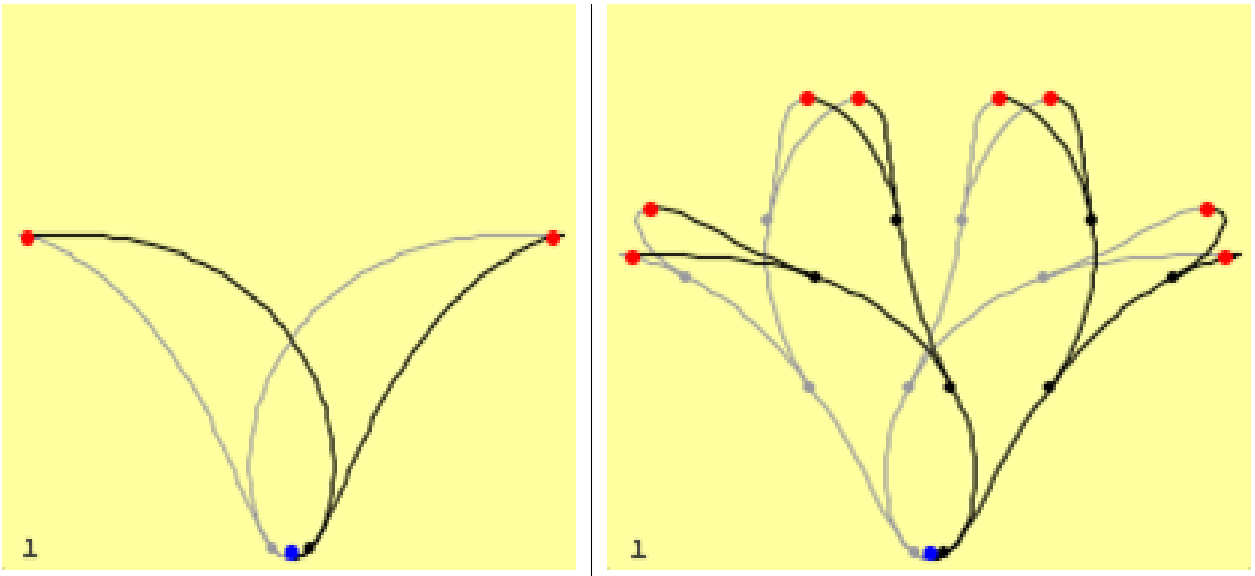
2.2.1 Growth-rate dependence of the main cellular processes

As stated previously, the cell needs to regulate its macromolecular content to fit to the resource availability to adapt to nutritional changes. These changes are growth-rate dependent and occur at the level of DNA replication, transcription and translation. They can be referred to as "global regulation".

2.2.1.1 Growth-rate dependence of DNA replication

The time for the cell to divide is shorter than the time for the chromosome to be entirely replicated at fast growth, a phenomenon observed in both gram positive and gram negative bacteria [Bremer and Dennis, 2008]. Indeed, the rate of DNA replication increases with growth rate until it saturates when a certain growth rate value is reached [Bremer and Dennis, 2008][Klumpp et al., 2009]. The bacterial cell-cycle model of Cooper and Helmstetter (1968) states that there exists overlapping rounds of chromosome replication (*i.e.* a round of replication is initiated before the previous one ends) (Figure 2.7 (b)). To increase DNA replication efficiency, two replication forks are created starting at the same origin but progressing clockwise and counter clockwise on the circular chromosome (figure 2.7). Thus, the number of replication forks will increase along with the growth rate given that doubling time shortens whereas replication time remains constant (figure 2.7).

It is important to note that genes near the origin of replication will quickly outnumber the ones at the terminus of replication which is very likely to influence the proportion of their transcripts and proteins. This suggests that genes needed for growth tend to be localized close to the origin of replication. Indeed, four out of the seven rRNA operons are clustered near the origin of replication in *E. coli* [Jin et al., 2012].



(a) **(b)**

Figure 2.7: DNA replication(a) DNA replication with only one replication fork (b) DNA replication with multiple replication forks for an increased growth rate compared to (a)

2.2.1.2 Growth-rate dependence of transcription

In *E. coli*, the free RNAP (*i.e.* available for transcription) concentration is growth rate dependent in a positive manner but also saturates at high growth rates (above 2 doublings/hour) (Figure 2.8 A)[Klumpp and Hwa, 2008]. Moreover, the transcription of constitutively expressed (unregulated) promoters directly reflects the free RNAP concentration. Indeed, their relative transcription rate increases similarly to the free RNAP (Figure 2.8 A). In *B. subtilis*, the total RNA abundance also increases twofold when the growth rate doubles and the total mRNA abundance represents a constant fraction of the total RNA abundance at different growth rates [Borkowski et al., 2016]. Concerning transcription of rRNAs, in both bacteria, the increase in their expression level reflects a growth-rate-dependent regulation, isolated from the change in free RNAP concentration (Figure 2.8 B) [Klumpp and Hwa, 2008][Borkowski et al., 2016].

Moreover, the Promoter Activity (PA) of a constitutive gene can be described by a Michaelis-Menten type rate law as a function of the growth rate (μ) and two promoter-specific parameters V_{max} and K_m [Klumpp and Hwa, 2008][Gerosa et al., 2013], such that:

$$PA = V_{max} * \frac{\mu/K_m}{1 + \mu/K_m}$$

where V_{max} quantifies the maximal promoter activity sustained by the promoter and K_m the growth rate at which promoter activity is half-maximal (Figure 2.9). This relationship shows that there exists a maximum expression level that can be reached which is interpreted as promoter capacity. The governing parameters V_{max} and K_m are found to be promoter specific [Gerosa et al., 2013]. Consequently, the transcription machinery abundance specifically influences the expression of each gene according to their promoter sequence in a growth-rate dependent manner.

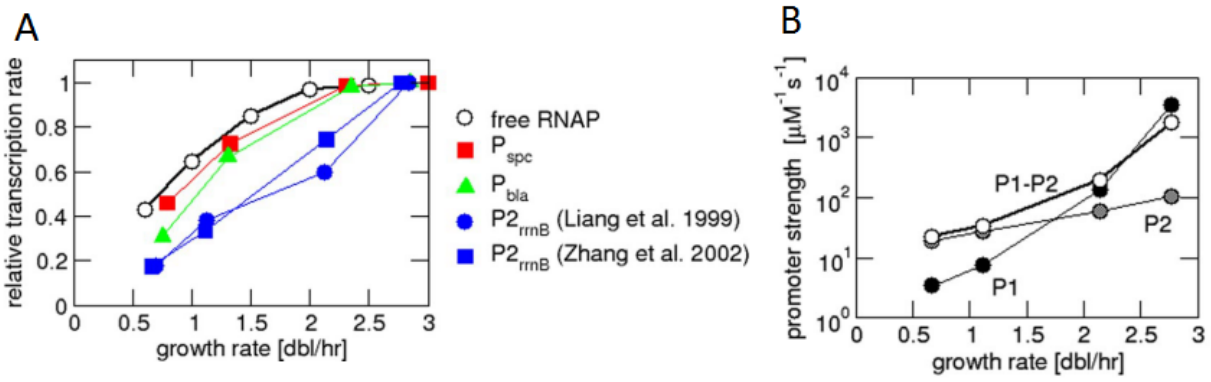


Figure 2.8: Growth-rate-dependent transcription from constitutive promoters. (adapted from [Klumpp and Hwa, 2008]). (A) Growth-rate dependence of the transcription rates from several constitutive promoters (*spc*, *bla*) and the *rrnB* promoter P2. The black curve indicates the free RNAP concentration, which is proportional to the predicted transcription rate from an unsaturated constitutive promoter. (B) Growth-rate-dependent regulation of the rRNA promoters: effective promoter strengths for the *rrnB* promoters P1(black), P2(gray), and the pair P1-P2 (white) as function of the growth rate.

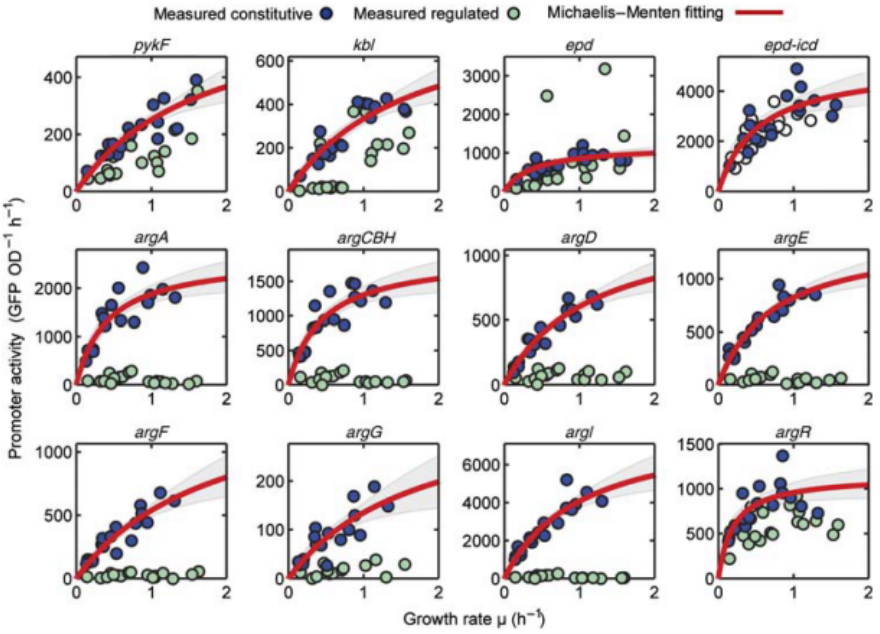


Figure 2.9: Global expression machinery regulation of promoter activity during exponential growth. (from [Gerosa et al., 2013]). Promoter activity of constitutive (blue dots) and native, specific regulated (green dots) promoters as a function of the steady-state growth rate under 18 nutritional conditions. Red lines show optimal least-square fitting of a Michaelis-Menten rate law.

2.2.1.3 Growth-rate dependence of translation

At the translation level, previous works on *E. coli* stated that global regulation alter the production of proteins only through growth-related dilution [Liang et al., 2000][Klumpp et al., 2009]. The translation efficiency, defined as the number of proteins produced per mRNA per hour, was thus considered constant with growth rate. But data obtained with high-resolution technologies suggest the contrary [Nicolas et al., 2012][Goelzer et al., 2015]. In *B. subtilis*, the translation efficiency drops when growth rate increases [Borkowski et al., 2016]. Furthermore, the level of production of a protein varies according to growth rate through their gene-specific TIR (Figure 1.5). Indeed, the transcript-specific translation efficiency (λ_i) depends on the concentration of the free ribosomes (*i.e.* ribosomes available to initiate translation) such that:

$$\lambda_i = \frac{K_{1i}[R_{free}]}{K_{2i} + [R_{free}]}$$

and

$$\lambda_i = \frac{\mu[P_i]}{[m_i]}$$

with two transcript-specific constants K_{1i} and K_{2i}; where R_{free} (*free* ribosome) abundance corresponds to the fraction of ribosomes ready to initiate translation; μ is the growth rate; $[P_i]$ is the protein concentration; and $[m_i]$ is the concentration of the mRNA translated to produce P_i. These relationships mean that a drop in translation efficiency (as observed) implies that R_{free} abundance decreases with increasing growth rate. Thus, an increased growth rate due to higher nutrient quality decreases the number of ribosomes available for translation and consequently negatively impacts the translation efficiency (in addition to growth dilution).

In conclusion, the translation efficiency appears to be either a constant [Liang et al., 2000][Klumpp et al., 2009] function of growth rate in *E. coli* and a decreasing function of growth rate in *B. subtilis* [Borkowski et al., 2016]. However, the experimental set ups used for collecting data on *E. coli* are very old and there should not be particular reasons for having such difference with *B. subtilis*. Thus, it is very likely that translation efficiency actually varies according to growth rate in *E. coli*. Moreover, for both bacterial species, the transcription efficiency is an increasing function of growth rate in both bacteria. Thus, Borkowski *et al.* (2016) propose that the global regulation of transcription and translation may allow prokaryotes to fine-tune the abundance of each protein as a function of the growth rate in absence of dedicated regulators (Figure 2.10).

2.2.2 Resource allocation: general approaches

Cells constantly face environmental changes and thus have to optimally reallocate their macromolecular resources to achieve survival and growth. This process is typically referred to as 'resource allocation' in the review of [Yang et al., 2018]. It has been proposed as the cornerstone for limiting growth rate under fast-growth conditions [Goelzer and Fromion, 2011]. Moreover, the production of certain proteins greatly changes between low and fast growth rates according to the cellular process they are

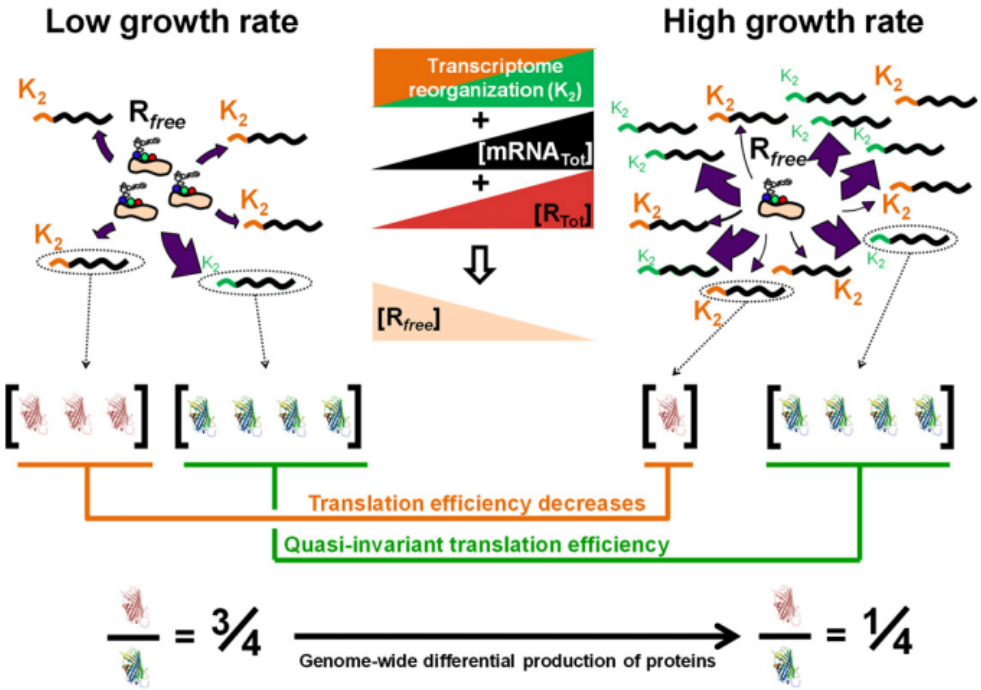


Figure 2.10: Fine-tuning of protein expression through global regulation (from [Borkowski et al., 2016]). Two different proteins (green and orange) are differentially produced according to growth rate. Concerning transcription, their respective mRNA levels differ according to growth rate: the transcription rate of the green protein is more increased at a higher growth rate as compared to the orange protein. When R_{free} concentration drops at a higher growth rate, the intrinsic properties of the proteins' TIR lead to a higher translation efficiency for the green protein as compared to the orange protein (green K_2 < red K_2). *In fine*, the ratio of orange protein on green protein is decreased from low to high growth rate (3/4 to 1/4) without the intercession of a dedicated regulator.

involved in [Goelzer and Fromion, 2011][Molenaar et al., 2009]. Indeed, each cellular process represents a certain cost for producing the proteins involved in it, referred to as "protein cost". This cost has been mathematically formalized to better understand how the cell optimally allocates its resources [Molenaar et al., 2009][Goelzer et al., 2009][Goelzer et al., 2011][Scott et al., 2010][Weiße et al., 2015].

The first approach which is constraint-based and consists in a self-replicator model, predicts a shift in growth strategies according to the nutrient availability [Molenaar et al., 2009]. A basic self-replicating system consists of one catalyst (a kind of ribosome) which synthesizes itself from a substrate. Thus, at constant substrate concentration this catalyst will exponentially replicate itself. The self-replicator model built by Molenaar *at al.* (2009) predicts that bacteria display a gradual shift instead of a switch from one to other type of metabolism according to growth rate. This prediction has been experimentally observed in *L. lactis* (Figure 2.11). This shift from metabolically to catalytically efficient metabolism when substrate concentration increases is the results of optimizing the cellular economy for growth rate. From this model, they were able to derive the first growth law.

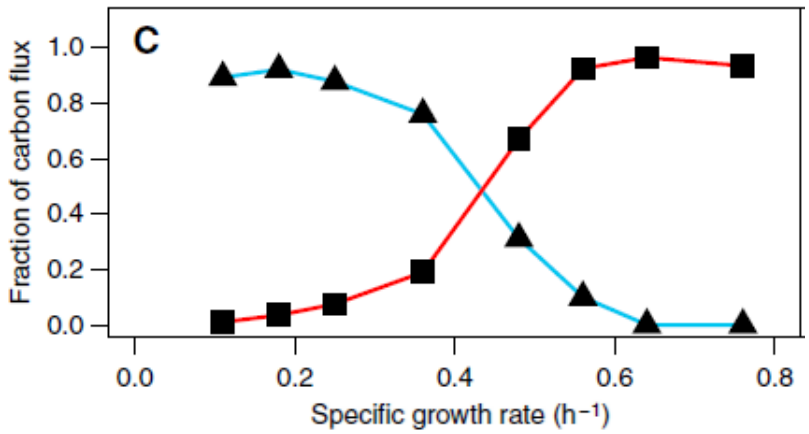


Figure 2.11: Shifting of optimal metabolic strategies (from [Molenaar et al., 2009]). *L. lactis* shifts from the metabolically efficient mixed-acid fermentation (blue) at low growth rates to lactic acid fermentation (red) at high growth rates.

The second model, a coarse-grain approach, implements three cellular trade-offs (*i.e.* finite energy, ribosomes and proteome) by considering two core biochemical processes: gene expression and nutrient import and metabolism [Weiße et al., 2015]. Thus, it is more complex than the model developped by Moleenar *et al.* since it also takes into account the transcription process through additional equations. Their model shows a hyperbolic dependence of growth rate on levels of extracellular nutrients and it implies that the growth rate is proportional to the ribosomal mass fraction. Thus, their model allowed to derive the two first growth laws.

A coarse-grain approach also predicts resource allocation of the proteome according to nutrient quality [Scott et al., 2010]. The total proteome can be partitioned into one growth-rate independent fraction and two growth-rate dependent fractions; one includes the ribosomal and other translational proteins and the other one includes metabolic proteins (*i.e.* transporters, catabolic and anabolic enzymes, etc.) (Figure 2.12 A). In addition, the ribosome-affiliated fraction exhibits a positive linear correlation with growth rate when it is modulated by the nutrient quality (Figure 2.12 A) as stated by the third growth law. Cell growth can also be controlled by translational inhibition using dedicated antibiotics. This leads to a negative correlation between ribosome abundance and growth rate (Figure 2.12 A). The ribosome-affiliated fraction can be subdivided into a ribosomal protein fraction and a tRNA- or translation speed-affiliated proteins [Klumpp et al., 2013] (Figure 2.12 B). These sub-fractions will evolve similarly (increase or decrease) according to the growth rate. These observations are referred to as the growth law of proteome partitioning in the review of [Jun et al., 2018] and these proteins partitions are similar to the one found in Goelzer *et al.* (2011).

Similarly to the model previously described, the model from Weiße *et al.* (2015) also explains in term of energy distribution the transcription rate of the different categories of genes. If the energy level is low, more enzyme-coding mRNAs (enzyme mRNAs) are expressed leading to more successful bindings

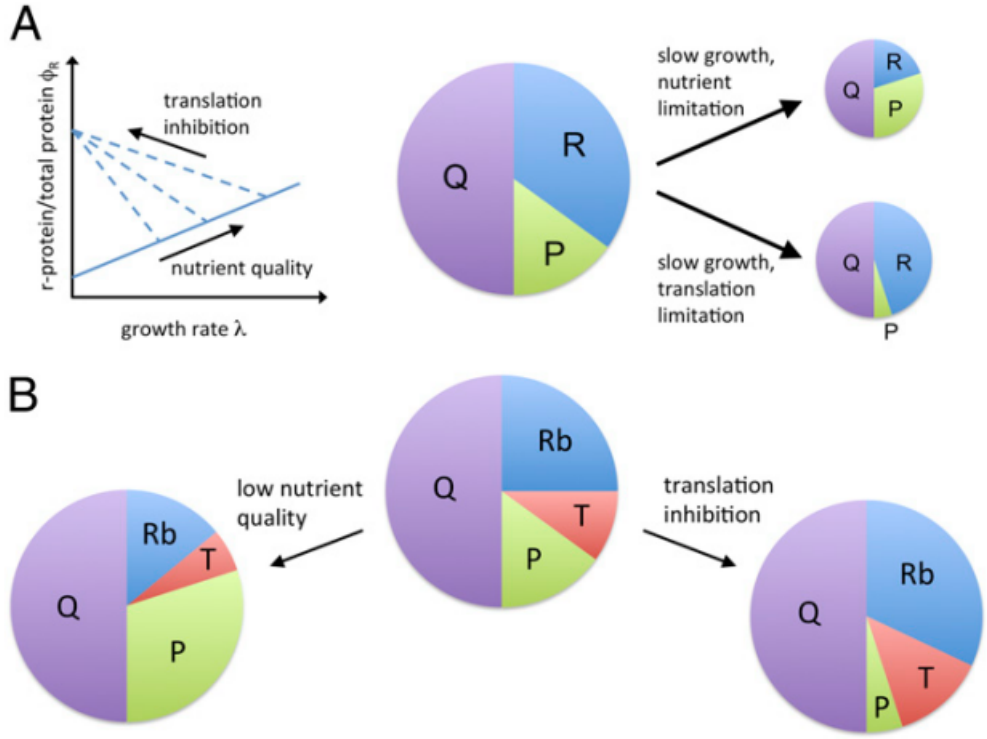


Figure 2.12: Coarse-grain approach: proteome partitioning (from [Klumpp et al., 2013]). (A) Three-component model of the proteome divided in three sectors: a fixed protein fraction (Q), a ribosome-affiliated fraction (R) and a metabolic as well as other proteins fraction (P). R exhibits a linear growth rate dependencies for nutrient variation and translation inhibition as sketched on the left. R and P have an opposite growth-rate dependence. (B) Four-component model of the proteome: the ribosome affiliated-fraction is split in two parts: a ribosomal protein fraction (Rb) and a fraction of tRNA- or translation speed-affiliated proteins (T), which increase and decrease together.

of enzyme mRNAs on ribosomes (P(e) in Figure 2.13). Conversely, if the enzyme level rises, then energy levels also increase and enzyme mRNAs are less successful in binding ribosomes; which leads to decreasing levels of enzymes. This is in accordance with the third growth law. Similarly, if ribosome levels decrease, then the translation rate decreases and thus energy rises leading to more ribosomal transcription (P(r) in Figure 2.13). On the opposite, an increase in ribosomes is counteracted by a decrease in ribosomal transcription due to changes in energy levels. These feedbacks set an equilibrium between energy influx and consumption, and *in fine* stabilize energy levels.

An other approach, the Resource Balance Analysis (RBA) [Goelzer et al., 2009][Goelzer et al., 2011][Goelzer and Fromion, 2011] is a method which is in an extension of the Flux Balance Analysis (FBA) framework, a genome-scale and constraint-based approach that analyses the flow of metabolites through a metabolic network [Orth et al., 2010]. RBA captures the resource allocation between cellular processes. It states that the distribution of resources among the different cellular processes is subject to four constraints:

i. The "Metabolic capability constraint": the capability of the metabolic network must be sufficient

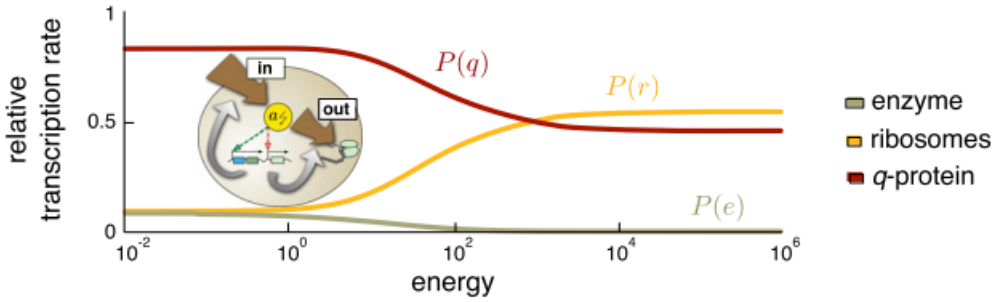


Figure 2.13: Coarse-grain approach: energy distribution between different mRNA categories (adapted from [Weiße et al., 2015]). (*Inset*) The brown arrows represent the flux of nutrients which provide the energy a (yellow round) for gene transcription; q-protein refers to house-keeping proteins like cytoskeletal proteins. The relative transcription rate which is plotted represents the ability of an mRNA to compete with ribosomes. The relative abundance of mRNA changes with the level of intracellular energy because of different transcriptional responses of ribosomal and non ribosomal genes.

to produce all the metabolic precursors required for biomass production.

ii. The "Translation capability constraint": the capability of the translation apparatus needs to be sufficient to keep the concentration of all the cell proteins constant at a given growth rate.

iii. The "Density constraint": the intracellular density must remain constant to ensure the suitable diffusion of all the cell components (meaning that the cytosolic density and the membrane protein occupancy are limited).

iv. The mass conservation law is satisfied.

Compared to the models of Moleenar *et al.* (2009) and Weiße *et al.* (2015), this model is more complex since it also integrates the metabolic pathways. This approach predicts the flux distribution, the maximal growth rate and the concentration of ribosomes, enzymes and transporters. The concentration of ribosomes increases with growth rate while the concentration of non ribosomal proteins (enzymes, transporters, ...) decreases (Figure 2.14). Consequently, the cell size and the translation machinery activity strongly increase with the growth rate. These predictions were experimentally validated [Goelzer et al., 2015].

Moreover, the RBA predictions and the coarse-grain approach show that production of unneeded protein decreases the fraction of proteins allocable for ribosomes (R fraction) and for the processes providing the nutrients needed for growth (P fraction) [Scott et al., 2010][Goelzer et al., 2009][Goelzer and Fromion, 2011]. This leads to a decrease in growth rate, but it can be seen as a way for the cell to prepare for a medium richer in nutrients by saving proteins with specific functions required for growth. A genome-scale model of *E. coli* also suggests that the cell pre-allocates its proteome toward alternative carbon sources, thus providing a fitness benefit when such sources are encountered [O'Brien et al., 2016]. These trade-offs between optimal repartition for maximizing growth but also preparing for nutritional changes have been studied through ME (Metabolism and macromolecular Expression)

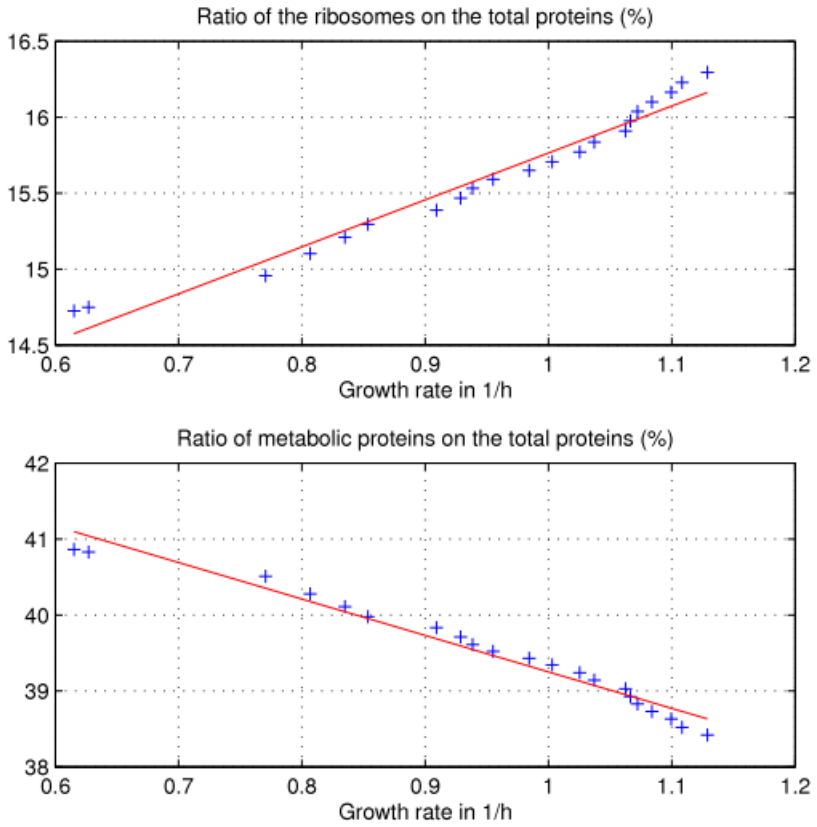


Figure 2.14: RBA prediction of the evolution of the ribosomes and metabolic proteins ratios as function of growth rate (from [Goelzer et al., 2011]). Ratio ribosomes/total proteins versus growth rate (top) and ratio metabolic enzymes/total proteins versus growth rate (bottom). The red line is the linear approximation of these predictions.

models [Yang et al., 2016]. They predict two proteome configurations: an "optimal" for maximizing growth versus a "generalist" which includes gratuitous proteins needed to anticipate potential stresses.

Experimentally, it has been observed that under favourable growth conditions, cells produce an apparent excess of protein which should help them to cope with potential stressful conditions such as a nutritional shift [Goelzer et al., 2015]. The metabolic processes which appear to express above demand are the central carbon pathway, the short metabolic pathways as well as the vitamin/cofactor and fatty-acids synthesis pathways [Goelzer et al., 2015]. Enzymes involved in sporulation are also likely to be "overproduced". Given that these gratuitous proteins represent less than 1% of the total proteome, ensuring extra flux through these pathways could provide a significant evolutionary benefit to the cell at a relatively low cost. Thus, cell resource allocation is a fine-tuned but complex process which consists in a trade-off between allocating proteins for growth optimization and for anticipating potential nutritional changes as well as other stresses.

In conclusion, these models provide similar results but they have different level of complexity. The simplest one is the self-replicator model from Moleenar *et al.* (2009) which is not specific to any

microorganism while the most complex model is the one of Goelzer *et al.* (2009 and 2011) which is specie specific and particularly focuses on explaining how *B. subtilis* optimally allocate its resources for survival. Moreover, these four different mathematical approaches were able to predict the three empirical growth laws.

Overview: To optimally allocate resources among the different cellular processes according to growth rate, global regulations occur at the level of DNA replication, transcription and translation. This can be achieved through the intrinsic properties of the genes' promoter and TIR sequences which allow the cell to fine-tune the abundance of each protein as function of the growth rate in absence of dedicated regulators.
Different modelling approaches (coarse-grain or constraint-based) have been proposed to better understand how resource allocation affects the cell growth rate during the steady-state phase. The self-replicator model of Molenaar *et al.* (2009) shows that bacteria display a graduate shift to switch from one type of metabolism to another. The coarse-grain approach of Weiss *et al.* (2015) explains that there exists feedback regulations which lead to an equilibrium between energy influx and consumption across the different cellular processes that impacts the level of the different RNAs. The resource balance analysis of Goelzer *et al.* (2009 and 2011) demonstrate that the cell size and the translation machinery activity strongly increases with the growth rate. The coarse-grain model of Scott et *et al.* (2010) divides the proteome in three fractions and states that the fraction of ribosome-affiliated proteins exhibits a positive linear correlation with growth rate when it is modulated by the nutrient quality. Certain of these models also predict that the bacteria produce "gratuitous" proteins that can help them to cope with potential environmental changes.
These four different mathematical approaches were able to predict the three empirical growth laws. The simplest model from Molenaar *et al.* (2009) provides predictions that can account for any microorganisms; while the more complex one from Goelzer *et al.* (2009 and 2011) is *B. subtilis* specific since it takes into account its metabolic pathways.

2.3 The molecular mechanisms used by the cell to adapt to nutritional changes

As seen previously, there exists global regulations which allow the cell to adapt to different nutritional conditions. But how these regulations take place at the molecular level? The following part details the molecular mechanisms which allow the cell to reorganise its macromolecular content.

2.3.1 The key metabolites to bacterial adaptation: the alarmones GTP and (p)ppGpp

The abundance of the phosphorylated metabolite GTP positively correlates with steady-state growth rate and tunes growth in a dose-dependent manner (Figure 2.15 (a)) [Bittner et al., 2014]. When the cell faces amino acid depletion, growth is slowed down or even stopped (Figure 2.15 (c)). This leads to

a reduced rate of protein synthesis and RNA accumulation [Maaløe and Kjeldgaard, 1966] [Neidhardt, 1966]. This phenomenon is referred to as "stringent response". Almost 50 years ago, Cashel and Gallant studied the change in phosphorylated metabolites abundance after an amino acid starvation during *E. coli* growth [Cashel and Gallant, 1969]. From extracts of *E. coli* responding to the stress of amino acid starvation, two spots appeared on radioautograms (Figure 2.15 (b) and (c)). These spots, first called magic spots, are derivatives of GTP and GDP that differ only by the presence of a pyrophosphate esterified to the ribose 3' carbon, abbreviated as pppGpp and ppGpp respectively. They are together designated as "(p)ppGpp". Since then, further works have been done to describe the stringent response mechanisms.

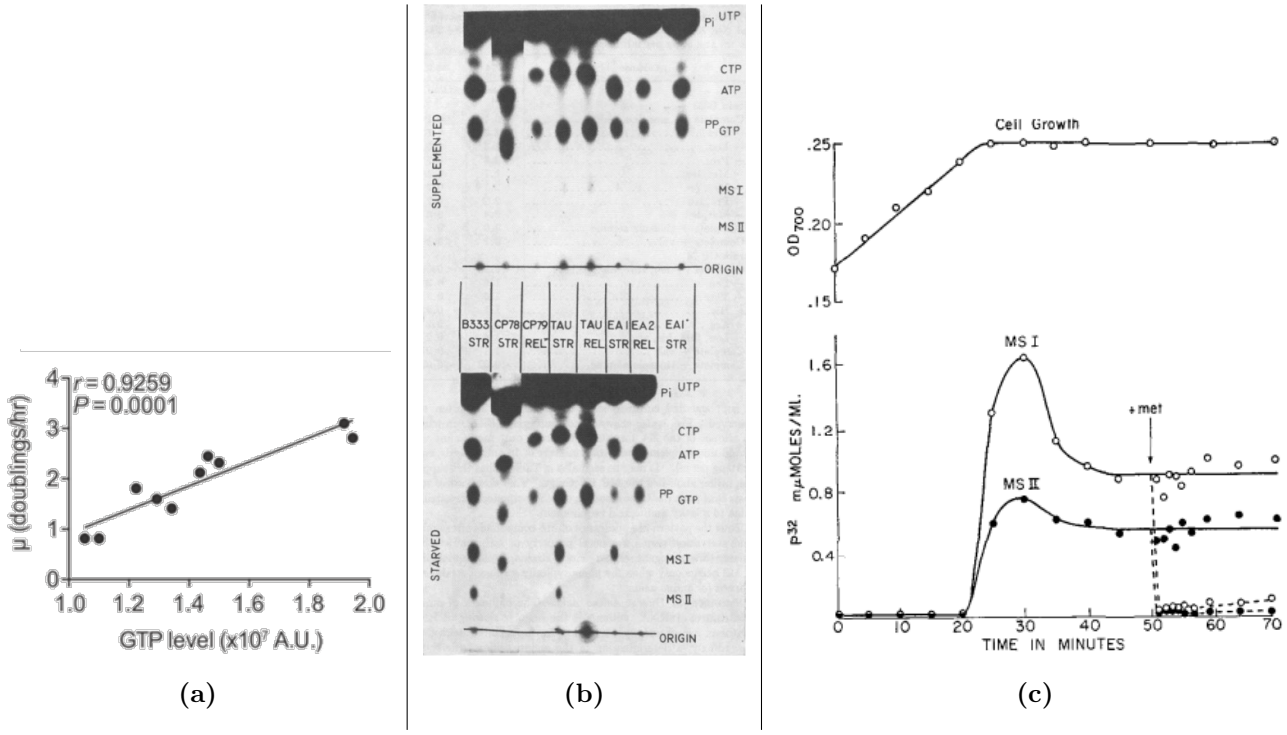
More precisely, upon amino acid shortage, the number of uncharged tRNAs stalled in the ribosomal A-site increases, which signals to the ribosome-associated RelA (a protein possessing (p)ppGpp synthase activity in Gram negative bacteria and both (p)ppGpp hydrolase and synthase activity in Gram positive bacteria, see section 2.3.2.2) to synthesize (p)ppGpp from GTP [Wendrich et al., 2002][Potrykus and Cashel, 2008]. In *E. coli* and more generally in Gram negative bacteria, the increased level of (p)ppGpp signals to the RNAP to stop the transcription of genes involved in the translation machinery. Consequently, the RNAP starts to preferentially transcribe biosynthetic genes which activate pathways involved in amino acid synthesis [Potrykus and Cashel, 2008][Dalebroux and Swanson, 2012]. Conversely, when there are many nutrients available and so amino acids, (p)ppGpp is degraded by SpoT (a protein possessing hydrolase and synthase activities and able to sense many sources of nutrient stress in gram negative) or by RelA (for Gram positive like *B. subtilis*) and thus the RNAP is directed to genes that are crucial for bacterial replication such as the ones encoding tRNA and rRNA in *E. coli* [Potrykus and Cashel, 2008][Dalebroux and Swanson, 2012]. How these regulations take place will be discussed in details in this part.

2.3.2 Biosynthesis of the alarmones GTP and (p)ppGpp

2.3.2.1 Biosynthesis of GTP

Two biological pathways lead to GTP production[1]: one is the "*de novo* pathway" and the other one is the "salvage pathway". For the *de novo* pathway, the D-ribose 5-phosphate is converted into 5-phospho-α-D-ribose 1-diphosphate (PRPP) by the phosphoribosylpyrophosphate synthetase (Prs)) which leads to the production of inosine 5'-phosphate (IMP) after multiple reaction steps (Figure 2.16). IMP can also be produced from hypoxanthine, a reaction catalyzed by the hypoxanthine phosphoribosyltransferase (HprT). In *E. coli* but not in *B. subtilis*, IMP can be directly produced from inosine by the Guanosine kinase (Gsk). Then, the IMP dehydrogenase GuaB catalyzes the reaction from IMP to xanthosine 5'-phosphate (XMP) that is then transformed into guanosine 5'-monophosphate (GMP) by the GMP synthase GuaA. XMP can also be produced through xanthine by the action of HprT. The salvage pathway consists in the synthesis of guanine from guanosine

[1]pathways' description obtained from the websites https://ecocyc.org/ and https://bsubcyc.org/



(a) **(b)** **(c)**

Figure 2.15: The discovery of the "magic spot". (a) Pearson correlation of growth rate with GTP level in a mutated *B. subtilis* strain where the intracellular GTP concentration could be modulated (adapted from [Bittner et al., 2014]). (b) Radioautogram of *E. coli* extracts in a medium supplemented with all amino acids (up) or with one amino acid lacking (down) (c) Growth of *E. coli* cells that are starved for methionine (up) and evolution of the level of ppGpp (MSI) and pppGpp (MSII) during this growth (down). The growth arrest correlate with the strong increase in (p)ppGpp alarmones. The figures (b) and (c) were adapted from [Cashel, 1969].

(reaction catalysed by the purine nucleoside phosphorylases PupG and DeoD) that reacts with PRPP to lead to GMP. This last reaction is also catalysed by HprT. Guanosine can come from the growth medium, and its entry is made possible by the hypoxanthine/guanine permeases PbuO and PbuG. Then, GMP is tranformed into guanosine 5'-diphosphate (GDP) via the GMP kinase Gmk. Finally, GDP is converted to GTP by NDP kinase (Ndk). It is important to note that hypoxanthine is also produced from inosine whose precursors is adenosine (derived from adenine) and recent results suggest that the purine pathway compete with the GTP synthesis pathway for substrates [Bittner et al., 2014].

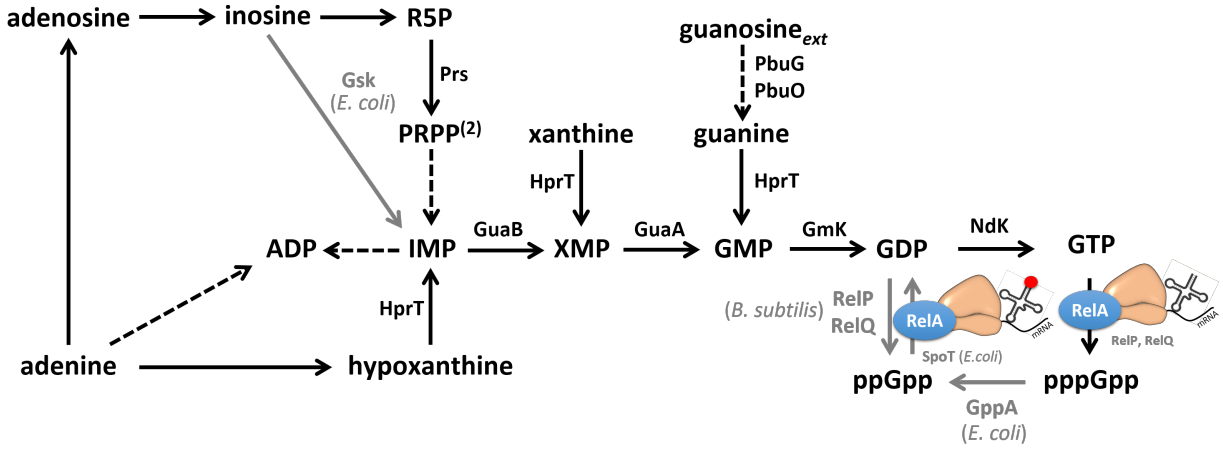


Figure 2.16: Biosynthesis of the alarmones GTP and (p)ppGpp. The full black arrows represent reactions catalysed by the enzyme/transporter mentioned next to it. The dashed arrows represent a pathway with at least two reaction steps. In grey are represented the reactions and their respective enzymes that are specific to either *E. coli* or *B. subtilis*.

2.3.2.2 Biosynthesis of (p)ppGpp

The synthesis of (p)ppGpp does not require the same enzymes in Gram negative and Gram positive bacteria. In *E. coli*, (p)ppGpp can be synthesized by two different enzymes: RelA and SpoT which respond to different stress signals [Hauryliuk et al., 2015]. SpoT is a bifunctional enzyme since it can both synthesize and hydrolyse (p)ppGpp. It senses a limitation of several nutrients (carbon sources, phosphate, iron and fatty acids). The *E. coli* RelA has only a (p)ppGpp synthetase activity and senses amino acid limitation (Figure 2.16). As compared to RelA, SpoT has a weak synthase activity and strong hydrolytic activity for (p)ppGpp. *E. coli* also possesses the GppA enzyme which hydrolyzes the 5'-γ-phosphate of pppGpp to form ppGpp [Steinchen and Bange, 2016].

In *B. subtilis*, the main enzyme which produces (p)ppGpp is RelA, considered as a SpoT homologue and part of the RelA/SpoT Homology (RSH)-type. It is a long RSH-type synthetase which is a multi-domain protein that can be divided into a catalytic domain (RSH-NTD) followed by a carboxy-terminal domain (RSH-CTD) [Steinchen and Bange, 2016]. RSH-NTD consists of a hydrolase followed by a synthetase domain, which are active according to the amino acid availability. The synthetase

domain is active in the case of the stringent response, while the hydrolase domain is active under favourable conditions. In its synthetase domain, RelA catalyzes phosphate transfer from ATP to the 3' OH group of the ribose moiety of GDP and GTP to synthesize ppGpp and pppGpp respectively (Figure 2.16). RelA also possesses an hydrolase activity which degrades ppGpp by removing the pyrophosphate from the 3' position of the ribose moiety, leading to GDP production.

In *B. subtilis*, there also exists Short RSH-type Synthetases/Small Alarmone Synthetases (SAS) to produce (p)ppGpp in *B. subtilis*: SAS1 (referred to as YjbM/RelQ) and SAS2 (referred to as YwaC/RelP) [Nanamiya et al., 2008]. They share similarity of nearly 50% on the amino acid sequence level. Moreover, they share a highly similar (p)ppGpp synthetase domain and both establish highly similar homotetrameric complexes [Steinchen et al., 2018].
Their activity depends on other stress signals that have not been yet fully determined. RelP is much more active than RelQ for producing (p)ppGpp. RelP always has a high synthetase activity state while RelQ can switch from a "passive state" (low (p)ppGpp synthetase activity) to an "active state" (high (p)ppGpp synthetase activity) [Steinchen et al., 2018]. RelQ is predominantly transcribed during logarithmic growth [Nanamiya et al., 2008]. It has been suggested that its activity is stimulated by amino acid starvation and also by the cell energy imbalance (i.e. if there is a great excess of GDP) [Arenz et al., 2016] and its activity could be intensively coupled to RelA's activity [Steinchen et al., 2018]. Concerning RelP, the transcription of its corresponding gene *ywaC* is likely to be dependent on σ_M, an extracytoplasmic function (ECF) sigma factor that is activated by various stress conditions that affect the cell wall [Tagami et al., 2012]. Furthermore, RelP transcripts only appear during early stationary phase and under specific treatments such as the addition of antibiotics, ethanol, high salt and acidic or alkalic pH stress conditions [Geiger et al., 2014][Thackray and Moir, 2003][Zweers et al., 2012].

2.3.2.3 How the stringent response triggers the synthase activity of RelA

When RelA is not bound to the ribosome and in the absence of a deacylated tRNA, RelA exists in an autoinhibited state that produces only low levels of (p)ppGpp [Potrykus and Cashel, 2008][Hauryliuk et al., 2015]. This autoinhibition of RelA activity is likely to be due to the interaction between its C-terminal domain (CTD) and its N-terminal domain (NTD), referred to as the 'closed' conformation [Arenz et al., 2016].
Data suggest that RelA binding to the ribosome is strong under favourable conditions while it is reduced when (p)ppGpp is synthesized under starvation conditions [Wendrich et al., 2002]. Thus, a "hopping" model has been proposed where RelA is hopping from one blocked ribosome to another ribosome bearing a 3' extension of the mRNA and a deacylated tRNA at the A site [Wendrich et al., 2002] (Figure 2.17).
English *et al.* (2011) have proposed an "extended hopping" model where many (p)ppGpp molecules are produced when RelA is rather off than on the ribosome. After producing several (p)ppGpp molecules,

RelA binds again another ribosome and if a deacylated cognate tRNA is detected in the A-site then it starts producing (p)ppGpp after dissociating from the ribosome.

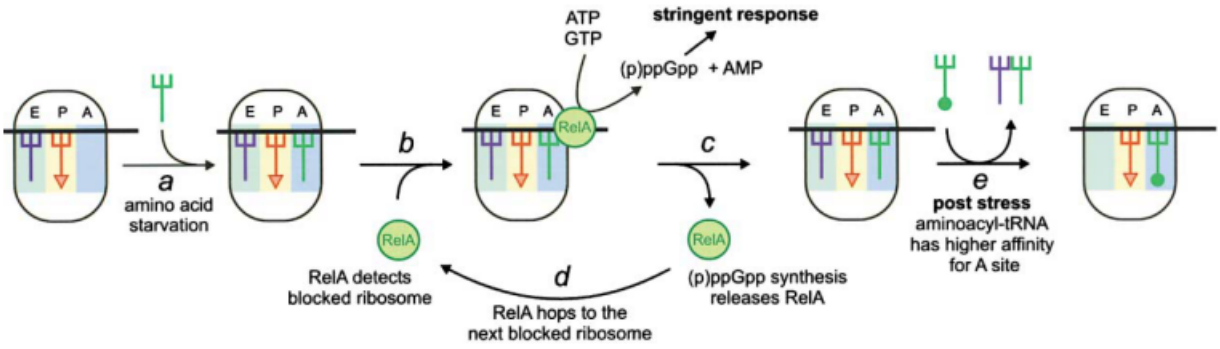


Figure 2.17: Hopping model of the RelA-mediated (p)ppGpp synthesis (from [Wendrich et al., 2002]) (a) Upon amino acid starvation, large pools of deacylated tRNAs are produced and they bind to the ribosomal A-site with low affinity and block the ribosome. (b) RelA detects a blocked ribosome. (c) In presence of a a deacylated tRNA at the A-site, RelA produce (p)ppGpp from GTP/GDP through conversion of ATP to AMP. While synthesizing (p)ppGpp, RelA is released but not the A-site-bound deacylated tRNA. (d) Then RelA can "hop" to the next blocked ribosome, and the synthesis of (p)ppGpp is repeated. The production of high levels of (p)ppGpp will then activate the stringent response. (e) Following post-stress conditions, aminoacylated tRNAs are replenished. An aa-tRNA has a higher affinity over a deacylated tRNAs for the A-site. This enables displacement of the deacylated tRNAs, and rescues blocked ribosomes which can then continue translation.

However, recent works have contradicted this extended model [Li et al., 2016][Loveland et al., 2016][Kudrin et al., 2018]. The ribosome binding of RelA is actually stronger under starvation conditions using similar fusion protein constructs as in English *et al.* (2011) [Li et al., 2016]. Moreover, the (p)ppGpp synthesis by RelA only occurs when a cognate deacylated tRNA interacts with mRNA-programmed ribosomes in the A-site [Loveland et al., 2016]. In particular, the ribosomal protein L11 is required for triggering (p)ppGpp synthesis, even if it is not needed for the ribosome binding of RelA. Upon binding of cognate deacyl-tRNA to the ribosome, the RelA synthetase domain is exposed in the vicinity of the 30S spur (or "needle") and is activated for (p)ppGpp synthesis by alleviation of RelA autoinhibition and interactions with the ribosome (Figure 2.18) [Loveland et al., 2016]. RelA can bind a stalled ribosome before the arrival of a deacylated cognate tRNA in the A-site. But, as also stated by Wendrich *et al.* (2002), the deacyl-tRNA binding is required to stabilize the extended RelA conformation. Furthermore, RelA does not form a stable complex with a deacylated tRNA and is not activated by it when it is off the ribosome [Kudrin et al., 2018]. This favours a model where RelA first binds the empty A-site and then recruits the tRNA, leading to (p)ppGpp production while RelA is on rather than off the ribosome.

Structurally, the RelA binding onto the ribosome and the presence of a deacylated t-RNA in the A-site lead to the releaving of the interaction that the RelA CTD exerts on RelA NTD [Arenz et al.,

2016][Brown et al., 2016]. Thus, the NTD is uninhibited and therefore it can catalyse the synthesis of (p)ppGpp from GTP/GDP and ATP. Concerning how RelA detects that the tsRNA is empty, it seems that the TGS (Thr-RS, GTPase and SpoT) domain of RelA contacts the 3'-end of the deacylated tRNA [Arenz et al., 2016]. A very recent study also favours a model in which (p)ppGpp synthesis occurs when RelA is stably bound to the ribosome [Winther et al., 2018]. More precisely, RelA interacts with uncharged tRNA off the ribosome and this interaction is prerequisite to ribosome binding and activation of RelA's (p)ppGpp synthetic activity.

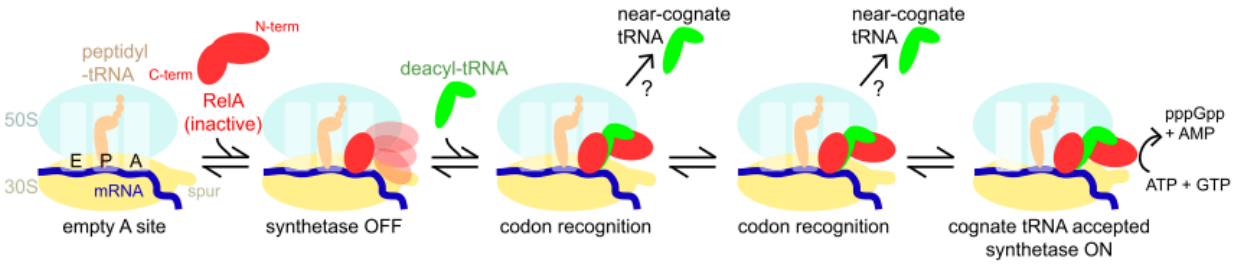


Figure 2.18: Second model of the mechanism of RelA activation by the ribosome and cognate deacyl-tRNA (adapted from [Loveland et al., 2016]). RelA, in its inactivated state, binds to the ribosome where the A-site is empty. Then a deacylated-tRNA reaches the A-site and will be accepted by the ribosome if it is cognate to the A-site codon. Since the tRNA is empty, this will trigger the synthetase activity of RelA, leading to (p)ppGpp production.

2.3.3 Effects of (p)ppGpp on enzymes synthesizing GTP and (p)ppGpp

In *B. subtilis*, (p)ppGpp strongly inhibits GmK as well as HprT, and slightly GuaB which are enzymes involved in GTP synthesis (Figure 2.22) [Lopez et al., 1981][Kriel et al., 2012][Liu et al., 2015b]. GmK is the enzyme catalysing the reaction from GMP to GDP and thus it is an important target to efficiently decrease the GTP level [Liu et al., 2015b]. HprT catalyses different reactions involved in GTP synthesis and this is why it is also a favoured target for (p)ppGpp [Kriel et al., 2012]. GuaB is less strongly inhibited by (p)ppGpp but remains a strategic target to decrease GTP abundance [Lopez et al., 1981][Kriel et al., 2012]. Consequently, the high production of (p)ppGpp during the stringent response lead to the inhibition of enzymes involved in GTP production and as a consequence the GTP level drops [Cashel and Gallant, 1969][Cashel, 1969][Kriel et al., 2012].

The IMP dehydrogenase, GuaB, involved in GTP biosynthesis was initially identified as the main target of (p)ppGpp in *E. coli* [Gallant et al., 1971][Pao and Dyes, 1981]. (p)ppGpp also inhibits the activity of PRT-I enzymes involved in purine and pyrimidine salvage pathways such as the xanthine-guanine phosphoribosyltransferase (Gpt) which is involved in the first step of the subpathway that synthesizes XMP from xanthine and the hypoxanthine PRTase HprT, similarly to *B. subtilis* [Kanjee et al., 2012][Gaca et al., 2015].

Moreover, (p)ppGpp also enhances its own production in both Gram positive and negative. In *B. subtilis*, pppGpp (but not ppGpp) facilitates the synthesis of (p)ppGpp by SAS1/RelQ and thus stimulates its own production [Steinchen et al., 2015]. In *E. coli*, pppGpp fine-tune RelA's synthetic activity by stimulating the production of ppGpp from GDP as well as its own production from GTP [Kudrin et al., 2018]. *In fine*, this pppGpp's stimulation maximizes the overall (p)ppGpp production by RelA.

2.3.4 Effects of (p)ppGpp on the DNA replication process

Observations made in the early 1990s conclude that in bacteria, (p)ppGpp synthesis is required in order to couple DNA replication to the growth rate that is slowed down during the stringent response [Chiaramello and Zyskind, 1990]. Upon starvation, (p)ppGpp slows down or even stops DNA replication by direct inhibition of the essential replication component primase DnaG in *B. subtilis*, an enzyme which synthesizes oligonucleotides needed to start DNA synthesis [Wang et al., 2007]. Inhibition of the primase activity should cause a decrease in both lagging- and leading-strand synthesis of chromosomal DNA. Similarly, (p)ppGpp also inhibits the activity of *E. coli* DnaG primase *in vitro*; ppGpp being more efficient than pppGpp to inhibit the protein activity [Maciag et al., 2010] [Maciag-Dorszynska et al., 2013]. But strong evidences for *in vivo* inhibition of *E. coli* DnaG still lack [DeNapoli et al., 2013].

In addition, Wang *et al.* (2007) suggest that binding of (p)ppGpp to DnaG might result in allosteric inhibition of other components of the replication complex through protein-protein interactions. Such possibility is supported by DeNapoli *et al.* (2013) and could explain why the DNA replication elongation is more strongly inhibited in *B. subtilis* than in *E. coli*.

Concerning elongation, it appears that the DNA replication elongation rate in *E. coli* is reduced upon amino acid starvation and that this reduction requires (p)ppGpp [DeNapoli et al., 2013]. Moreover, inhibition of replication elongation by (p)ppGpp is found to be dose-dependent in both *E. coli* and *B. subtilis*.

2.3.5 Effects of GTP and (p)ppGpp on the transcription process

GTP and (p)ppGpp influence transcription by different mechanisms and have antagonistic roles. They can directly target specific proteins such as transcription factors or impact the GTP/ATP ratio which influences the level of genes' expression according to their transcription start site (TSS).

2.3.5.1 Interaction of GTP with CodY to control transcription initiation of target genes in *B. subtilis*

In *B. subtilis*, the GTP pool influences transcription of a range of biosynthetic genes through its interaction with the protein CodY, which has been first identified as a repressor of the *B. subtilis* dipeptide permease operon (*dppABCDE*) [Slack et al., 1993][Slack et al., 1995]. *In vitro*, CodY is a GTP-binding protein that senses the intracellular GTP concentration as an indicator of nutritional

conditions and regulates the transcription of early stationary phase and sporulation genes, allowing the cell to adapt to nutrient limitation [Ratnayake-Lecamwasam et al., 2001]. Concerning the affinity of CodY for target DNA *in vitro*, it is increased upon GTP addition, and not with other nucleotides [Handke et al., 2008]. However, CodY has a relatively low affinity for GTP but this allows the protein to distinguish between 2 to 3 mM GTP (the concentration in fast growing cells) and 300 μM GTP (the concentration in early-stationary phase). The CodY activation by GTP has been confirmed *in vivo* through genetic evidences in mutant strains [Brinsmade and Sonenshein, 2011], and in particular for repressing the expression of certain genes involved in the synthesis of Branched-Chain Amino Acids (BCAAs) [isoleucine, leucine and valine (ILV)] [Kriel et al., 2014].

CodY directly or indirectly regulates more than 100 genes, the products of which are generally involved in the adaptation of bacteria to media poor in nutrients required for growth [Molle et al., 2003]. Genes encoding proteins involved in the biosynthesis of branched-chain amino acids are directly regulated by CodY (*ilvBHC-leuABCD* operon and the *ilvA*, *ilvD*, and *ybgE* genes) [Shivers and Sonenshein, 2004][Tojo et al., 2008] as well as the arginine biosynthesis genes *argG* and *argJ* [Kriel et al., 2014] (Figure 2.19). CodY mediates regulation of the *ilvB* operon by GTP and BCAAs and bind to the *ilvB* promoter region [Shivers and Sonenshein, 2004]. In the presence of BCAAs, its affinity for promoters of BCAA biosynthetic genes is increased. But the expression of the *ilv-leu* operon is derepressed during the stringent response through detachment of the CodY protein from its *cis* elements upstream of the *ilv-leu* promoter [Tojo et al., 2008]. This detachment of CodY is due to the drop in GTP levels. The same occurs for genes involved in the transport of amino acids, amino sugars and dipeptides [Shivers and Sonenshein, 2004].
The binding sites of CodY have been identified on three different genes respectively involved in protein degradation (*ispA*) and sporulation initiation (*rapA* and *rapE*) [Belitsky and Sonenshein, 2013].

Furthermore, the gene encoding the GuaB enzyme involved in GTP synthesis, is also regulated by CodY and thus, GTP may influence its own production through its interaction with CodY [Molle et al., 2003]. This possibility is supported by the fact that in a mutant strain deprived of CodY and of genes involved in (p)ppGpp production (*i.e.* the strain cannot produce (p)ppGpp) the overall growth rate is reduced, likely because deletion of CodY decreases GTP level [Kriel et al., 2012][Kriel et al., 2014].

2.3.5.2 GTP and ATP levels influence transcription through the TSS

In *B. subtilis*, GTP concentration influences the transcription level of certain genes according to their TSS (Figure 2.20)[Krásnỳ and Gourse, 2004]. rRNA promoters are directly regulated by the concentration of GTP and all initiate with it through a G residue at their TSS. Given that RNA synthesis is the rate-limiting step in ribosome synthesis [Henkin and Yanofsky, 2002] [Paul et al., 2004b] and that all rRNAs promoters initiate with GTP [Krásnỳ and Gourse, 2004], this makes GTP

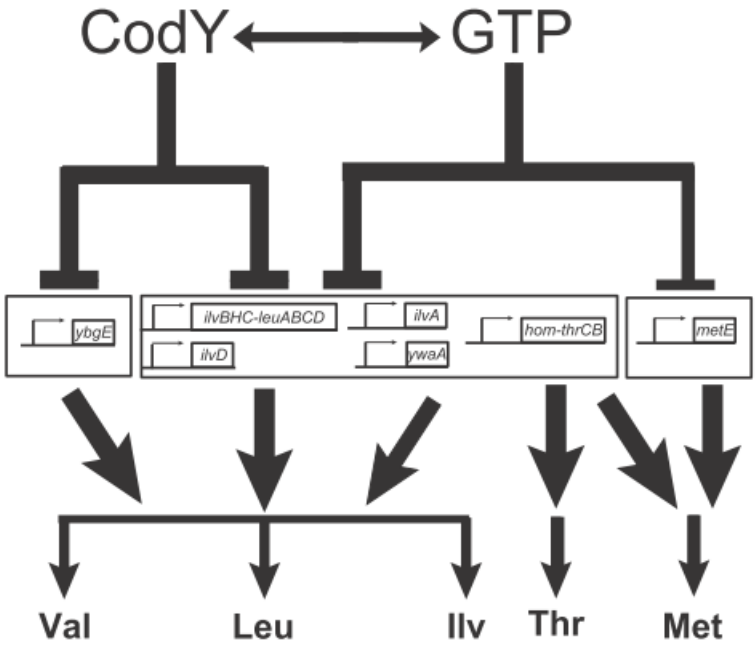


Figure 2.19: The transcriptional regulation by GTP and CodY of genes encoding enzymes involved in amino acid synthesis (adapted from [Kriel et al., 2014]). The transcription of *ybgE* is controlled by CodY, transcription of the *ilvB* and *ilvD* operons is controlled primarily by CodY but is also affected by GTP levels, and transcription of *ywaA* and *ilvA* is likely directly regulated by CodY but also controlled by GTP levels. Transcription of the *hom-thrCB* operon is likely directly regulated by CodY but is also strongly affected by GTP levels. Transcription of *metE* is also strongly affected by GTP levels. The operons *ybgE*, *ilvB*, *ilvD*, *ywaA* and *ilvA* encodes enzymes involved in the biosynthesis of valine (Val), leucine (Leu) and isoleucine (Ilv). These BCAAs [isoleucine, leucine and valine (ILV)] also interact with codY and thus repress their own transcription. The operons *hom-thrCB* and *metE* encode enzymes involved in the biosynthesis of respectively threonine (Thr) and methionine (Met).

concentration one of the driving forces for ribosome production and *in fine* growth [Bittner et al., 2014]. In *E. coli*, the rRNA transcription is not NTP specific, it can initiate with different ATP, GTP or even cytosine 5'-triphosphate (CTP) [Haugen et al., 2008].

In *B. subtilis*, other promoters linked to growth enhancement possess a guanine as TSS and their expression positively correlates with GTP concentration: the shared promoter P*str* of the genes encoding the ribosomal proteins RpsL and RpsG; the *tuf* gene which encodes EF-Tu [Krásnỳ et al., 2008] ; promoters of the *ptsGHI* and *pdhABCD* operons which respectively encode enzymes involved in glucose and pyruvate metabolism [Tojo et al., 2010]. Conversely, GTP can negatively affect transcription through the TSS when it is an adenine (Figure 2.20); especially for genes coding for enzymes involved in amino acid production: *ilvB*, *ywaA*, the BCAA biosynthesis genes *ilvA* and *ilvD* as well as the the threonine and methionine biosynthesis genes *hom-thrCB* and *metE* [Krásnỳ et al., 2008][Kriel et al., 2014] (Figure 2.19). These genes are predicted to have an adenine as TSS and thus tend to be activated by an increased ATP level (Figure 2.20) [Krásnỳ et al., 2008][Tojo et al., 2010][Kriel et al., 2014].

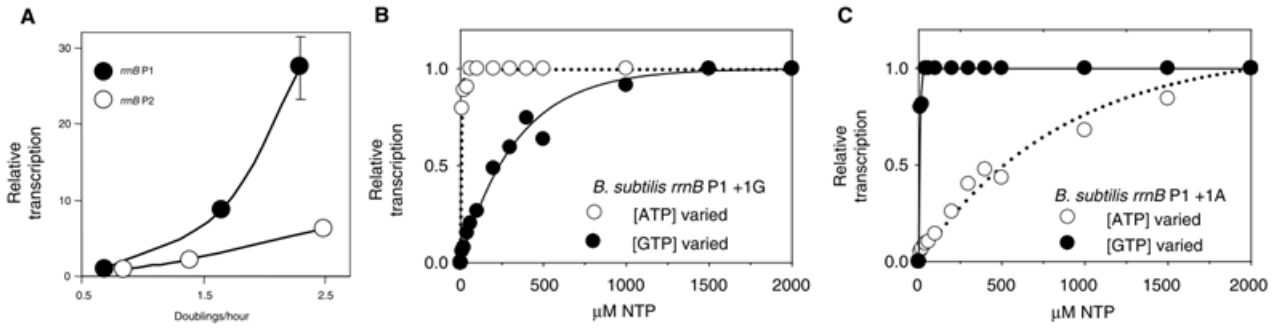


Figure 2.20: The growth rate and GTP-dependent control of *B. subtilis rrnB* promoters (adapted from [Krásnỳ and Gourse, 2004]). (A) The promoter activities were normalized to the activity of its own promoter at the lowest growth rate (the first point is equal to 1 for the minimal growth rate). The plot represents the *B. subtilis rrnB* promoters' relative transcription as function of growth rate. (B) and (C) *In vitro* experiments show how the relative transcription rate evolves with the NTP (GTP or ATP) concentration. In (B) the TSS of *rrnB* is G like for the native promoter while in (C) the TSS of *rrnB* is A.

Indeed, during the stringent response, the ATP and GTP levels change reciprocally and upregulated promoters initiate mostly with ATP while downregulated promoters initiate mostly with GTP [Krásnỳ et al., 2008]. The rRNA promoters which possess a guanine as TSS are less transcribed when the (p)ppGpp pool is increased [Natori et al., 2009]. The *ptsGHI* and *pdhABCD* operons encoding the sugar phosphotransferase system and the pyruvate dehydrogenase complex, which have a guanine as TSS, are also negatively regulated during the stringent response [Tojo et al., 2010]. Other genes and operons for which the TSS is adenine are found to have a transcription level which is upregulated under starvation conditions: the *ilv−leu* operon [Tojo et al., 2008] (Figure 2.19), the *pycA* gene encoding pyruvate carboxylase and the *alsSD* operon for synthesis of acetoin from pyruvate [Tojo et al., 2010].
This effect might also be completed by other complex regulatory mechanisms since the promoter activity of the *ilv* genes is not strongly upregulated with ATP under starvation conditions even if it has an adenine as TSS [Krásnỳ et al., 2008] (Figure 2.19).

The nucleotide found at +2 position also appears to influence transcription during the stringent response [Tojo et al., 2010]. Tojo *et al.* (2010) propose that a rate-limiting transcription initiation step might involve the formation of the first phosphodiester bond between the nucleotides at positions 1 and 2, which is assumed to be most affected by the concentrations of GTP and ATP, which greatly change during stringent response.

In conclusion, there exists strong evidences in *B. subtilis* that the concentration of GTP and ATP influences transcription levels through the promoter region sequence, and more particularly through the +1 (or even +2) position. Moreover, it appears that the mechanisms for regulating transcription

through NTPs concentration tend to differ between *E. coli* and *B. subtilis*. Indeed, (p)ppGpp directly affects transcription initiation in *E. coli* by interacting with the RNAP as described in the next part.

2.3.5.3 Direct interaction between (p)ppGpp and the RNAP to redirect transcription in *E. coli*

In *E. coli*, (p)ppGpp interacts directly with the RNA Polymerase (RNAP) to affect transcription of a broad range of promoters during the stringent response and other stress events [Chatterji et al., 1998] [Toulokhonov et al., 2001]. The binding site of (p)ppGpp on the RNAP is at an interface of the two RNAP subunits ω and β' suggesting an allosteric mechanism of action involving restriction of motion between two mobile RNAP modules [Ross et al., 2013][Zuo et al., 2013]. From these observations, it is proposed that the (p)ppGpp binding could potentially facilitate the escape of DNA from these complexes if the time required for nucleotide addition becomes longer than the lifetime of the open complex; leading to an inhibitory effect on the initial RNA synthesis [Zuo et al., 2013]. Consequently, the RNA synthesis rate should be slowed down when (p)ppGpp binds the RNAP.
Transcription initiation is also impacted through the effects of (p)ppGpp on sigma factors by balancing the actions of the different σ factors according to the operons they regulate [Jishage et al., 2002]. For instance the RpoS regulon is induced by the stress sigma factor σ^S under high (p)ppGpp concentration to induce the expression of stress survival genes [Gaca et al., 2015].

Moreover, transcription repression is greatly enhanced by the presence of the small protein DksA. In particular, together with (p)ppGpp, it negatively regulates transcription of rRNA promoters and increases the amino acid promoter activity during stringent response [Paul et al., 2004a][Paul et al., 2005]. Recent works show that there exists a second binding site on the RNAP where (p)ppGpp can interact with DksA [Ross et al., 2016]. This second (p)ppGpp binding site (referred to as "site 2") is responsible for the majority of the effects of (p)ppGpp on transcription initiation, both *in vitro* and *in vivo*. Indeed, strains lacking site 2 are severely impaired for growth following nutritional shifts.
DksA might also act as a transcription elongation factor during the stringent response since it prevents transcription arrest upon ribosome stalling [Zhang et al., 2014]. Thus, it would modulate both transcription initiation and elongation. Another view is proposed by Roghanian, Zenkin, and Yuzenkova [2015]: they suggest that DksA coupled with (p)ppGpp increases the fidelity of transcription elongation by slowing down misincorporation events.

2.3.6 Effects of GTP and (p)ppGpp on the translation process

As seen in part 1.2.3, the translation process requires GTP to proceed. For each different translation step, GTP must be converted to GDP so that the ribosome proceeds to the next step. First, the translation will initiate thanks to the translation initiation factors (IF1, IF2 and IF3). Once translation has started, the translating process will require the action of different elongation factors (EF-Tu, EF-Ts, EF-G and EF-P). These factors appear to be strategic targets for (p)ppGpp.

2.3.6.1 Roles of GTP and (p)ppGpp in translation initiation

Upon optimal nutritional conditions, IF2 is expected to bind the 30S ribosomal subunit almost exclusively in the GTP form since the alarmone concentration is high while GDP concentration is low (IF2 has a similar affinity for both GTP and GDP) [Milon et al., 2006]. The IF2-GTP complex possesses a higher affinity for the 30S ribosomal subunit as compared to the IF2-GDP complex or the free IF2. Consequently, high concentration in GTP results in a stimulated translation initiation.

Furthermore, IF2 binds (p)ppGpp at the same nucleotide-binding site as GTP. The (p)ppGpp interferes with the formation of the IF2-dependent initiation complex and thus, severely inhibits initiation of dipeptide formation. *In vitro* data showed that the affinities of IF2 for ppGpp and GDP are similar and about 2.5-fold to 5-fold higher than the affinity of IF2 for GTP [Mitkevich et al., 2010]. ppGpp does not simply freeze IF2 in its *apo* form (*i.e.* when IF2 is not bound to GTP or GDP) but induces a third conformational change. The strong effects of (p)ppGpp on translation initiation is also due to the fact that the initiator $tRNA^{fMet}$ has even a much higher affinity for ppGpp-bound IF2 on the 30S subunit. Consequently, (p)ppGpp plays an important role in rapidly blocking protein synthesis.

2.3.6.2 Effects of (p)ppGpp on Translation Elongation Factors

(p)ppGpp also targets the elongation factor EF-G by binding to it [Hamel and Cashel, 1973] but its relative affinity is nearly ten times lower than the one it has for IF2, making EF-G a secondary target of (p)ppGpp [Mitkevich et al., 2010]. Similarly to IF2, ppGpp does not simply freeze EF-G in its *apo* form (*i.e.* when EF-G does not bind GTP or GDP) but induces a third conformational change. EF-Tu activity also appears to be inhibited by ppGpp [Miller et al., 1973], but this inhibition seems to be influenced by other factors such as the GTP level and more importantly by the interaction with another elongation factor EF-Ts [Rojas et al., 1984].

2.3.6.3 Effects of (p)ppGpp on Translation Termination Factors

The release factor RF3 activity also appears to be inhibited by (p)ppGpp and it is likely that during stringent response the (p)ppGpp-bound form is the major fraction of RF3 molecules [Kihira et al., 2012]. Thus, (p)ppGpp may decelerate the recycling of RF1 under stringent conditions while GDP accelerates the recycling of RF1 under nutrient-rich growth conditions.

2.3.6.4 Effects of (p)ppGpp on ribosome assembly

Different studies show that, under stress conditions, (p)ppGpp interacts with several proteins involved in ribosome assembly.

• BipA (BPI-inducible protein A) is a GTPase factor conserved among bacteria which has been associated with ribosomes under stress conditions and is implicated in the regulation of numerous cellular processes such as stress response [Robinson et al., 2008]. It is an essential factor for bacterial survival under nutrient limitation. A recent work using *Thermus thermophilus* suggests that BipA is a

translational factor which associates with either the 70S ribosome or the 30S subunit depending on the relative intracellular abundance of GTP and (p)ppGpp [Kumar et al., 2015]. Under rich nutritional conditions it associates with the ribosome in a GTP-bound state while it dissociates under stringent conditions [Robinson et al., 2008].

• In *E. coli*, the ObgE protein is a tRNA structural mimic involved in ribosome assembly. (p)ppGpp enhances ObgE binding to the 50S subunit which consequently prevents the formation of the 70S ribosome under nutrient limitation or other cellular stresses [Feng et al., 2014].

• A recent work has shown that in different bacteria, the activity of several GTPases is inhibited by (p)ppGpp [Corrigan et al., 2016]. *In vitro*, (p)ppGpp inhibits the *Staphylococcus aureus* activity of the GTPases RsgA, RbgA, Era, HflX, and ObgE. RbgA has also been characterized *in vivo* and it is indeed involved in ribosome assembly. The *B. subtilis* and *Enterococcus faecalis* enzymes ObgE, RsgA, RbgA, HflX and Era are also likely to be involved in ribosome biogenesis based on the homology regions their share with their *S. aureus* counterparts. *In vitro*, (p)ppGpp strongly inhibits *B. subtilis* RbgA, HflX and Era. Consequently, it is very likely that another role for (p)ppGpp during stringent response in bacteria is to prevent the assembly of the ribosome which contributes to slowing down the translation machinery.

2.3.7 Effects of (p)ppGpp on metabolic processes

In *B. subtilis*, (p)ppGpp inhibits the activity of YybT, a phosphodiesterase that hydrolyzes cyclic c-di-AMP and cyclic c-di-GMP to generate the linear dinucleotides 5'-pApA and 5'-pGpG [Rao et al., 2010]. The hydrolysis of c-di-AMP is fully suppressed during stringent response which should lead to an increase in c-di-AMP level. In *E. coli*, c-di-GMP and (p)ppGpp together control biofilm formation in response to translational stress [Boehm et al., 2009]. Thus, an increase in (p)ppGpp should also lead to an increase in c-di-GMP by inhibiting YybT activity so they can synergistically respond to certain stresses [Liu et al., 2015a].

In *E. coli*, (p)ppGpp inhibits the activity of enzymes involved in fatty acid synthesis or glycogen biosynthesis [Kanjee et al., 2012]. It also inhibits certain metabolic enzymes and the amino acid decarboxylases to store nutrients and amino acids under starvation conditions. Indeed, (p)ppGpp slows down amino acid consumption by, for instance, inhibiting the activity of the *E. coli* inducible lysine decarboxylase, LdcI, which catalyses the reaction that converts lysine into cadaverine and carbon dioxide [Kanjee et al., 2011].

Furthermore, inorganic polyphosphate (polyP) molecules accumulate during the stringent response so the cell can cope with stress (i.e. mutants lacking the polyP synthesizing enzyme fail to adapt to stress, and do not survive in stationary phase) [Kuroda et al., 1997]. This accumulation is due to the inhibition by (p)ppGpp of the polyP degradation (PPX) enzyme activity.

2.3.8 (p)ppGpp is a key regulator to optimally reallocate resources

Recent works used models to explain resource allocation by integrating the different effects of (p)ppGpp (described in section 2.3). Two approaches conclude that (p)ppGpp's regulatory actions maintain an optimal allocation of the resources during steady-state growth [Bosdriesz et al., 2015] as well as during dynamic nutritional shifts [Giordano et al., 2016]. The third approach conclude that during bacterial growth transition, the central metabolite pool (like amino acids and ketoacids) is the driver of global regulatory control, to which the (p)ppGpp signaling pathway is tightly linked [Erickson et al., 2017]. Thus, (p)ppGpp production through sensing of amino acid levels by RelA is considered to be a key metabolite to ensure optimal resource allocation.

(p)ppGpp also appears to play a role in fine-tuning GTP levels even under favourable growth conditions (*i.e.* growth medium supplemented with all amino acids) [Bittner et al., 2014]. Indeed, when extracellular guanosine is added to a medium where all amino acids are present (CAA medium), GTP levels increase unchecked in (p)ppGpp-deficient *B. subtilis* cells (termed (p)ppGpp^0 [Potrykus and Cashel, 2008]) but not in wild-type (WT) strains which are able to produce (p)ppGpp [Kriel et al., 2012]. Moreover, in mutant strains where the *guaB* gene is either mutated (leading to low GuaB activity) or under an isopropyl-β-D-1-thiogalactopyranoside (IPTG)-inducible promoter (P-*spac*), adding extracellular guanosine leads to increased GTP levels [Kriel et al., 2012][Kriel et al., 2014][Bittner et al., 2014]. In a strain able to synthesize (p)ppGpp, the GTP level reaches a maximum value (even upon high guanosine concentrations) leading to an optimal growth rate according to the growth medium; while for the (p)ppGpp^0 strain low guanosine concentrations increase the cell GTP level and growth rate but higher concentrations lead to cell death [Kriel et al., 2014][Bittner et al., 2014]. From these observations, Bittner *et al.* (2014) suggest that overloaded GTP levels can induce stress and inhibit growth instead of enhancing it (Figure 2.21).

This raises the question of why under non stringent conditions, cells can still be subject to stress due to GTP "excess". A possible explanation would be that in absence of the (p)ppGpp's regulatory actions, the translational error rate rises leading to non-functional protein production which is detrimental to cell growth and can provoke death. How translational errors occur in bacteria and how they can be linked to GTP/(p)ppGpp levels is discussed in the next chapter.

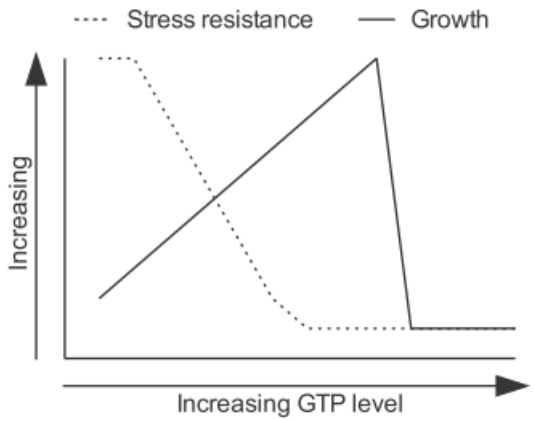


Figure 2.21: Model of the influence of GTP abundance on stress survival and growth in the absence of (p)ppGpp (from [Bittner et al., 2014]). The curve representing stress resistance is in dashed lines and the curve representing growth is in solid line. When the GTP level rises, the cell's growth also increases. However, Bittner *et al.* (2014) propose that without (p)ppGpp's actions, the uncontrolled GTP level increase also leads to a decrease in stress resistance until the GTP level reaches a threshold that results in cell death.

Overview: To efficiently reallocate its resources according to nutrient quality and availability, the bacterium can rely on two alarmones: GTP and (p)ppGpp. GTP is the driver of growth. In *B. subtilis*, it influences the transcription according to their transcription start site (TSS) of genes and by interacting with the transcription factor CodY. It is also required for translation initiation and elongation. Upon amino acid depletion, (p)ppGpp is produced in large amounts and triggers the stringent response. It directly targets enzymes involved in GTP production to decrease the intracellular GTP level. The drop in GTP abundance (in *B. subtilis*) or the synergetic interaction between RNAP, DksA and (p)ppGpp (in *E. coli*) decrease transcription initiation of genes involved in the translation machinery and other biosynthetic pathways; while it upregulates genes involved in amino acid synthesis and stress management. (p)ppGpp inhibits DNA replication elongation as well as the translation and ribosome assembly processes. It strongly inhibits IF2 and to a lesser extent elongation factors (EF-G, EF-Tu) and recycling factor RF3. (p)ppGpp also directly targets metabolic processes such as the fatty acid or glycogen biosynthesis. Consequently, the growth is slowed down during the stringent response. The main regulations operated by GTP and (p)ppGpp in bacteria are summarized in figure 2.22. (p)ppGpp production through sensing of amino acid levels by RelA is considered as a key metabolite to ensure optimal resource allocation. In the absence of (p)ppGpp, uncontrolled GTP levels appear to decrease bacterial stress resistance which can potentially lead to cell death.

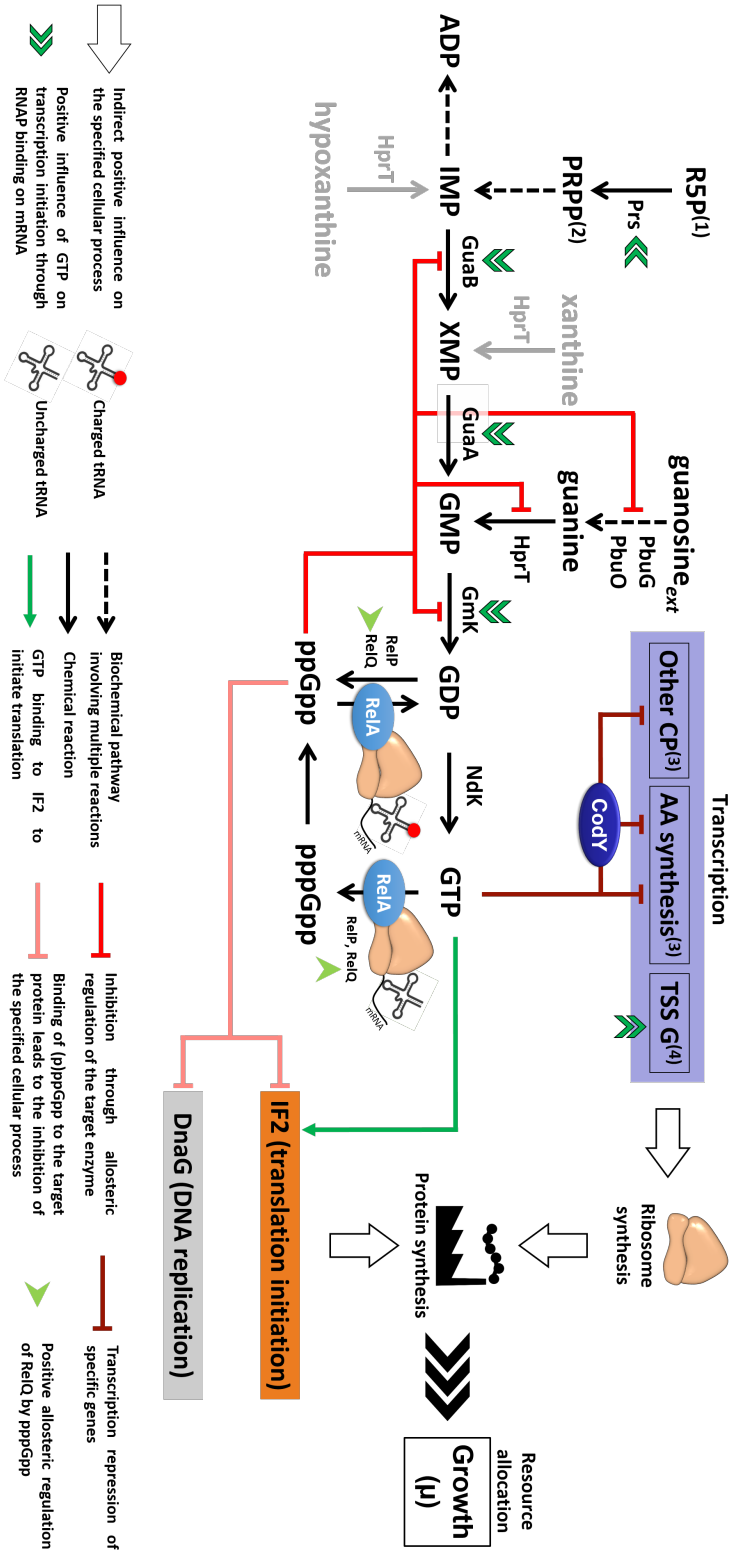


Figure 2.22: The main regulations operated by GTP and (p)ppGpp on bacterial cellular processes. GTP and (p)ppGpp levels vary according to resource availability. They influence the DNA replication, the translation machinery (including transcription, translation) and their own metabolic production. These regulations impact the ribosome synthesis, the protein synthesis and *in fine* the cell growth. Abbreviations: (1) Ribose-5-phosphate (2) Phosphoribosylpyrophosphate (3) CodY is a global regulator involved e.g. in the regulation of amino acid (AA) synthesis pathways and other cellular processes (CP) such as sporulation. (4) Transcription Start Site (TSS) which is a guanine.

Chapter 3

Translational errors

3.1 Introduction

The DNA replication process is a quite accurate process: the level of misincorporated nucleotides is about $\simeq 10^{-10}$ mutations/base pair/chromosome duplication in *E. coli* [Schaaper, 1993]. DNA transcription is less accurate: the rate of transcription errors in bacteria is estimated in the range of 10^{-4}-10^{-5} mutations/nucleotide in both Gram positive and negative [Traverse and Ochman, 2016]. The translation process, is where the highest level of errors can occur even if there exists proofreading steps during translation as described in section 1.2.4.

Two different types of errors can occur during translation: the missense and the frameshift errors [Parker, 1989]. Missense errors result from (i) an erroneously charged tRNA (*i.e.* misacylation), or (ii) from an anticodon-codon mismatch on the ribosome (*i.e.* misreading error). The frameshift errors usually correspond to a 2-base or 4-base translocation. The 2-base translocation is considered as a backward or 5' slip by the ribosome referred to as -1 frameshift (or leftward frameshift) and likewise, the 4-base translocation is referred to as +1 frameshift (or rightward frameshift). Even if the +1 and -1 frameshifts are the most observed frameshift events, other works have also observed slipping sizes of -2, -4, +2, +5 and +6 nucleotides [Weiss et al., 1987][Yan et al., 2015][Tsai et al., 2017].

Concerning misacylation, the level of errors involving closely related amino acids may occur in the frequency range of 10^{-6} [Kramer and Farabaugh, 2007] to 4 x 10^{-4} [Parker, 1989] per codon. Misreading errors are estimated to range from 10^{-3} to 10^{-4} per codon based on multiple studies with different reporter systems [Parker, 1989] [Kramer and Farabaugh, 2007]. In *B. subtilis*, higher levels of misincoporations have been measured and represent 0.4 $\pm$0.1% of the total of proteins produced during exponential growth in rich medium [Meyerovich et al., 2010]. During the stationary phase the frameshift errors are more heterogeneous among cells and increase to reach around 11 $\pm$2 % of the total of proteins produced [Meyerovich et al., 2010].

Frame-shift errors can spontaneously occur, an event estimated to range from 5x10^{-5} to 3x10^{-3} per codon [Parker, 1989]. These values have been obtained from the measurement of the β-galactosidase activity in *E. coli* strains where the *lacZ* gene has been randomly mutated by mutagenesis [Newton, 1970] [Atkins et al., 1972]. The obtained *E. coli* mutants possess *lacZ* mutations which correspond

to insertion or deletion of base pairs in the gene sequence [Newton, 1970]. Measurement of the β-galactosidase activity of these different mutants provides a spontaneous frameshift frequency range of 10^{-5} to 10^{-3} per codon [Atkins et al., 1972]. In *B. subtilis*, higher levels of frameshift errors have been measured which represent 2.4 $\pm$0.4 % of the total of proteins produced during exponential growth in rich medium [Meyerovich et al., 2010]. During the stationary phase the frameshift errors are more heterogeneous among cells and increase to reach around 11 $\pm$2 % of the total of proteins produced. Nevertheless, different studies have highlighted higher frameshift frequencies that correspond to programmed frameshift errors, a phenomenon used to regulate protein synthesis [Farabaugh, 1996].

3.2 Programmed translational frameshift errors and their utility

Different works have shown that the frameshift errors can be a way for living organisms (prokaryote and eukaryotes) to regulate the synthesis of certain proteins. They are referred to as Programmed Translational Frameshift (PRF). The focus will be on the regulations found in bacteria but PRFs are also found in eukaryotes such as the yeast and also in retroviruses [Parker, 1989].

3.2.1 The *prfB* gene in *E. coli*

A widely studied case of a programmed frameshift is the regulation of the peptidase release factor 2 (RF2) encoded by the *prfB* gene in *E. coli*. RF2 recognizes the stop codons UAA and UGA which leads to translation termination. The expression of *prfB* involves an autogenous regulatory loop [Farabaugh, 1996]. This feedback regulation takes place thanks to a nucleotide sequence in the *prfB* coding sequence (CDS) which is prone to frameshift events during translation [Craigen and Caskey, 1986][Curran and Yarus, 1988][Donly et al., 1990][Farabaugh, 1996]. Indeed, at the beginning of the RF2 CDS, there is an UGA terminator at codon 26. In the +1 shifted frame after codon 26, there are then 340 codons more before encountering a stop codon. To produce a complete and functional RF2 protein, a +1 frameshift must occur at the site where the first UGA stop codon is encountered which is CUU UGA C. The mechanism of frameshift regulation is as follow: when the UGA codon is in the A-site after the ribosome has translated CUU into leucine, if there is enough of RF2, RF2 binds to the ribosome which leads to premature termination of *prfB* translation. If insufficient RF2 is present, the ribosome will pause, allowing the peptidyl-tRNA$^{\mathrm{Leu}}$ to shift reading frames +1 from CUU to UUU, a near-cognate codon.

There are three elements on the sequence that are mandatory for an efficient frameshift at this site:

i. A stop codon (UGA) is required since replacing it by a sense codon importantly decreases the shift [Craigen and Caskey, 1986]. UGA can still be replaced by one of the two other stop codons UAA or UAG. Thus, the presence of a stop codon makes the ribosome pausing when there are insufficient levels of RF2.

ii. The last decoded zero-frame codon must be encoded by an aa-tRNA able to make a stable interactions with the newly formed codon after a +1 rightward shift of the sequence.

iii. The sequence immediately upstream of the frameshift site AGGGGG resembles the SD sequence of ribosome-binding sites and experiments showed that it was critical for producing an efficient frameshift. Its efficiency also relies on the spacing between the SD-like sequence and the frameshift site that must be of three base-pairs. Thus, this implies that formation of the SD-like interaction may strain the ribosome, pushing the peptidyl-tRNA into the +1 frame [Farabaugh, 1996].
The *prfB* sequence prone to frameshifts has been widely used to better understand the sequence features that stimulate frameshifts during mRNA translation.

3.2.2 The *dnaX* gene in *E. coli*

The γ subunit of the DNA polymerase III encoded by the *dnaX* gene is also regulated by ribosomal frameshifting in *E. coli* [Blinkowa and Walker, 1990] [Flower and McHenry, 1990] [Tsuchihashi and Kornberg, 1990]. The sequence prone to leftward frameshift is A AAA AAG. Similarly to what is observed for the frameshift stimulation of the *prfB* sequence, the ribosomal frameshift on *dnaX* sequence is stimulated by a SD-like sequence and a stem-loop downstream of the frameshift site [Larsen et al., 1994] [Larsen et al., 1997]. This slippery sequence has also been used to better understand the stimulatory elements of frameshifting.

3.2.3 The *cdd* gene in *B. subtilis*

In *B. subtilis*, the gene *cdd* encoding cytidine deaminase (CDA) has a CDS whose end is frameshift prone [Mejlhede et al., 1999]. The slippery site is CGA AAG and it is stimulated by an SD-like sequence located 14 bp upstream of the shift site, as observed for the *E. coli* genes *prfB* and *dnaX*. This results in a CDA subunit extended by 13 amino acids. This translational frameshift has a frequency of about 16% and leads to the formation of heterotetrameric forms of the enzyme CDA along with the dominant homotetrameric species. These two forms have approximately the same specific activity. Thus, it is not clear why a non-negligible portion of CDA is found to be heterotetrameric.

3.2.4 Occurrence of translational frameshift prone sequences in the genome

In *E. coli*, only two genes possessing frameshift prone sequences have been studied but it is possible that these slippery sequences are found in other parts of the genome. Indeed, the *dnaX* slippery sequence A AAA AAG is also present in 68 other genes in *E. coli*, and the frameshift levels of 12 of them have been measured, varying from 1.2 to 25.5 % which does not seem to affect the bacterium's fitness [Gurvich et al., 2003]. When looking closer to the role of these frameshifts there are no clear evidence of specific regulations. Nevertheless, it seems that evolution has shaped *E. coli*'s genome so it minimizes the occurence of this slippery sequence in its genome: its occurrence is lower (present 70 times in 68 genes) than for randomly built *E. coli* genomes (the mean occurrence in each genome generated is 97.6 times).
However, the frameshift-prone sequence CCC TGA has an occurrence higher in *E. coli*'s genome (found in 19 genes) as compared to an occurrence found for a randomly built *E. coli* genome (6.9

genes should have this sequence). Some genes seem to possess this sequence to regulate the synthesis of the corresponding protein but for other genes, there are no evidence for any translational regulations. Thus, synthesis of a small amount of dysfunctional proteins as a result of frameshift errors might be significantly harmful [Gurvich et al., 2003].
In *B. subtilis*, nearly 300 regions potentially contain frameshift prone sequences [Médigue et al., 1999]. They may correspond to genes regulated by a programmed frameshift which have not been studied yet or to nonfunctional genes such as pseudogenes (*i.e.* remaining fragments of functional genes which have undergone modifications such as mutations or duplications).
In conclusion, bacteria might have driven strong selection against frameshift-prone sequences at least in moderately expressed genes [Gurvich et al., 2003].

Overview: The different works on programmed frameshifting indicate that this can be a powerful tool for the bacteria to quickly regulate protein synthesis. However, frameshift sequences seem to be avoided across the genome of both *E. coli* and *B. subtilis*. An interpretation would be that undesired frameshift events during translation lead to nonfunctional protein production, which can be harmful for the cell's fitness and survival. Nevertheless, the slippery sequences of programmed frameshifting are a good tool for understanding the overall mechanisms of frameshift events during translation.

3.3 Translational frameshift stimulators

Frameshift events are in most cases to be avoided but several studies have shown that there are different elements that can stimulate it.

3.3.1 Ribosome pausing

Concerning rightward frameshift, several reviews state that there must be a pause in the A-site due to lack of aa-tRNA encoding the corresponding codon [Farabaugh and Björk, 1999] [Urbonavičius et al., 2001] [Harger et al., 2002]. This pause allows the peptidyl-tRNA to slip +1 if it can form interactions with the newly formed mRNA codon in the A-site. First, as mentioned in part 3.2, the A-site pausing is stimulated by the SD-like sequence. Moreover, mutations or genetic modifications which affect the availability of the mature aa-tRNA corresponding to the codon present in the A-site lead to a significant increase in frameshifting [Jäger et al., 2013].
Similarly, the stalling of the ribosome at a "hungry" codon calling for an aminoacyl-tRNA in short supply is also a stimulator for leftward frameshifting, which means that the A-site must be empty when the -1 shift occurs [Barak et al., 1996]. Even if two "hungry" codons code for the same amino acid, one leads to more shiftiness than the other but Barak *et al.* (1996) could not come up with an explanation.

3.3.2 Stability of the mRNA:tRNA base pairs before aa-tRNA slipping in the P-site

3.3.2.1 Near-cognate decoding of the P-site codon

In the yeast *Saccharomyces cerevisiae*, the presence of a near-cognate aa-tRNA in the P-site enhances frameshifting (Figure 3.1 A) [Sundararajan et al., 1999]. This favors the hypothesis that an abnormal codon-anticodon peptidyl-tRNA interaction deforms the structure of the aa-tRNA-mRNA complex in the P-site which makes the peptidyl-tRNA more likely to slip on the mRNA and to rephase the mRNA reading one base pair downstream.
Farabaugh and Björk (1999) suggest that leftward frameshifting is also stimulated when a near-cognate aa-tRNA is present in the P-site, similarly to rightward frameshifting.

3.3.2.2 Role of the tRNA nucleosides' modifications

As seen in section 1.2.4.1, the post-translational modifications of tRNA nucleosides seem to play a role in maintaining translation accuracy and in particular in preventing rightward frameshifting. Indeed, hypomodified nucleoside are very slow to enter the A-site which induces a pause [Urbonavičius et al., 2001] and consequently likely triggers frameshifting as explained in section 3.3.1 (Figure 3.1 B). Moreover, a modified anti-codon nucleoside stabilizes the interaction between the mRNA and the tRNA anticodon when in the P-site [Urbonavičius et al., 2001] (Figure 3.1 C). Surprisingly, nucleoside tRNA hypomodifications (*i.e.* the tRNA nucleoside does not undergo post-translational modifications) did not stimulate leftward frameshifting [Urbonavicius et al.]. To explain it, Urbonavicius *et al.* (2003) propose that the interactions between the E- and P-site tRNAs and various parts of the ribosome are more likely to influence -1 frameshift.
Moreover, the frameshift rate is higher when the third nucleotide of the anticodon of the P-site aa-tRNA forms a Wobble interaction with the last base of the peptidyl-tRNA anticodon, which is weaker than a Watson-Crick interaction [Curran, 1993]. Thus, this suggests that a destabilized interaction at the Wobble position stimulates frameshifting.

3.3.3 Stability of the ribosomal grip of the peptidyl-tRNA

Frameshifting is stimulated by a weakened interaction between the ribosome and the stem as well as the D-arm of the peptidyl-tRNA (DSL) (figure 1.10) present in the P-site [Näsvall et al., 2009]. Indeed, mutations altering the interaction between the peptidyl-tRNA and the P-site codon as well as the ribosomal grip on the peptidyl-tRNA enhances frameshifting. Frameshifting is stimulated by a truncated C-terminal of the S9 ribosomal protein. This part of S9 is known to penetrate the ribosome like a tentacle and the two last amino acids make a contact with the 5' phosphate of nucleotide 32 (R130) and the 5'-phosphates of positions 33 and 34 (K129) of peptidyl-tRNA (Figure 3.2). Thus, altering the ribosomal grip seems to enhance frameshifting. These conclusions have been confirmed by structural analysis of +1 frameshifting [Maehigashi et al., 2014]. During translocation to the P-

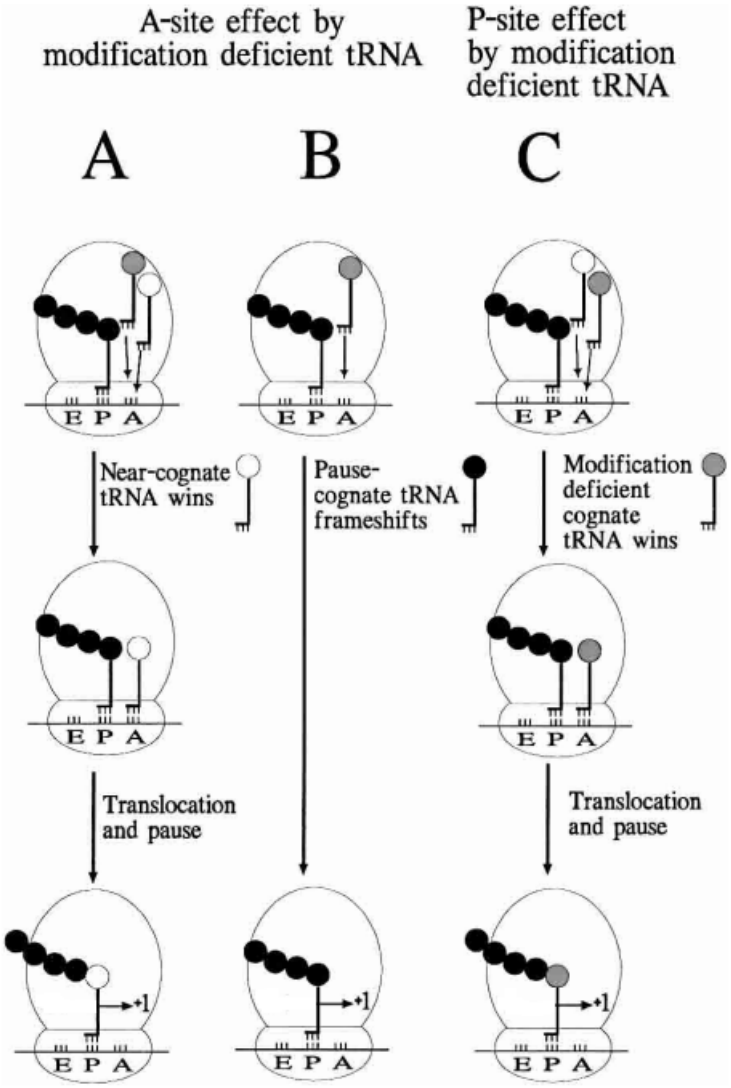


Figure 3.1: Model for frameshifting in presence of an hypomodified cognate aa-tRNA. (adapted from [Urbonavičius et al., 2001]). (A) An hypomodified cognate aa-tRNA is defective in the selection step of aa-tRNA, leading to the selection of a near-cognate tRNA instead at the A-site. After a normal three-nucleotide translocation, the unstable interaction between the near-cognate aa-tRNA and the mRNA provoke a +1 slipping of the aa-tRNA in the P-site. (B) The hypomodified cognate aa-tRNA is very slow to enter the A-site which induces a pause and consequently triggers frameshifting. (C) The hypomodified cognate aa-tRNA is accepted in the A-site and after a normal three-nucleotide translocation, the hypomodification induces slippage into the +1 frame. For clarity, only one tRNA is depicted as residing on the ribosome, but two tRNAs are always present in the A- and P-sites or P- and E-sites.

site, there are several ribosomal residues which interact with the stem of the ASL whereas very few interactions are made between the ribosome and the codon-anticodon helix. Thus, perturbing these interactions potentially leads to frameshifting.

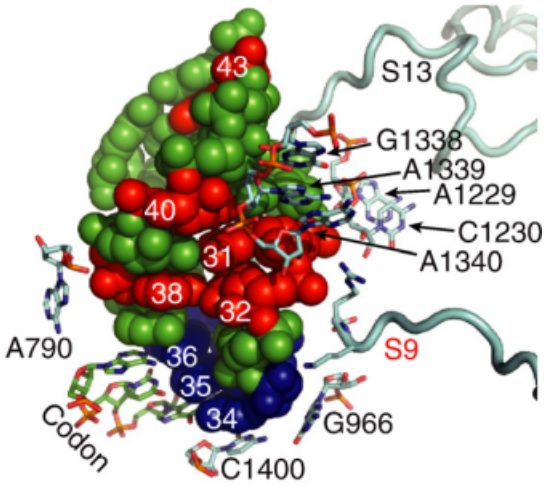


Figure 3.2: Interactions between peptidyl-tRNA and the ribosome in the 70S P-site. (adapted from [Näsvall et al., 2009]). The ASL (positions 26-44) of $tRNA^{Met}$ is shown as spacefill. The anticodon is blue and the rest of the nucleotides are in green and red (mutated positions in the sudy of [Näsvall et al., 2009]). Residues of rRNA and ribosomal proteins that have atoms within 3.8 Å of the tRNA are shown as stick representations, and the rest of the proteins are shown as tubes.

3.3.4 Stability of the mRNA:tRNA base pairs after aa-tRNA slipping in the P-site

Frameshifting efficiency also tends to depend on stable mRNA:tRNA base pairing after rephasing [Curran, 1993]. The more the P-site peptidyl-tRNA is able to form interactions (from 3 to 0) with the newly formed P-site codon, the more likely the frameshift event will occur (Figure 3.3).

3.3.5 Abundance of the tRNA decoding the next codon in the +1 frame

When overexpressing the gene encoding the aa-tRNA decoding the next codon in the +1 frame (the new codon formed just after the +1 shift occurred), the frameshifting level is increased [Pande et al., 1995]. Pande *et al.* (1995) state that a frameshift-competent aa-tRNA must occupy the P-site of the ribosome, and the aa-tRNA which decodes the first +1 frame codon must transiently enter the ribosomal A-site. The more abundant this aa-tRNA is, the more likely the translation read will shift.

3.3.6 Influence of the E-site

Several studies suggest that the E-site plays a crucial role in the efficiency of +1 PRF in *E. coli*, showing that RF2 programmed frameshifting is inversely correlated with E-site stability in *E. coli* (reviewed in [Liao et al., 2008]). Indeed, premature release of the E-site tRNA from the ribosome correlates with high levels of frameshifting products.

Shift site	Message:anticodon pairs for the +1 realignment	β-gal	F(%)
UAU CUU UAG	U:G, U:A, U:G	8047	48
CCC	C:G, C:G, U:G	1874	11
UUU	U:A, U:A, U:G	1811	11
GUU	U:C, U:A, U:V*	1746	10
CCU	C:G, U:G, U:V*	833	5
UGU	G:A, U:C, U:G	260	1.5
CAU	A:G, U:U, U:Q	250	1.5
AUU	U:U, U:A, U:G	240	1.4
UUC	U:A, C:A, U:G	103	0.6
GCC	C:C, C:G, U:V	84	0.5
UCU	C:A, U:G, U:V	80	0.5
ACC	C:U, C:G, U:G	69	0.4
GAU	A:C, U:U, U:Q	57	0.3
CAG	A:G, G:U, U:C	53	0.3
CAA	A:G, A:U, U:S	50	0.3
GAA	A:C, A:U, U:S	41	0.2
GAG	A:C, G:U, U:S	40	0.2
GUC	U:C, C:A, U:G	39	0.2
GAC	A:C, C:U, U:Q	38	0.2
AGC	G:U, C:C, U:G	38	0.2
UUA	U:A, A:A, U:D	37	0.2
GCU	C:C, U:G, U:V	36	0.2
UAC	A:A, C:U, U:Q	35	0.2
GGA	G:C, A:C, U:E	30	0.2
GUG	U:C, G:A, U:V	26	0.15
ACG	C:U, G:G, U:V	23	0.14
CCA	C:G, A:G, U:V	20	0.12
ACA	C:U, A:G, U:V	14	0.08
CUG	U:G, G:A, U:C	12	0.06
UCA	C:A, A:G, U:V	10	0.06
AUA	U:U, A:A, U:F	9	0.05
GUA	U:C, A:A, U:V	7	0.04

Figure 3.3: Effects of tRNA:message stability on frameshift frequency at the *E. coli* RF2 programmed frameshift site (adapted from [Curran, 1993]). A, G, C and U are the standard bases; Q = queuosine; S = 5-methylaminomethyl-2-thiouridine; and V = uridine-5-oxyacetic acid; D is an unidentified derivative of A that pairs with G and A; E is an unidentified derivative of U that pairs with A; and F is a modified pyrimidine that pairs with A. β-gal are β-galactosidase units. F is the frameshift frequency, relative to a pseudowildtype *lacZ* allele that produces 16800 β-galactosidase units. The more interactions the anticodon of the P-site peptidy-tRNA will make with the newly formed mRNA codon, the more likely the shift will occur.

Overview: The elements that favor translational frameshifts are the ribosomal pausing; the stability of the interaction between the P-site peptidyl-tRNA anticodon and the mRNA codon before the shift; the presence of a near-cognate aa-tRNA in the P-site; a weakened ribosomal grip on the P-site peptidyl-tRNA; the stability of the interaction between the P-site peptidyl-tRNA and the newly formed codon after the shift; a relatively high abundance of the charged peptidyl-tRNA corresponding to the newly formed codon present in the A-site; and a relatively rapid departure of the tRNA present in the E-site.

3.4 Specificities of the rightward and leftward frameshifts

As seen previously, frameshift can be forward or backward. But what elements provoke preferentially forward instead of backward frameshift and the other way around? Which mechanisms specifically

trigger either rightward or leftward translational frameshift?

3.4.1 Sequence specificity

To better understand the elements that favour leftward or rightward frameshifting, a sequence prone to frameshift (C UUC AAG) has been used and some of its base pairs have been changed [Lindsley and Gallant, 1993]. In this case, under lysyl-tRNA limitation, the ribosome pauses during translation when the hungry codon AAG corresponding to lysine is in the A-site. This stimulates frameshifts (as explained in section 3.3.1) which can be leftward or rightward. Lindsley and Gallant (1993) conclude from their experiments that the critical heptanucleotide for rightward frameshifting includes three bases to the left of the hungry codon and one to its right; and for the leftward frameshifting the critical heptanucleotide includes four bases to the left of the hungry codon and none to the right. They could not state exactly on the nucleotide bases required. Another study also demonstrates that the leftward frameshifting is strongly influenced by the identity of the bases two, three and four positions to the left of the frameshift site [Kolor et al., 1993].

3.4.2 Kinetic mechanisms of the rightward and leftward frameshifts

3.4.2.1 Possible mechanisms for rightward frameshift events

A kinetic model suggests that a combination of stimulatory signals leading to the release of deacylated tRNA in the E-site, aa-tRNA slippage in the P-site, and the hungry codon effect in the A-site synergistically promotes efficient +1 ribosomal frameshifting [Liao et al., 2008]. Moreover, this model shows that the rate of P-site aa-tRNA slippage is the dominant factor, while the effect of hungry codon in the A-site and E-site tRNA destabilization further enhances +1 PRF. The analysis also suggests that rightward frameshift more likely occurs after the E-site tRNA dissociates while the codon recognition step is occurring (A-site occupied). The +1 slippery event occurs after the A-site codon has been rejected, leading to only the P-site occupied and thus slippage is more likely to occur since the interaction between the mRNA and the ribosome is weakened.
This kinetic model has been used to build a mechanistic-based genetic algorithm search for potential +1 frameshift sites in *E. coli* called FSscan [Liao et al., 2009]. The program assigns scores for a 16-nt window along a gene sequence according to different effects of the stimulatory signals (*i.e.* an anti-SD sequence six bp upstream of the E-site position) and interactions of the E-, P- and A-site in the ribosome. Additional scores are also attributed according to the availability of the cognate tRNA encoding A-site codon. This algorithm can detect potential frameshift sites which have been experimentally confirmed for most of them [Liao et al., 2009], thus validating the kinetic approach [Liao et al., 2008].

Two shift-prone mechanisms have also been proposed by using the slippery sequence CCC-C found in the sequence AUG-CCC-CGU-U, and where CCC codes for proline, the native codon CGU codes for arginine while the codon GUU formed after shifting codes for valine [Gamper et al., 2015]. The first

one is a slow mechanism where the frameshift occurs while the $tRNA^{Pro}$ is stalling at the P-site next to an empty A-site (Figure 3.4 a and c). The second mechanism is a fast one where the frameshift takes place during the translocation of the $tRNA^{Pro}$ into the P-site (Figure 3.4 b and d). The slow mechanism is likely to occur under starvation conditions. Indeed, when there is a lack of amino acids, the A-site remains more often empty.

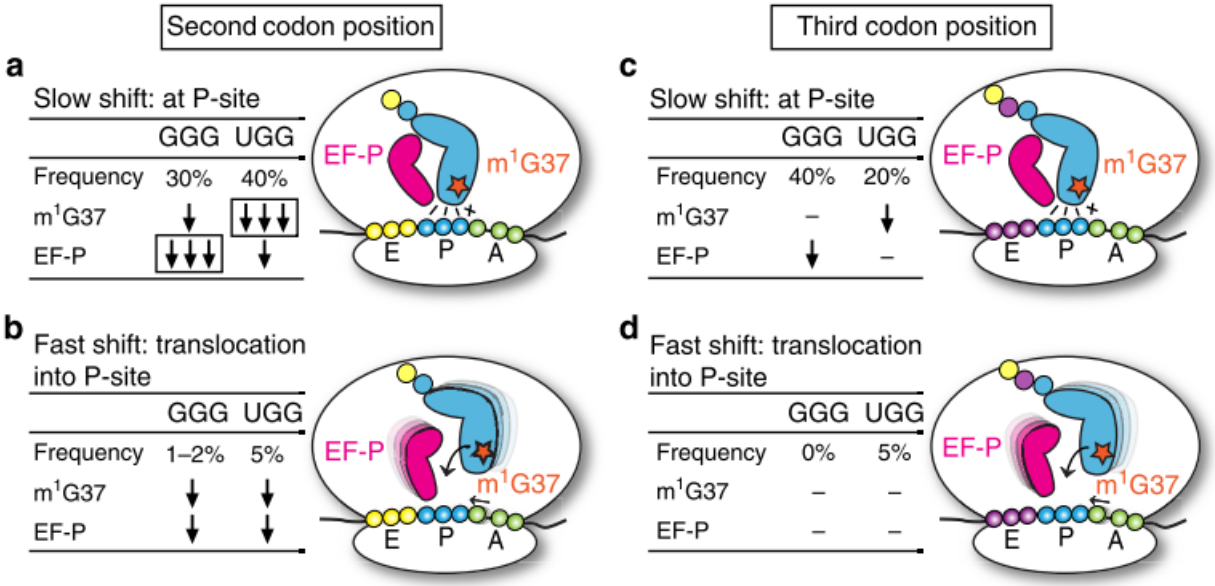


Figure 3.4: A model of +1FS on CCC-C by GGG and UGG $tRNA^{Pro}$ (from [Gamper et al., 2015]). (a) and (c) In the slow mechanism, the shift occurs when $tRNA^{Pro}$ stalls at the P-site next to an empty A-site. (b) and (d) In the fast mechanism, the shift occurs when $tRNA^{Pro}$ is translocated onto the P-site. (a) The high frequencies of slow shifts at the second codon are suppressed primarily by m^1G37 for the UGG tRNA and by EF-P for the GGG tRNA. (b) The low frequencies of fast shifts at the second codon are suppressed by both m^1G37 and EF-P for UGG and GGG tRNAs. (c) The high frequencies of slow shifts at the third codon are suppressed by m^1G37 for the UGG tRNA and by EF-P for the GGG tRNA. (d) The low frequencies of fast shifts in the early elongation phase are not effectively suppressed by m^1G37 or EF-P. Open arrows indicate suppression of error frequencies, whereas boxed arrows indicate suppression of both frequencies and kinetics of error formation. One arrow indicates a reduction of 2- to 3-fold, two arrows indicate 3- to 30-fold, three arrows indicate greater than 30-fold and a '–' indicates less than 2-fold effects. Percent frequencies are rounded up to the closest approximation.

Importantly, Gamper *et al.* (2015) show that the cell has developped strategies to prevent a rightward frameshifting event during translation initiation. Indeed, if the slippery sequence is found at the second codon (just after AUG), the frameshift event can be prevented by stabilization of the peptidyl-tRNA present in the P-site through the recruitment of the translation factor EF-P or through post-translational modification of the aa-tRNA (*i.e.* the tRNA's guanine base at position 37 G37 has its nitrogen atom N^1 which is methylated and is referred to as m^1G37) (Figure 3.4). EF-P, which is

known to relieve ribosomes stalling at poly-Proline sequences [Doerfel et al., 2012][Ude et al., 2013], inhibits slow shifts of the proline GGG tRNA. The nucleoside modification m^1G37 also inhibits slow shifts but for the UGG tRNA. When the corresponding proline codon is at position 2, both m^1G37 and EF-P have an important effect on the kinetics of the shift concerning UGG tRNA and GGG tRNA respectively. However, on the other codon positions such as position 3, they mainly act on the frequency of the shift. Concerning the fast shift mechanism, the effects of m^1G37 and EF-P are not specific to the proline codon present in the A-site. Gamper *et al.* (2015) suggest that m^1G37 can reduce the shift frequency by pre-organizing the anticodon loop, whereas EF-P can help to position tRNA correctly upon entering the P-site. Thus, avoiding frameshifting right after the translation initiation to ensure a correct reading frame might be an essential aspect of translation [Gamper et al., 2015].

3.4.2.2 Possible mechanisms for leftward frameshift events

An 'integrated model' for programmed frameshifting in the leftward direction has been proposed [Harger et al., 2002] which states that -1 slippery site tend to have X XXY YYZ as nucleotide sequence. The P-site is occupied by the aa-tRNA encoding XXY and the A-site is occupied by the aa-tRNA encoding YYZ, and the ribosome pauses due to the presence just after the A-site codon of an mRNA pseudoknot which consists of two nested stems, the loop of one stem forming the base-pairs of the second [Harger et al., 2002]. On this slippery site, thanks to pausing, the non-wobble bases of both the A- and P-site aa-tRNAs can re-pair with the new -1 frame codons.

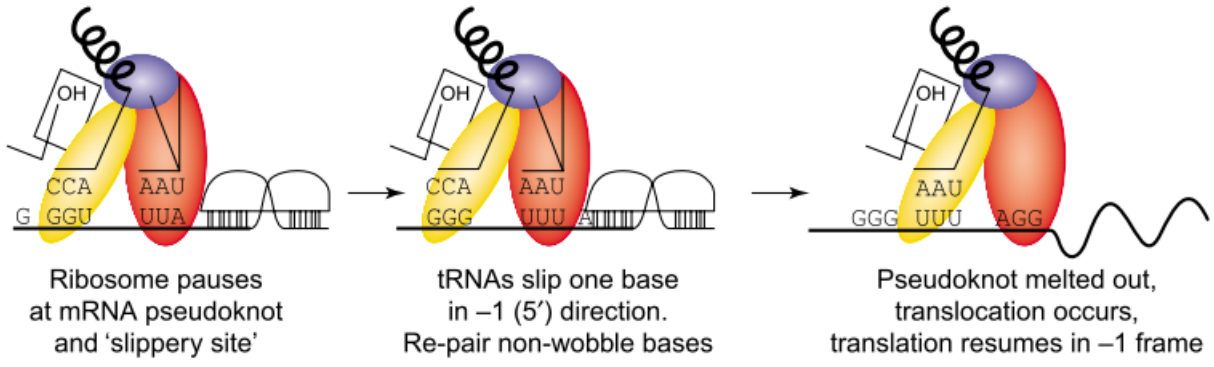


Figure 3.5: An integrated model of the leftward programmed framshift. (adapted from [Harger et al., 2002]). The slippery site X XXY YYZ here is G GGU UUA found in the L-A dsRNA virus of yeast [Giedroc et al., 2000]. The A-site is in red, the P-site is in yellow, the petidyl-transferase center is in purple and the nascent peptide chain is in black. The pseudoknot is represented on the right of the A-site.

A fully reconstituted in vitro translation system has been developed to study real-time kinetics of -1PRF and to identify the steps that control frameshifting [Caliskan et al., 2014]. A kinetic analysis of the data obtained lead to the construction of a kinetic model of -1PRF (Figure 3.6). This model

states that the frameshift event occurs when two aa-tRNAs are respectively present in the P- and A-site. First, after the cognate aa-tRNA is present in the A-site and EF-Tu has been released, EF-G is recruited to shift the reading frame by three nucleotides to the right in order to add the next amino acid. Thus, the aa-tRNAs move rapidly to a chimeric state. The aa-tRNA once in the P-site is now in the E-site and then moves apart from the 30S subunit. The pseudoknot impacts the speed of this step: it is quicker for ribosomes which switch to the -1 frame as compared to the ones which remain in the 0 frame. Then the EF-G is released once the tRNA has departed from the E-site and translation continues to proceed. If the slippery sequence is absent and there is a pseudoknot, the progression of the ribosome is stalled, leading to an extremely slow EF-G release. On the contrary, if the slippery sequence is present, the reframing it provokes leads to a quicker EF-G release and thus elongation can rapidly continue to proceed. Caliskan *et al.* conclude that with both stimulatory elements present, the slippery sequence provides the necessary freedom for the ribosome to change its position with respect to the pseudoknot, allowing for the completion of translation and continuation of translation in the new frame. Thus, they suggest that -1 slippage allows for a 3-fold faster movement of the ribosome through the pseudoknot base.

3.4.3 Structural insights

To better understand how rightward frameshifts occur, the frameshift suppressor $tRNA^{SufA6}$ and the $tRNA^{Pro}$ have been used to determine X-ray crystal structures of the ASL^{SufA6} and ASL^{Pro} when they are bound to the *Thermus thermophilus* 70S A-site [Maehigashi et al., 2014]. Frameshift suppressor $tRNA^{SufA6}$ is a derivative of $tRNA^{Pro}_{CGG}$ and contains an inserted guanosine (referred to as G37.5) between positions 37 and 38 in the anticodon loop (Figure 3.7 A and B). The ASL^{SufA6} manages to have formation of a single hydrogen bond between the two mismatched cysteines C•C at the third position (Wobble position) when it is bound to the mRNA codon CCC. The geometry of the mismatch at the Wobble position is similar to the modified wobble base interactions obtained with ASL cmo^5U34 that is a nucleotide shaped by evolution. This suggests that the wobble position achieves plasticity due to conformational changes imposed by the unusual base pairing.

The interaction between the bases 32 and 38 is disrupted in the ASL of $tRNA^{SufA6}$ as compared to $tRNA^{Pro}_{CGG}$ which seems to promote +1 decoding (Figure 3.7 C and D). Given that EF-G interacts with the 5' stem ASL proximal to U32, Maehigashi *et al.* (2014) hypothesize that during translocation from the A-site to the P-site this interaction causes a rearrangement of the anti-codon loop in the P-site, where ASL also interacts with the 16S rRNA and the S9 residues (Figure 3.7 E). This P-site remodeling is likely to propagate to the codon-anticodon interaction, which leads to a readjustment in the +1 frame. In summary, the 32-38 interaction disruption impacts translocation where rearrangements must occur to stabilize the $tRNA^{SufA6}$ in the P-site, thus leading to +1 frameshifting. From a structural point of view, it seems that changes in reading frames are likely to occur when a reorganization takes

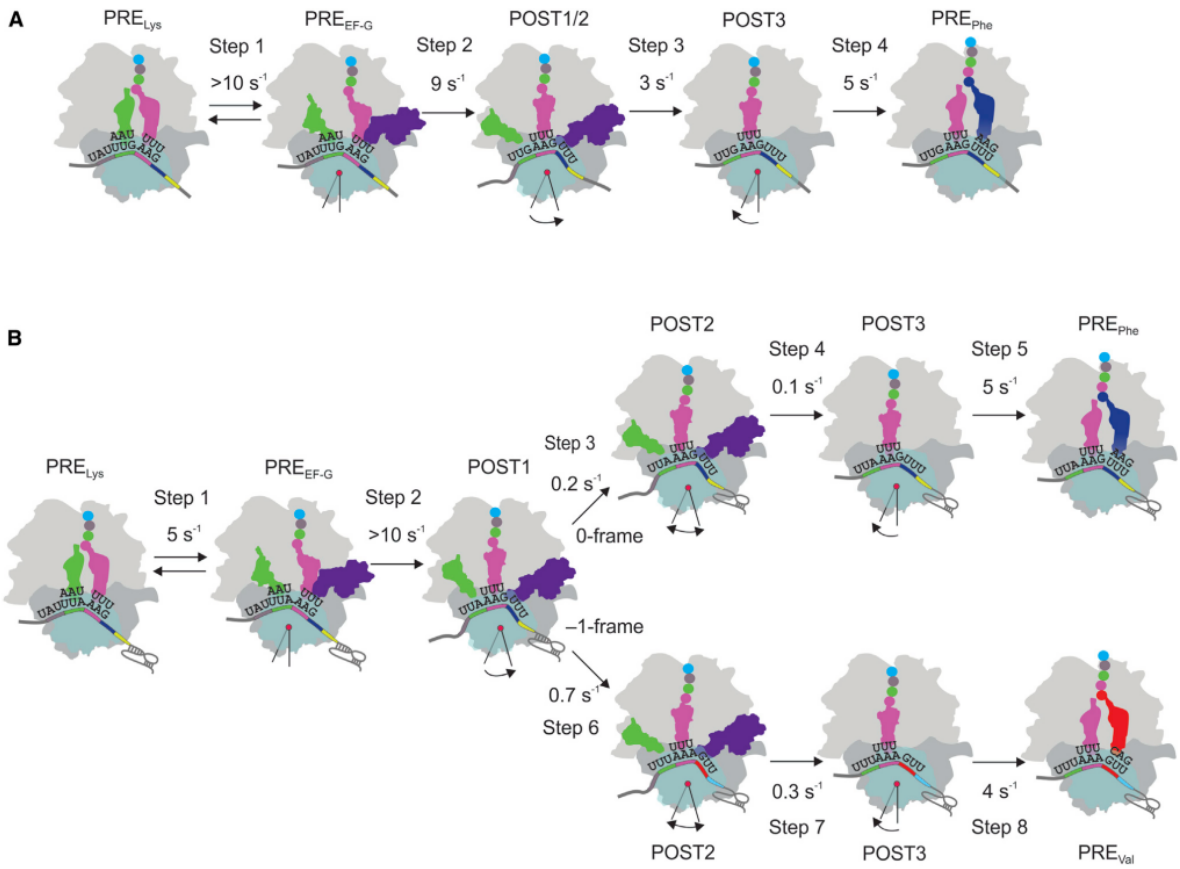


Figure 3.6: Kinetic model of the leftward programmed frameshift (from [Caliskan et al., 2014]) (A) In the absence of -1PRF stimulatory elements, EF-G binds rapidly to the PRE ribosomal complex (step 1). Subsequently, the tRNAs move into a chimeric state (POST1/2) in which both deacylated tRNA and aa-tRNA move relative to the 50S subunit to translocate respectively in the E- and P-sites, whereas their contacts with the 30S subunit are not disrupted (step 2). In step 3, $tRNA^{Leu}$ detaches from the 30S head, probably during the backward 30S head rotation, and EF-G is released. In step 4, EF-Tu-GTP-Phe-$tRNA^{Phe}$ binds to the A-site, and Phe is incorporated into the peptide chain. (B) Kinetic mechanism of the -1PRF. The slippage occurs during translocation of the two tRNAs bound to the slippery sequence ($tRNA^{Leu}$ and MYLK-$tRNA^{Lys}$). Recruitment of EF-G (step 1) to the ribosomal PRE complex facilitates rapid tRNA movement (step 2) into a chimeric state (POST1). However, the following steps are inhibited by the presence of the pseudoknot. Further movement of $tRNA^{Leu}$ proceeds in two steps. First, $tRNA^{Leu}$ moves on the 50S subunit into a POST2 state while the distance to the 30S subunit is not changed (steps 3 and 6 in 0 frame and -1 frame, respectively). Second, $tRNA^{Leu}$ and the 30S subunit move apart (steps 4 and 7) into a POST3 state. Steps 3 and 4 are particularly slow for the tRNA that remains in 0 frame, which limits the rate of the following Phe-$tRNA^{Phe}$ binding (step 5). In contrast, $tRNA^{Leu}$ movement on those ribosomes which switched to the -1 frame, is faster (step 6), followed by dissociation of $tRNA^{Leu}$ from the 30S subunit, 30S head rotation, and dissociation of EF-G (step 7) and binding of Val-$tRNA^{Val}$ (step 8).

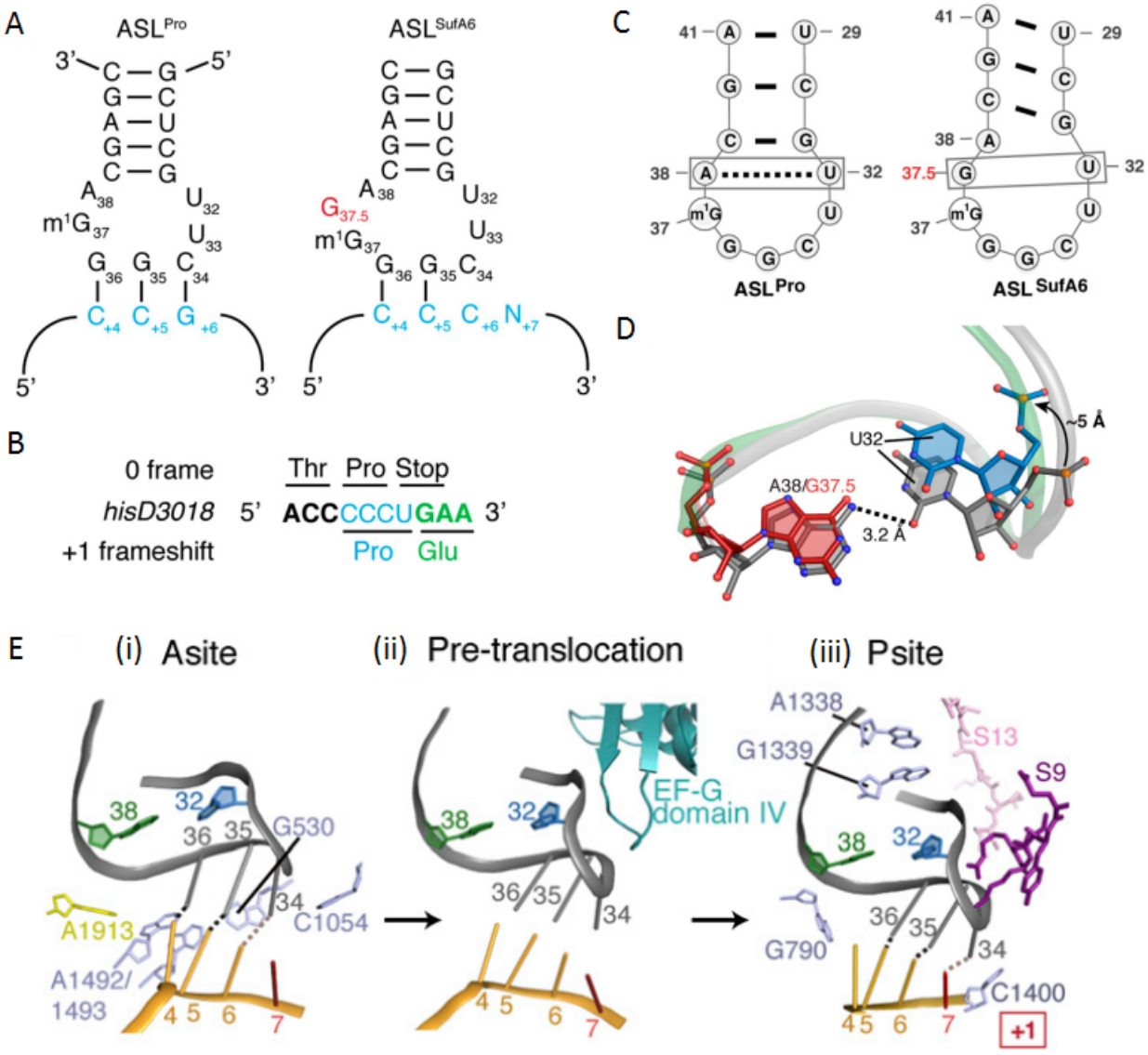


Figure 3.7: Structural study of the rightward frameshift using the frameshift suppressor ASL^{SufA6} which decodes 4-nt codon to produce a +1 frameshift (adapted from [Maehigashi et al., 2014]). (A) base-pairing of the ASL of tRNA^{Pro} and tRNA^{SufA6} with their respective mRNA codons (blue). The insertion of the anticodon loop of the +1 frameshift suppressor tRNA^{SufA6} is between nucleotides 37 and 38; G37.5 is red, and the m^{1}G37 modification is shown in both. (B) The frameshift suppressor tRNA^{SufA6} decodes the proline codon as CCC-U (blue) rather than the cognate 3-nucleotides proline codon CCC (0 frame). (C) Secondary representations of both ASL^{Pro} and ASL^{SufA6} show that expansion of the ASL^{SufA6} anticodon loop to 8 nucleotides abrogates the conserved hydrogen bond between U32 and A38 (boxed) and alters the incline of base pairs in ASL^{SufA6}. (D) The U32-A38 interaction in ASL^{Pro} (gray) is shown alongside U32 (blue) and G37.5 (red) of ASL^{SufA6}, emphasizing the lack of interaction. (E) (i) In the A-site, the ASL (gray) interacts with a 3-nucleotide codon (yellow) with the additional nucleotide (7) that is read as the preceding codon shown in red. (ii) Domain IV of EF-G (teal) interacts with the ASL 5' stem and nucleotide 32, potentially inspecting the integrity of the anticodon loop. (iii) On movement into the P-site by EF-G, the 16S rRNA residues A1338, G1339 (light purple), S9 (purple), and S13 (pink) inspect the stem of the ASL, whereas very few interactions are made with the codon-anticodon helix, which is now in a +1 frame.

place in the P-site in order to achieve optimal stability.
Such structural studies have not been done for leftward frameshifts.

Overview: Even if the forward and backward frameshifts share common features, the nucleotides of the slippery sequence will determine which way the ribosome will slip. The critical bases for the rightward frameshift are the three bases to the left of the hungry codon and one to its right. The critical bases for the leftward frameshift are the four bases to the left of the hungry codon and none to its right.
It appears that both rightward and leftward frameshifts occur during translation through two different types of mechanisms: a fast one where the frame is shifted during translocation so when both the P- and A-site are occupied but the E-site is empty; and a slow one which happens when only the P-site is occupied, which requires a lack of aa-tRNA corresponding to the codon present in the A-site. The fast mechanism seems to mainly occur in the case of a programmed frameshift, which is partly triggered by the presence of either a SD-like sequence few base pairs upstream of the slippery site or an mRNA pseudoknot/stem-loop few base pairs downstream of the slippery site. This leads to ribosome pausing which, in this case, can also be amplified by the presence of a "hungry" codon present in the A-site whose corresponding aa-tRNA is poorly available. In the slow mechanism, the ribosome pausing is only due to the latter case.
For both mechanisms, it seems that frameshift is also due to the rearrangement of the conformation of the aa-tRNA present in the P-site so that its ASL optimizes its interaction with the ribosome to reach optimal stability. The more unstable the P-site aa-tRNA is, the more likely it will shift the frame to adopt a more stable position. This instability can be from different origins as explained in section 3.3.

3.5 Ribosomal frameshift under starvation conditions

3.5.1 Frameshift errors are increased under aa-tRNA limitation

A recent study by Caliskan *et al.* (2017) on the effect of aa-tRNA limitation on the *dnaX* -1PRF shows that the P-site peptidyl-$tRNA^{Lys}$ can slip into the -1 frame when the A-site is vacant. This aa-tRNA depletion-stimulated frameshifting (ADF) pathway requires only a slippery tetranucleotide sequence (A AAA or A AAG), does not require the Stem Loop stimulator, and is slow compared to the predominant -1 frameshifting. The ADF route depends on a 'hungry' codon in the A-site. In the presence of the cognate aa-tRNA, the pathway does not operate and translation proceeds in the 0 frame without a translational pause. Moreover, when the -1-frame aa-tRNA is lacking, the ribosomes switches into the -2 frame. From their results and previous one they suggest that ribosomes can employ different frameshifting pathways on the same slippery sequence, switching from PRF to ADF and extending the repertoire of accessible reading frames under certain cellular conditions such as starvation or bacterial infection [Lainé et al., 2008][Olubajo and Taylor, 2005]. Hence, changing the reading frame through

the availability of aa-tRNAs may provide an efficient way to modulate the cellular proteome to adjust to the cellular environment and to achieve alternative gene expression [Caliskan et al., 2017]. These observations are in agreement with an overall increase in the occurence of +1 frameshifts when *B. subtilis* is in stationnary phase (*i.e.* nutrients are depleted in that growth phase) [Meyerovich et al., 2010].

Moreover, when an mRNA with a leftward frameshift prone sequence is translated, protein products can also include products obtained through -4 and +2-nucleotide frameshifts [Yan et al., 2015]. Once the ribosome is on the slippery site, it may frameshift, but it is sensitive to mismatches that result from the pairing between the frameshifted codons and anticodons. These mismatches likely trigger a fidelity check mechanism that results in the ribosome either to proceed translation in a new frame or to prematurely abort translation. However, in the experiments conducted by Caliskan *et al.* (2017) using also the frameshift prone *dnaX* sequence, the amounts of putative -4 or +2 frameshifting products are very small and close to background, <5%. Nevertheless, to explain the observations made by Yan *et al.* (2015), Caliskan *et al.* (2017) suggest that the *in vitro* translation system used to accumulate the peptides for mass spectrometry and to perform optical tweezers experiments may over time become depleted of some aa-tRNAs, which may facilitate ribosome pausing and excursions into alternative frames.

3.5.2 Role of (p)ppGpp in translational errors

In *E. coli*, several works have studied the effects of the stringent response on translational errors and have focused on the role of (p)ppGpp. Most of them have used two *E. coli* strains: the *relA*$^+$ and the *relA*$^-$. *relA*$^+$ is able to detect amino acid limitation while *relA*$^-$ cannot since it lacks the *relA* gene which has been deleted, and thus it cannot produce (p)ppGpp under starvation conditions.

Missense errors and premature terminations are increased in a *relA*$^-$ strain [O'Farrell, 1978]. When starving *E. coli relA*$^+$ and *relA*$^-$ strains for a particular amino acid, more spots corresponding to mischarged proteins appear on electrophoresis gel. In the same manner, on SDS page gels, bands corresponding to shorter proteins appear for starved *relA*$^-$ strains but not for starved *relA*$^+$ strains. From these observations, O'Farrel (1978) suggest that (p)ppGpp's main role is to decrease the protein synthesis level so that it fits to the low availability of the limiting amino acid. Thus, this would prevent the deleterious effects (misreading and premature termination) caused by starvation.

Furthermore, in *E. coli*, the leftward frameshift level can be influenced by the *relA* genotype when the bacterium is subjected to amino acid starvation through the use of analogues [Masucci et al., 2002]. A mutated *lacZ* gene reporter (where known slippery sequences have been inserted) has been cloned in the *relA*$^+$ and *relA*$^-$ strains. Each frameshift-prone sequence possesses a 'hungry' codon encoding an amino acid for which a drug analogue has been used to mimic starvation (*i.e.* the corresponding aa-tRNA becomes limiting upon analogue injection). For both genotypes the frameshift level is enhanced

and this phenomenon is even more pronounced in the strain where the *relA* gene is deleted. Masucci *et al.* (2002) presume that the limiting aa-tRNA stimulates ribosome pausing at hungry codons and thus frameshifting, as seen in section 3.3.1. They propose that the differences observed between the two *relA* genotypes are due to a larger fraction of the aa-tRNA undermodified forms of the tRNA in question as compared to its overall population. Indeed, these forms are more prone to slippage as seen in section 3.3.2.2.

Several studies have tried to understand the role of (p)ppGpp in maintaining a high translational accuracy, especially concerning amino acid misincorporation. A first hypothesis concern the interaction of (p)ppGppp with EF-Tu which could be responsible for maintaining a high translational accuracy during the stringent reponse [Gerhart et al., 1982] [Dix and Thompson, 1986] but it has been refuted [Rojas and Ehrenberg, 1991]. The effects of (p)ppGpp on translation fidelity might actually be indirect: it leads to the reduction of the RNA synthesis rate so the mRNA level eventually decreases and becomes limiting for protein synthesis [Sørensen et al., 1994]. Consequently, there is a reduced demand for the aa-tRNA that are lacking which prevents a rise in mistranslation during stringent response. This explanation is supported by the fact that during a nutritional downshift, the protein synthesis rate is the same for *relA*$^+$ and *relA*$^-$ strains [Johnsen et al., 1977], which means that in both strains new proteins are produced at the rate allowed by the supply of the limiting substrate.

3.5.3 Impact of amino acid starvation on the charging level of tRNAs and consequences for the translational errors

The stringent response leads to an important reduction of the tRNA's charging level for which the amino acid is lacking [Sørensen, 2001]. This reduction varies according to the amino acid but mostly goes from 70-80% charging level during steady state and decreases to less than 5% charging level after starvation of the corresponding amino acid. These reductions in the charging levels of tRNA are important in both *relA*$^+$ and *relA*$^-$ strains; but after amino acid starvation the residual charging of tRNAs in the *relA*$^-$ strain are lower than for the *relA*$^+$ strain. Sorensen (2001) suggests that this difference could be responsible for the higher misincorporation level found in a *relA*$^-$ strain as compared to the *relA*$^+$ strain. He explains that in the *relA*$^-$ strain, RNA synthesis continues since there are no (p)ppGpp metabolites to inhibit transcription. Thus protein synthesis is only limited by the availability of the starved amino acid; and the high concentration of ribosomes which are in demand of the corresponding aa-tRNA causes a charging level of this particuler aa-tRNA close to zero. Consequently, the ribosomal A-site remains empty for a longer period, enhancing the probability that near-cognate aa-tRNAs are accepted, which results in a higher misincoporation level in the *relA*$^-$ strain than in the *relA*$^+$ strain. The ribosome pause at the codon corresponding to the starved amino acid present in the A-site has been confirmed and it leads to ribosome traffic jam at this specific site [Subramaniam et al., 2014]. This ribosome pausing can be deleterious to the cell but mechanisms exist to abort translation [Shoemaker and Green, 2012], which is particularly important during stress [Keiler and Feaga, 2014].

Moreover, the majority of tRNAs are degraded upon amino acid starvation, including both the cognate and non-cognate tRNAs of the amino acid *E. coli* is starved for [Svenningsen et al., 2016]. This occurs in both *relA*$^+$ and *relA*$^-$ strains which means that starvation induced tRNA degradation is independent of the (p)ppGpp-mediated stringent response. Inhibiting mRNA synthesis also results in tRNA degradation, which leads to the conclusion that tRNA degradation is a general part of *E. coli*'s response to nutritional downshift [Svenningsen et al., 2016]. The charging level of the tRNA corresponding to the amino acid the cell is starved for is decreased while the charging level of the tRNAs accepting other amino acid is slightly increased. This is likely due to the reduction of the total tRNA pool. Svenningsen *et al.* (2016) propose a passive model for tRNA degradation where tRNAs engaged in translation are not degraded while the excess tRNA (*i.e.* not involved in translation) is degraded by unknown factors. They add that this overall tRNA degradation upon amino acid starvation has for purpose to prevent a too high translational error rate.

Overview: Translational errors, including misincorporation as well as rightward and backward frameshifts, increase upon amino acid starvation. This phenomenon is amplified in a strain which cannot produce (p)ppGpp (*relA*$^-$). A nutritional downshift also provokes -4 and +2-nucleotide frameshifts, a drop in the uncharged tRNA levels as well as translation abortion due to a ribosome traffic jam at the starved amino acid. Besides, short (truncated) proteins are produced in important amount in a *relA* $^-$ strain upon amino acid starvation.

A possible explanation for (p)ppGpp's role in maintaining translation accuracy is that it decreases the mRNA level which decreases the demand for the starved amino acid. The charging level of the tRNA corresponding to the depleted amino acid drops for both *relA*$^+$ and *relA*$^-$ strains, and more importantly for *relA*$^-$. Thus, a high translation accuracy might also be maintained through the tRNA degradation upon nutritional shift.

A proposed explanation for a decreased translation accuracy in absence of (p)ppGpp is that mRNA synthesis remains high in a *relA* $^-$ strain. This increases the demand for the starved amino acid and leads to an aa-tRNA charging level of nearly zero. Consequently, the ribosome pauses longer which can lead to a higher level of misincorporation: near-cognate aa-tRNAs compete with the starved aa-tRNAs and are more likely to be chosen since their concentration is higher.

Part II

Results

Chapter 4

Problematic

The (p)ppGpp has been known for decades as an essential metabolite which helps bacteria to adapt to nutritional changes (section 2.3.1). Moreover, it has been shown that it is involved in the control of protein synthesis errors (section 3.5). Indeed, it is crucial for the bacterium to maintain its protein synthesis error at a certain level: low enough to prevent irreversible damages but high enough to provide the cell with fitness advantages.

The two laboratories implicated in this work have conducted researches on how bacteria adapt to environmental changes at the replication, transcription and translation levels. They particularly focused on how these cellular processes are affected during steady-state growth, and how this can be translated in terms of resource allocation. Here, we support the view that resource allocation is orchestrated by the GTP and (p)ppGpp upon nutritional changes but also in steady-state growth conditions. Indeed, the GTP and (p)ppGpp regulations presented in section 2.3 can be seen as feedback loops: the GTP positively impacts the transcription of genes involved in the translation machinery and in particular stimulates its own production (positive feedback loop); while (p)ppGpp inhibits translation initiation (fast negative feedback loop) and GTP production which in fine represses the translation machinery (slow negative feedback loop).
When RelA encounters an uncharged tRNA, it produces (p)ppGpp from GTP which directly inhibits translation and indirectly transcription in *B. subtilis*. In the absence of RelA, and consequently of intracellular (p)ppGpp, the uncharged tRNAs are not sensed which impacts resource allocation. Indeed, after *E. coli* strains were starved for a particular amino acid, reductions in the charging levels of the corresponding tRNA were important in both *relA*+ and *relA*- strains; but after amino acid starvation the residual charging of tRNAs in the *relA*- strain were lower than for the *relA*+ strain [Sørensen, 2001]. Similar observations are expected to be found for *B. subtilis* RelA^+ and $(\text{p})\text{ppGpp}^0$ strains. Sorensen (2001) suggested that the lower tRNA charging level of the *relA*- strain as compared to the *relA*+ strain could be responsible for the higher misincorporation level found in a *relA*- strain as compared to the *relA*+ strain [O'Farrell, 1978]. Sorensen (2001) explained that the ribosomal A-site remains empty for a longer period, enhancing the probability that near-cognate aa-tRNAs are accepted, which results in a higher misincoporation level in the *relA*- strain than in the *relA*+ strain. The same reasoning

can also be applied to translational frameshift errors. Indeed, the longer the ribosome is stalled at an empty A-site, the more likely a frameshift event can occur [Urbonavičius et al., 2001][Caliskan et al., 2017]. Misincorporation of a near-cognate tRNA in the A-site also stimulates frameshifting [Jäger et al., 2013] and misincorporation levels are higher in an *E. coli relA*- strain (and so we assume in a (p)ppGpp0 strain) [O'Farrell, 1978]. To summarize, translational errors (misincorporations and frameshifts) are expected to be higher when the charging levels of tRNAs is lower and the other way around; which means that monitoring the translational errors level would also provide information on the charging levels of tRNAs.

Hence, the goal of this project was to better understand how the translational errors, the transcription and the translation are affected during nutritional shifts by the GTP/(p)ppGpp regulatory feedback loops.

To better understand the link between resource allocation and translational errors, we divided this work in two main axes. In a first part, we studied the occurrence of translational frameshift errors during growth in different media when perturbing the GTP/(p)ppGpp regulatory feedback loops. In this context, we built fluorescent translational frameshift error reporter systems and transformed them into WT and (p)ppGpp-deficient cells (referred to as "(p)ppGpp^0" and obtained by deleting the three genes encoding the (p)ppGpp synthetases RelA , RelP, and RelQ); and then monitored the fluorescence level. We modulated the intracellular level of GTP by inserting the gene encoding GuaB under the control of the isopropyl-β-D-1-thiogalactopyranoside inducible hyperspank promoter (P_{hs}). Eventually, we mimicked the role of (p)ppGpp on IF2 by using drugs known to inhibit translation initiation and injected them before the cells entered the stationary phase.
In a second part, we looked at how these perturbations affect the transcription process as well as the ribosome synthesis. We modulated the intracellular level of GTP, similarly to the first part, to observe the effects of variations in GTP concentration on the transcription initiation of synthetic genetic constructs with transcription start sites of different composition in nucleotides (guanine versus adenine). Indeed, the effects of varying the GTP concentrations on transcription initiation through the transcription start site sequence have only been demonstrated in vitro [Krásnỳ et al., 2008]. Then, we focused on how ribosome production is affected by the deregulation of the GTP/(p)ppGpp regulatory feedback loops by monitoring the synthesis of ribosomal RNAs. Indeed, all *rrns* possess a guanine as transcription start site, thus their transcription is expected to be impacted by the modulation and the regulation of intracellular GTP levels; which in fine should impact ribosome synthesis.

Overall, these results provide insights into how the very same perturbation of the GTP/(p)ppGpp regulatory feedback loops affect nutrient supply and demand and consequently translational error rates.

Chapter 5

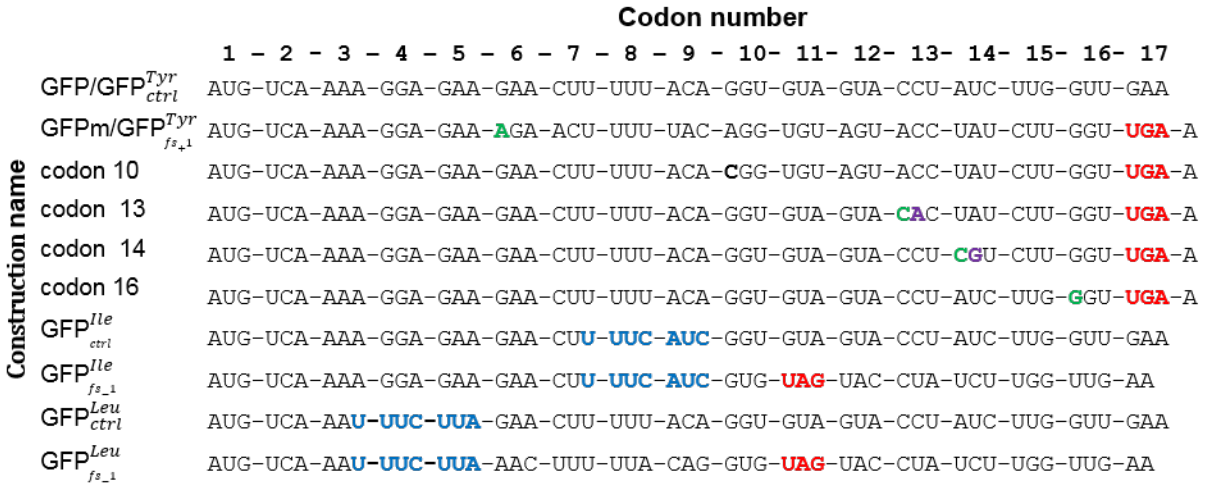
The (p)ppGpp alarmone prevents translational frameshift errors via its feedback control on translation initiation in *B. subtilis*

5.1 Design of distinct reporter systems for the detection of leftward and rightward translational frameshift errors in *B. subtilis*.

In order to explore the relative contributions of GTP and (p)ppGpp to the occurrence of translational frameshift errors (TFEs) in *B. subtilis*, our strategy was to build intrinsically distinct TFE reporter systems. To do so, we designed leftward (-1) and rightward (+1) TFE reporter systems making use of a chromosomally encoded reporter fusion between a strong constitutive promoter (P_{veg}) and the gene coding the superfolder Green Fluorescent Protein (GFP) (sfGFP, simply named GFP from here and below) [Guiziou et al., 2016]. To monitor leftward (-1) TFEs, the *gfp* gene was modified by the introduction of frameshift prone sequences previously characterized in *E. coli* (Figure 5.1A) [Kolor et al., 1993, Barak et al., 1996, Lindsley and Gallant, 1993]. These sequences trigger frameshift events through a ribosomal pause when either the leucine or isoleucine codons (i.e. the hungry codons) are present in the ribosomal A-site. The leftward (-1) TFE reporter systems will therefore be referred to as the GFP^{Leu}_{fs-1} and GFP^{Ile}_{fs-1} reporter systems, respectively. The biosynthesis of branched chain amino acids in *B. subtilis* is under the control of the GTP-regulated CodY transcription factor [Sonenshein, 2007]. We thus expect the GFP^{Leu}_{fs-1} and GFP^{Ile}_{fs-1} reporter systems to be specifically influenced by variations in the corresponding amino acid-charged tRNA (i.e. $tRNA^{Ile}$ and $tRNA^{Leu}$) following variations in (p)ppGpp and GTP abundances. To monitor rightward (+1) TFEs, the *gfp* gene was modified based on a TFE reporter system previously described in *B. subtilis* [Meyerovich et al., 2010], which consisted in the insertion of one nucleotide upstream the 6^{th} codon (Figure 5.1A). During translation, the ribosome will rapidly encounter an UGA stop codon at position 17 and GFP will

be produced only upon a +1 frameshift event upstream the 17th codon. Using a series of mutated *gfp* sequences, the TFE prone site has been located somewhere in between the 10th and 17^{th} codons [Meyerovich et al., 2010]. In order to precisely map the TFE prone site, we established a ranked list of the putative frameshift sites including and following the 10^{th} codon (Figure 5.1B) on the basis of commonly-accepted criteria: (i) the shift event is induced by a ribosome pause likely resulting from the lack of the amino acid-charged tRNA (aa-tRNA) corresponding to the codon present in the A-site [Farabaugh and Björk, 1999, Urbonavičius et al., 2001, Harger et al., 2002], (ii) the aa-tRNA anticodon present in the P-site must form a stable interaction with the newly formed codon following the shift event [Curran, 1993] and (iii) the frameshift event is facilitated if the newly formed codon in the A-site is decoded by an abundant aa-tRNA [Pande et al., 1995]. The ranked list pointed out to four key candidate frameshift prone sites (ranked 1 to 4; Figure 5.1B). We modified the *gfp* sequence so that one nucleotide was inserted right upstream each of the corresponding codons (Figure 5.1A). As expected, a one-nucleotide insertion upstream the 10th codon led to a GFP production similar to that of the one-nucleotide insertion upstream the 6th codon (Figure 5.1C). However, a one-nucleotide insertion upstream the 13^{th}, 14^{th} or 16^{th} codon did not yield GFP (Figure 5.1D), which indicated that each of these one-nucleotide insertions suppressed the frameshift event. We concluded that the frameshift event is most likely triggered by a ribosomal pause when the fourteenth codon (which encodes for tyrosine) is present in the ribosomal A-site (Figure 5.1E). The rightward (+1) TFE reporter system will thus be referred to as the GFP^{Tyr}_{fs+1} reporter system. As the intracellular tyrosine abundance was shown to be (p)ppGpp independent [Kriel et al., 2012], the GFP^{Tyr}_{fs+1} reported TFEs are not expected to be additionally influenced by variation in the tRNA^{Tyr} abundance with respect to variations in GTP and (p)ppGpp abundances.

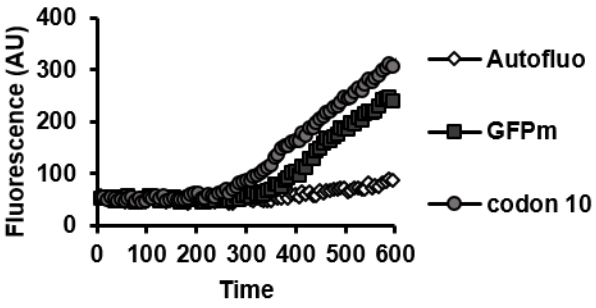
A



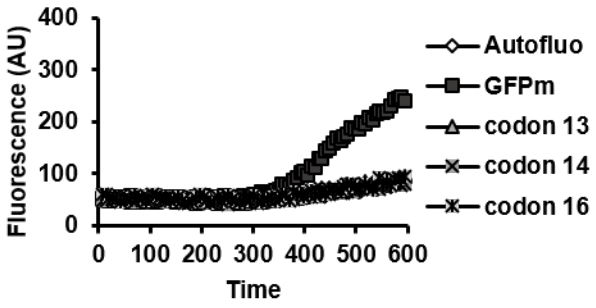
B

Potential shift sites (black = E-site, green = P-site, red = A-site)	P-site corresponding tRNA	mRNA:anticodon pairs for the +1 realignment	Number of stable interactions	tRNA encoding the "hungry" codon	Expected intracellular abundance of the tRNA encoding the following codon (A-site)	Frame-shift likelihood ranking
9 – 10– 11 UAC–AGG–UGU	tRNA$^{Arg}_{CCU}$	G:U G:C U:C	2	tRNA$^{Cys}_{GCA}$	Low	1
10 – 11– 12 AGG–UGU–AGU	tRNA$^{Cys}_{GCA}$	G:A U:C A:G	0	tRNA$^{Ser}_{GCU}$	Medium	7
11 – 12– 13 UGU–AGU–ACC	tRNA$^{Ser}_{GCU}$	G:U U:C A:G	1	tRNA$^{Thr}_{GGU}$	Low	6
12 – 13– 14 AGU–ACC–UAU	tRNA$^{Thr}_{GGU}$	C:U C:G U:G	2	tRNA$^{Tyr}_{GUA}$	Medium	1
13 – 14– 15 ACC–UAU–CUU	tRNA$^{Tyr}_{GUA}$	A:A U:U C:G	2	tRNA$^{Leu}_{GAG}$	Low	1
14 – 15– 16 UAU–CUU–GGU	tRNA$^{Leu}_{GAG}$	U:G U:A G:G	2	tRNA$^{Gly}_{UCC}$ and tRNA$^{Gly}_{GCC}$	High	4
15 – 16– 17 CUU–GGU–UGA	tRNA$^{Gly}_{UCC}$ and tRNA$^{Gly}_{GCC}$	G:C U:C U:U and G:C U:C U:G	2	RF2 (UGA stop codon)	Doesn't apply	5

C



D



E

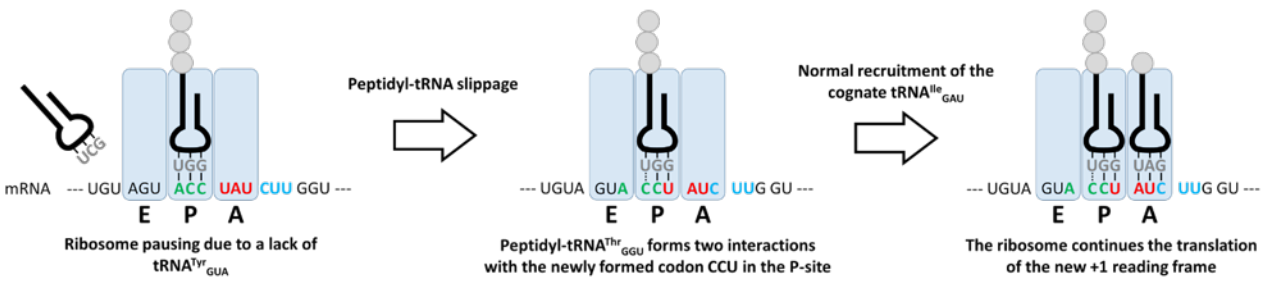


Figure 5.1: Frame-shift prone sequences used in this study and mapping of the +1 frameshift-prone site. A. Native and modified sequences of the *gfp* gene used to report translational frameshift errors during translation and to map the +1 frameshift prone site. The first sequence GFP corresponds to the native one; the GFPm sequence is the one used by Meyerovich *et al.* (2010) to report +1 frameshifts. The sequences referred to as codon 10, codon 13, codon 14 and codon 16 are the ones used to map the +1 frameshift prone site. The nucleotides added which lead to the formation of a potential frameshift prone sequence are in green. The nucleotide in purple is a nucleotide modified from the native sequence to ensure that the newly formed sequence is not also slippery (table 8.9). The nucleotides in blue correspond to the -1 frameshift prone sequence where the codon present in the A-site encodes either an isoleucine (GFP^{Ile}_{ctrl} and GFP^{Ile}_{fs-1}) or a leucine (GFP^{Leu}_{ctrl} and GFP^{Leu}_{fs-1}). The stop codons formed after adding a nucleotide to or suppressing a nucleotide from the GFP sequence are in bold. B. Potential frameshift sites of the modified GFP sequence (GFPm). The fourth column indicates the number of stable interactions the P-site tRNA can make with the shifted mRNA sequence based on criteria described in Curran (1993). The sixth column is based on the number of genes encoding the tRNA corresponding to the "hungry codon" as well as its intracellular content in *B. subtilis* [Kanaya et al., 1999]. In the ultimate column, the sequences are ranked for their likelihood to be a frameshift site based on the information from columns 4 and 6. C and D. Fluorescence levels of the strains carrying the mutated GFP sequences as shown in panel A as function of time and where the Autofluo curve corresponds to the autofluorescence of the WT strain. E. Mechanism of the mapped +1 frameshift: the lack of $tRNA_{GUA}$ charged with tyrosine induces a ribosomal pause which can allow the P-site $tRNA_{GGU}$ to slip forward and form a quite stable interaction with the codon CCU. Thus, the reading frame is shifted +1 and the newly formed codon in the A-site is AUC. The ribosome resumes translation of the new reading frame by recruiting a charged $tRNA^{Ile}_{GAU}$ which corresponds to the AUC codon.

5.2 Translational frameshift errors are increased in the absence of (p)ppGpp in exponentially growing *B. subtilis* cells

We asked whether TFEs in *B. subtilis* are more frequent in a (p)ppGpp^0 strain (i.e. a Δ*relA* Δ*relP* Δ*relQ* triple mutant; Table 8.9) than in the wild-type (WT) strain as observed in *E. coli* [Masucci et al., 2002]. The GFP primary sequences consecutive to frameshift events differ for each of our reporter systems. To be able to compute the TFE rate previously defined as the ratio of the $GFP^x_{fs\pm1}$ and GFP^x_{ctrl} produced over a period of time [Meyerovich et al., 2010], we constructed the corresponding control GFP reporter systems (GFP^{Tyr}_{ctrl} , GFP^{Leu}_{ctrl}, GFP^{Ile}_{ctrl}; Figure 5.1A). The $GFP^x_{fs\pm1}$ and GFP^x_{ctrl} TFE reporter systems were transformed into WT, RelA$^+$ (i.e. a Δ*relP* Δ*relQ* double mutant) and (p)ppGpp^0 cells. Strains were grown in M9 minimal medium with glucose (M9G) as carbon source. The WT, RelA$^+$ and derivative strains showed growth rates of 0.70 ± 0.02 h^{-1}, higher than that of the (p)ppGpp^0 and derivative strains of 0.30 ± 0.02 h^{-1}(Figure 5.3A). GFP production was detected in

the exponential growth phase of all strains carrying the $GFP^{x}_{fs\pm1}$ reporter systems, with the GFP^{Leu}_{fs-1} reporter showing the highest fluorescence signal and the GFP^{Ile}_{fs-1} exhibiting the lowest one (Figure 5.2). These results indicated that leftward and rightward TFEs occurred in such growth conditions and yielded functional GFP molecules. In the WT and RelA$^+$ cells, the computed TFE rates were of 0.25 %, 0.35 % and 1.50 % when using the GFP^{Tyr}_{fs+1}, GFP^{Ile}_{fs-1} and GFP^{Leu}_{fs-1} reporter systems, respectively (Figure 5.3B, C and D). These results reinforced the idea that both the occurrence and level of TFEs depend on the primary sequence of the frameshift prone sites. In the (p)ppGpp^0 cells, the TFE rates were of 0.40 %, 0.55 % and 3.2 % when using the GFP^{Tyr}_{fs+1}, GFP^{Ile}_{fs-1} and GFP^{Leu}_{fs-1} reporter systems, respectively (Figure 5.3B, C and D). It is worth to note that the TFE rates have approximately doubled in the absence of (p)ppGpp. Similar results were obtained when growing the WT, RelA$^+$, (p)ppGpp^0 and derivative strains in M9 with malate as a gluconeogenic carbon source (M9M, Figure 5.4). Altogether, our results showed that TFEs in *B. subtilis* are more frequent in a (p)ppGpp^0 background than in WT and RelA$^+$ cells, whatever the frameshift prone sequence is used. As the TFE rates were identical between the WT and RelA$^+$ strains, we further concluded that the RelP and RelQ (p)ppGpp synthetases did not contribute to the occurrence of TFEs in exponentially growing cells.

5.3 The TFE rate increases together with the induction of GTP biosynthesis in steady-state growth of (p)ppGpp^0 cells.

As (p)ppGpp is known to negatively regulate GTP biosynthesis via the enzymes Gmk, GuaB and HprT (Figure 5.5A), an excess of GTP production in (p)ppGpp^0 cells grown in minimal media may be key to the emergence of TFEs. To test the hypothesis of whether controlling GTP biosynthesis might prevent TFEs in (p)ppGpp^0 cells, we constructed RelA$^+$ and (p)ppGpp^0 strains in which the gene coding GuaB (enzyme catalyzing the reaction leading to XMP, a GTP precursor; Figure 5.5A) has been deleted (ΔguaB) and ectopically reinserted under the control of an isopropyl-β-D-1-thiogalactopyranoside (IPTG) inducible promoter. These strains will be referred to as RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible}, respectively (Table 8.9). As the intracellular tyrosine abundance was shown to be (p)ppGpp independent [Kriel et al., 2012], we specifically selected the GFP^{Tyr}_{fs+1} reporter system to monitor the TFE rate in these two strains grown in M9G. As shown on Figure 5.5B, the growth rate of each strain was IPTG dependent until 15 μM IPTG, where it reached ≈0.69 h^{-1}. This value is similar to the growth rates of the WT and RelA$^+$ strains grown in M9G. Beyond 15 μM IPTG, the growth rate constantly decreased with increasing IPTG concentration for the (p)ppGpp^0 *guaB*^{inducible} strain, while it remained constant for the RelA$^+$ *guaB*^{inducible} strain. These data suggested that a 15 μM IPTG mediated induction of *guaB* expression in M9G-grown cells led to a level of *guaB* expression similar to that in wild-type cells, while additional overproduction of GuaB may have induced a severe deregulation. Consistently, for IPTG concentrations below or equal to 15 μM, the GFP^{Tyr}_{fs+1} associated TFE rate was of 0.25% for both the RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0

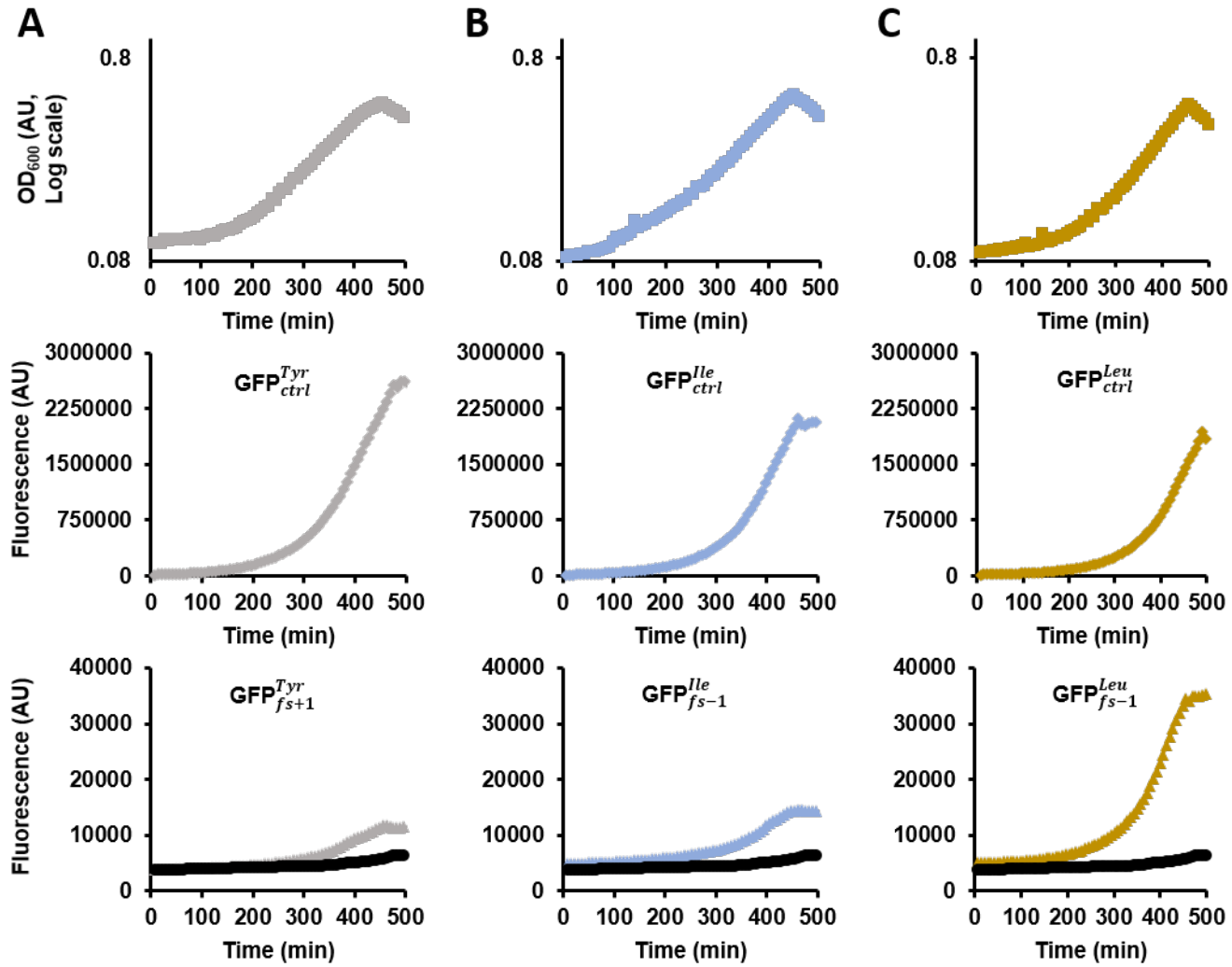


Figure 5.2: Growth and fluorescence levels of the strains carrying the different frameshift reporter systems grown in M9G. A, B and C. Top panels represent the growth curves of the RelA$^+$ strain. The middle panels represent fluorescence levels as function of time of the GFP used to report the overall GFP production for the different slippery sequences (GFP^x_{ctrl}). The bottom panels represent the fluorescence levels as function of time of the GFP used to report the occurrence of a frameshift error during translation for the different slippery sequences ($GFP^x_{fs\pm1}$). A. Representation of the data for the GFP^{Tyr}_{fs+1} reporter system. B. Representation of the data for the GFP^{Ile}_{fs-1} reporter system. C. Representation of the data for the GFP^{Leu}_{fs-1} reporter system.

guaB^{inducible} strains (Figure 5.5C), which was similar to the values obtained for the WT and RelA$^+$ strains grown in M9G. For higher induction levels of *guaB* expression, the GFP^{Tyr}_{fs+1} associated TFE rate in the (p)ppGpp^0 *guaB*^{inducible} strain increased more than twice up to ≈0.65% (when using 50 μM IPTG), while it remained constant at ≈0.25% for the RelA$^+$ *guaB*^{inducible} strain. Similarly, the growth rates of the two strains grown in M9M were IPTG dependent until 15 μM IPTG, where they reached ≈0.65 h^{-1}, and the GFP^{Tyr}_{fs+1} associated TFE rates were of about 0.2%, which is similar to that found in wild-type cells grown in M9M (Figure 5.5D and 5.5E). For higher induction levels of *guaB* expres-

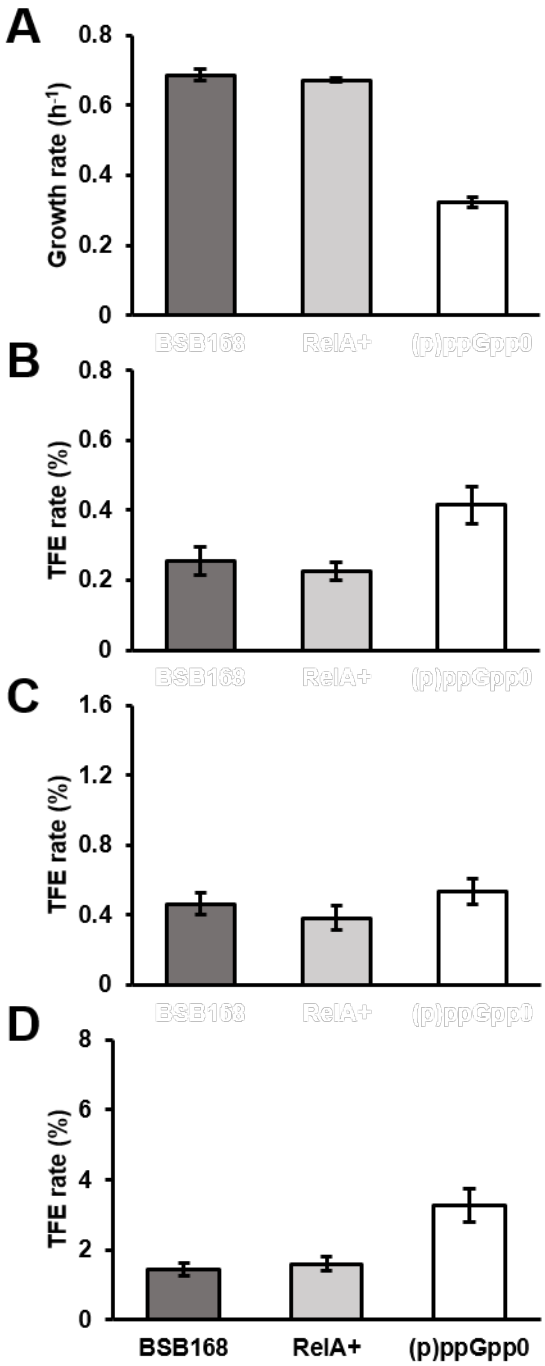


Figure 5.3: Translational frameshift errors during steady-state growth in M9G medium. A. Growth rates of the WT, RelA$^+$ and (p)ppGpp^0 strains grown in M9G. B. GFP^{Tyr}_{fs+1} TFE rate of the WT, RelA$^+$ and (p)ppGpp^0 strains during steady-state growth in M9G. C. GFP^{Ile}_{fs-1} TFE rate of the WT, RelA$^+$ and (p)ppGpp^0 strains during steady-state growth in M9G. D. GFP^{Leu}_{fs-1} TFE rate of the WT, RelA$^+$ and (p)ppGpp^0 strains during steady-state growth in M9G. A, B, C and D. The error bars correspond to the standard deviation of the growth rate or TFE rate calculated using the bootstrap method (see section 8.7.2.5).

sion, the GFP^{Tyr}_{fs+1} associated TFE rate remained constant in the RelA$^+$ $guaB^{inducible}$ strain, while it increased in the (p)ppGpp^0 $guaB^{inducible}$ strain up to ≈1.5% (when using 200 μM IPTG). TFEs

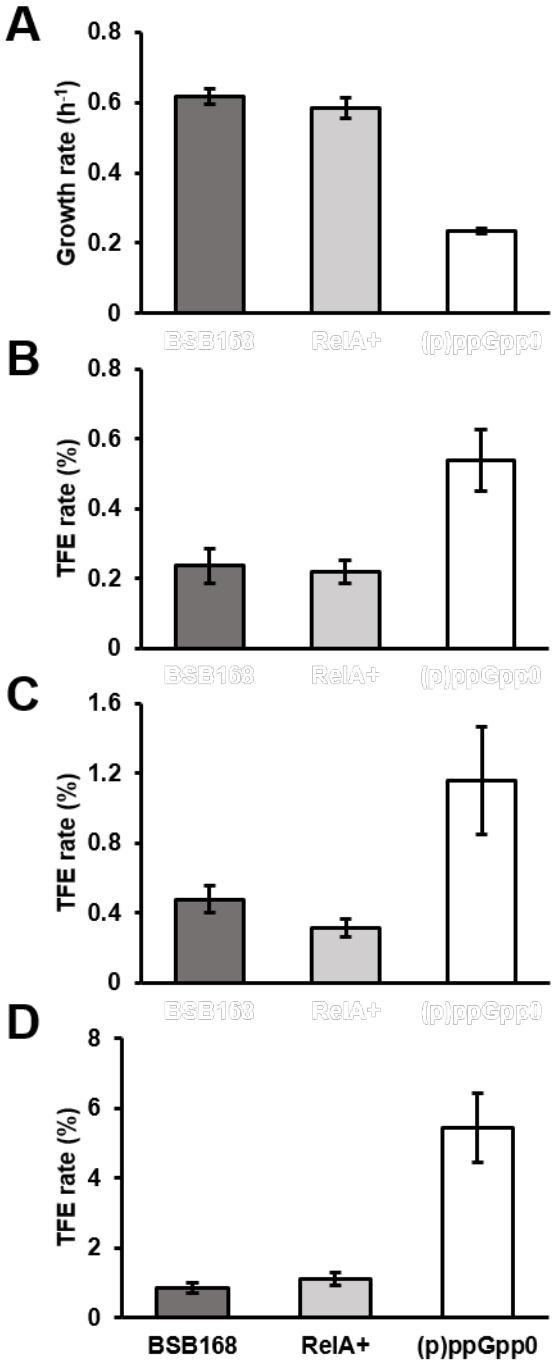


Figure 5.4: Translational frameshift errors during steady-state growth in M9M medium. A. Growth rates of the WT, RelA$^+$ and (p)ppGpp^0 strains grown in M9M. B. GFP^{Tyr}_{fs+1} TFE rate of the WT, RelA$^+$ and (p)ppGpp^0 strains during steady-state growth in M9M. C. GFP^{Ile}_{fs-1} TFE rate of the WT, RelA$^+$ and (p)ppGpp^0 strains during steady-state growth in M9M. D. GFP^{Leu}_{fs-1} TFE rate of the WT, RelA$^+$ and (p)ppGpp^0 strains during steady-state growth in M9M. A, B, C and D. The error bars correspond to the standard deviation of the growth rate or TFE rate calculated using the bootstrap method (see section 8.7.2.5).

occurrence was thus growth condition-dependent upon high induction levels of *guaB* (1.5% in M9M to be compared to 0.65% in M9G). Our results also indicated that even for highly sub-optimal induction

of GTP biosynthesis leading to much lower growth rates than that of the WT, TFEs occurred at a basal level equivalent to that of the WT (0.25%). Altogether, these findings prompted us to conclude that beyond the basal TFE level GTP itself may be the TFE trigger factor, and that (p)ppGpp was not strictly required to reduce TFE occurrence but most likely maintained lower TFE rates in WT cells via its feedback regulation on GTP biosynthesis.

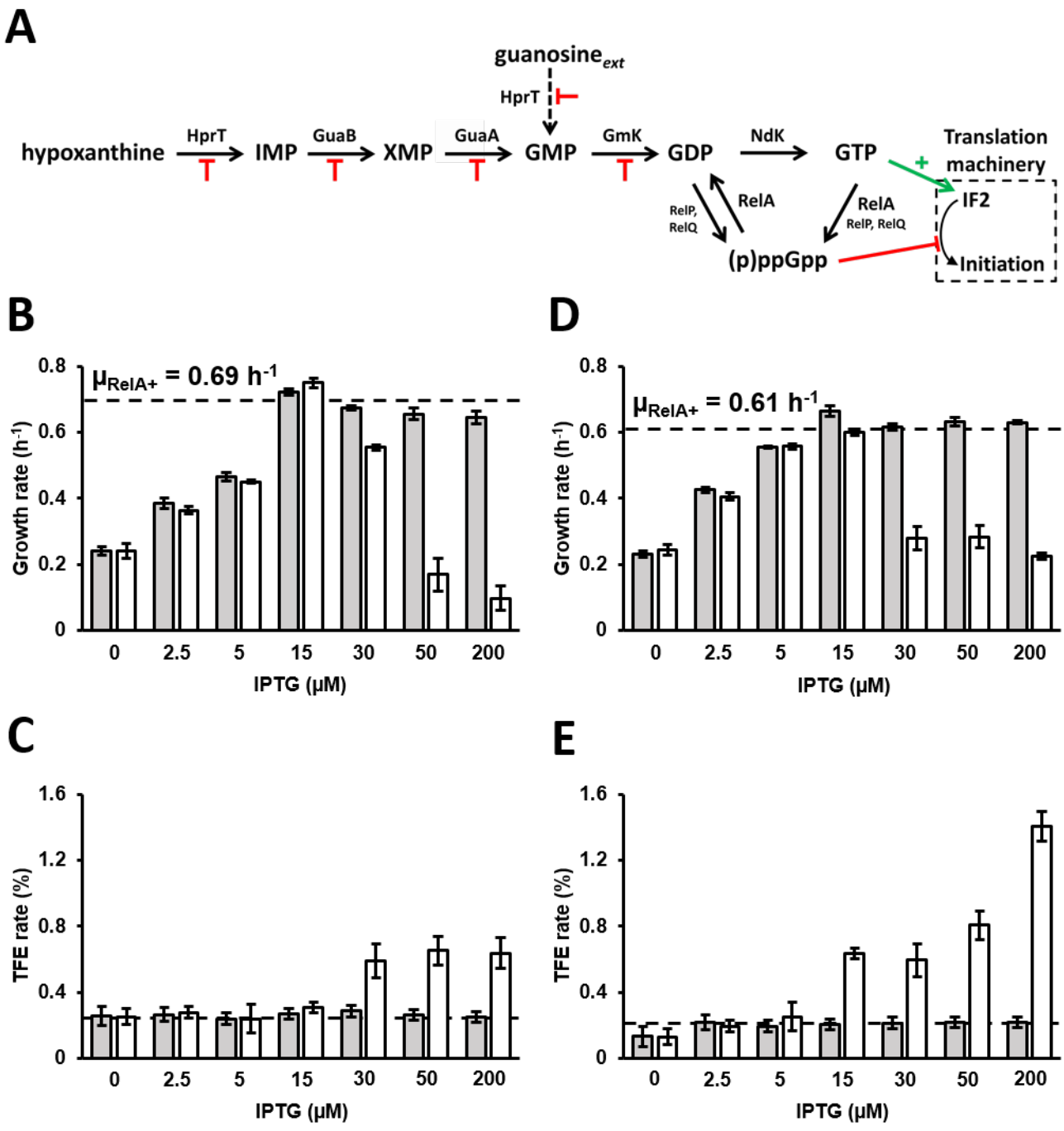


Figure 5.5: Impact of the control of GTP levels on the TFE rate and the growth rate during steady-state growth in poor medium (M9G and M9M). A. Scheme of the metabolic pathways which lead to GTP and (p)ppGpp synthesis. The enzymes which catalyze the reactions are written at the top of the reaction's arrow. The dashed arrow corresponds to multiple reactions which lead to GMP synthesis from extracellular guanosine. The red blunt-end arrows correspond to the inhibition of the corresponding enzymatic reaction. The green arrow indicates that GTP is required to initiate translation by binding to the translation initiation factor IF2. (p)ppGpp binds to IF2 with a higher affinity than GTP [Mitkevich et al., 2010] which inhibits translation initiation (red blunt-end arrow). B. Mean of the growth rate of the RelA$^+$ $guaB^{inducible}$ and (p)ppGpp^0 $guaB^{inducible}$ strains grown in M9G with different concentrations of IPTG (in μM). C. Mean of the TFE rate of the RelA$^+$ $guaB^{inducible}$ and (p)ppGpp^0 $guaB^{inducible}$ strains grown in M9G with different concentrations of IPTG (in μM). The dashed line corresponds to the TFE rate of the RelA$^+$ strain grown in M9G which is of $\approx 0.25\%$. D. Mean of the growth rate of the RelA$^+$ $guaB^{inducible}$ and (p)ppGpp^0 $guaB^{inducible}$ strains grown in M9M with different concentrations of IPTG (in μM). E. Mean of the TFE rate of the RelA$^+$ $guaB^{inducible}$ and (p)ppGpp^0 $guaB^{inducible}$ grown in M9M with different concentrations of IPTG (in μM). The dashed line corresponds to the TFE rate of the RelA$^+$ strain grown in M9M which is of $\approx$ 0.22%. In B, C, D and E the error bars correspond to the standard deviation of the growth rate or TFE rate calculated using the bootstrap method (see section 8.7.2.5).

5.4 The TFE rate peaks in the transition to the stationary phase in the absence of *relP* and *relQ*, a phenomenon significantly amplified in the absence of (p)ppGpp.

In response to amino acid starvation, the (p)ppGpp level rises in bacteria, which triggers the stringent response [Dalebroux and Swanson, 2012]. In a (p)ppGpp^{0}strain exhibiting a 'relaxed' phenotype, the stringent response does not take place. We therefore asked whether the absence of (p)ppGpp during the transition to the stationary phase leading to an uncontrolled GTP biosynthesis would generate more TFEs consecutive to a sudden aa-tRNA depletion. To answer this question and properly evaluate the TFE frequency across growth phases and conditions beyond the steady-state growth, we first had to redefine the TFE rate as the ratio of the specific production rates of $GFP^x_{fs\pm1}$ and GFP^x_{ctrl}. Indeed, we formally proved and experimentally demonstrated that the redefined computation method of the TFE rate is equivalent to the former computation method when applied to exponentially growing cells but can also be applied to non-steady state growth (see section 8.8). We next monitored TFEs across the different growth phases of the WT, RelA$^+$ and (p)ppGpp^0 strains grown in CH rich medium and computed the corresponding TFE rates. All strains exhibited similar growth phenotypes and exhibited growth rates of about 1.5 h^{-1} during the exponential growth phase (Figure 5.6). Growth rates then gradually decreased and the transition to the stationary phase was accompanied by a net growth rate inflexion leading to a lower growth rate at 0.3 $\pm$ 0.1 h^{-1} for over 5 hours before growth arrest

(Figure 5.6, second panels). The rightward TFE rates (monitored using the GFP^{Tyr}_{fs+1} reporter system) during the exponential growth and transition phases in the WT, RelA$^+$ and (p)ppGpp^0 strains were approximately constant around 0.2% (Figure 5.6 third panels). It is worth to note that this value is similar to the TFE rates observed in steady-state growth of WT and RelA$^+$ cells in M9G and M9M. Then, further in stationary phase, the WT TFE rate slowly increased to reach a plateau around 0.6-0.8% (Figure 5.6A, third panel). This observation is consistent with previous works where the rightward TFEs are found to be increased during stationary phase in *B. subtilis* [Meyerovich et al., 2010]. However, in the transition to stationary phase the TFE rate of the RelA$^+$ strain peaked up to ≈0.9% before decreasing down to 0.6% and reaching the same plateau as the WT TFE rate (≈0.6-0.8%) in the late stationary phase (Figure 5.6B, third panel). The TFE rate of the (p)ppGpp^{0}strain showed a burst up to ≈4.5% in the transition to the stationary phase (Figure 5.6C, third panel). Then, further in the stationary phase, the TFE rate decreased and stabilized at ≈0.8%, which is similar to the values found in WT and RelA$^+$ cells. Given the elevated values of the measured TFE rates, we repeated the experiments using the leftward GFP^{Ile}_{fs-1} and GFP^{Leu}_{fs-1} reporter systems. As shown on Figure 5.6 (fourth panels), the GFP^{Ile}_{fs-1} TFE rate of the WT, RelA$^+$ and (p)ppGpp^0 strains remained constant around ≈0.3% until the transition to stationary phase, where it slowly increased for the WT but peaked up to 0.4% in the RelA$^+$ strain and up to 1.0% in the (p)ppGpp^{0}strain. Then, further in stationary phase, the TFE rates slowly converged for all strains to a value of 0.6%. Similarly, the GFP^{Leu}_{fs-1} TFE rate in the WT, RelA$^+$ and (p)ppGpp^{0}strains remained constant around ≈1.0% until the transition to stationary phase, where it slowly increased for the WT but peaked up to 1.8% in the RelA$^+$ strain and up to 11% in the (p)ppGpp^0 strain (Figure 5.6C, fifth panel). Then, further in stationary phase, the TFE rates slowly converged to a value of 2.0%. Whatever the TFE reporter system was used, the WT TFE rates in stationary phase were higher than in steady-state growth, which probably reflects an evolutionary trade-off between the control of translational errors and optimal cell adaptation. The (p)ppGpp^0 cells exhibited a huge burst in TFEs in the transition to the stationary phase, which demonstrated that the (p)ppGpp synthetase RelA played a preponderant role in the control of TFEs during nutritional downshifts. However, as compared to WT cells, RelA$^+$ cells still exhibited a small peak in TFEs in the transition to the stationary phase. We concluded that the RelP and/or RelQ (p)ppGpp synthetases also contributed to reduce TFE occurrence in the transition to the stationary phase. Overall, we had assumed that (p)ppGpp maintained lower TFE rates in WT cells via its feedback regulation on GTP biosynthesis, we therefore expected TFE rates to be increased in rich media as the GTP level positively correlates with growth rate [Bittner et al., 2014]. Yet, we observed similar TFE rates in steady-state growth in poor and rich media. This observation challenged our assumption that GTP is by itself the TFE trigger factor.

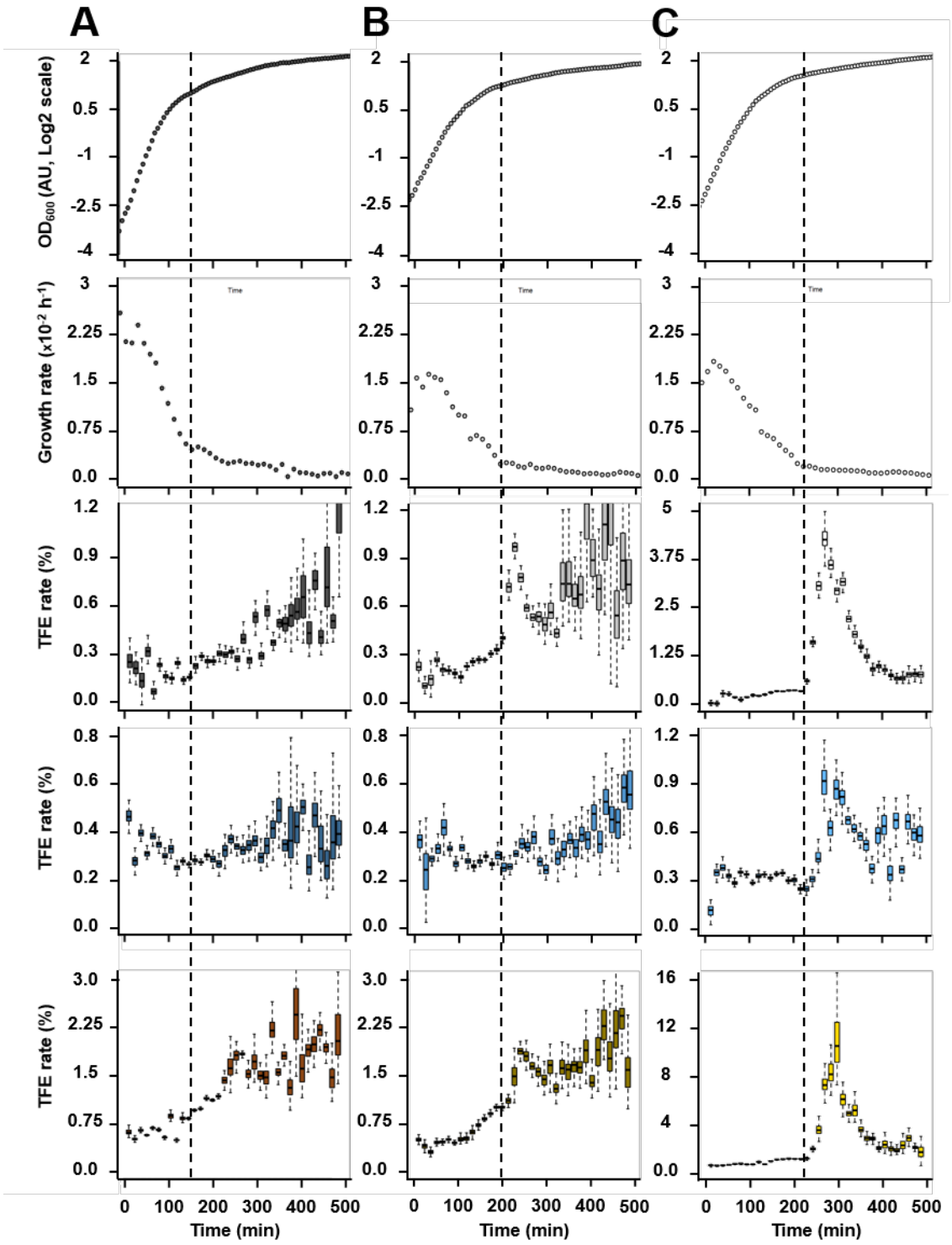

A
B
C
OD600 (AU, Log2 scale)
Growth rate (x10-2 h-1)
TFE rate (%)
TFE rate (%)
TFE rate (%)
Time (min)
Time (min)
Time (min)

Figure 5.6: Translational frameshift errors during the different growth phases in rich medium. A, B and C. The first panel represents the growth curve; the second one represents the instantaneous growth rate at each time point; and the third, the fourth and the fifth panels represents the TFE rate of the GFP^{Tyr}_{fs+1}, GFP^{Ile}_{fs-1} and GFP^{Leu}_{fs-1} reporter systems, respectively. The TFE rate was calculated using the second computation method and estimated by the bootstrap method (represented by boxplots). Note that different y-axis scale are used in the third to fifth panels. A. These five panels represent the growth, instantaneous growth rate and TFE rate of the WT strain grown in CH. B. These five panels represent the growth, instantaneous growth rate and TFE rate of the RelA$^+$ strain grown in CH. C. These five panels represent the growth, instantaneous growth rate and TFE rate of the (p)ppGpp^0 strain grown in CH.

5.5 GTP abundance is the trigger factor of the TFE occurrence in the exponential growth phase and in the transition to stationary phase.

Growth of (p)ppGpp^0 cells with down-regulated *guaB* expression in rich medium in the presence of guanosine (GUO, as a precursor of GTP synthesis; Figure 5.5A) allows to precisely control GTP abundance and growth rate [Bittner et al., 2014]. In order to verify whether GTP abundance triggers TFEs in *B. subtilis*, we monitored TFEs using the GFP^{Tyr}_{fs+1} reporter system in the RelA$^+$ $guaB^{inducible}$ and (p)ppGpp^0 $guaB^{inducible}$ strains grown in CH medium in the presence of different guanosine (GUO) concentrations. As shown on Figure 5.7A, the growth rate of the RelA$^+$ $guaB^{inducible}$ strain increased with increasing GUO concentration until reaching a value of 1.3 h^{-1} at concentrations equal or above 500 μM. On the contrary, the growth rate of the (p)ppGpp^0 $guaB^{inducible}$ strain increased up to 1.1 $^{-1}$ at 200 μM GUO and then decreased for higher concentrations (Figure 5.7A). For GUO concentrations of 50 μM or below, the TFE rates in steady-state growth remained constant at ≈0.2% in both RelA$^+$ $guaB^{inducible}$ and (p)ppGpp^0 $guaB^{inducible}$ strains (Figure 5.7B top panel). When higher GUO concentrations were used, the TFE rates in the RelA$^+$ genetic background remained constant, while higher TFE rates were detected in the (p)ppGpp^0 background with a maximal value of ≈0.8% at 1 mM GUO during steady-state growth (Figure 5.7B last panel). This indicated that in addition to the basal TFE rate, GTP abundance triggered the TFE increase in the exponential growth phase. In the late stationary growth phase, the TFE rates of both strains converged to a value of 1.0% whatever the GUO concentration used. When experiments were performed using GUO concentrations of 200 μM or below, the TFE rates slowly increased from exponential to stationary growth phases but no peak was detected. For GUO concentrations of 500 μM or above, the TFE rate peaked in the two strains in the transition to stationary phase. While the burst in TFE rates of the RelA$^+$ $guaB^{inducible}$ strain remained around ≈1.0% (which was similar to the peak observed in the RelA$^+$ strain grown in CH), its size increased up to a value of ≈8% at GUO 500 μM in the (p)ppGpp^0 $guaB^{inducible}$ strain (Figure 5.7B third panel). At GUO 1 mM, the burst size was reduced, which is most likely due to a GTP

excess that altered the overall cellular process of protein production, and which is consistent with the fact that (p)ppGpp^0 *guaB*^{inducible} cells died for GUO concentrations above 2 mM (not shown). Taken together our results demonstrated that GTP abundance triggered TFEs (in addition to the basal TFE level) in both the exponential growth phase and the transition to the stationary phase.

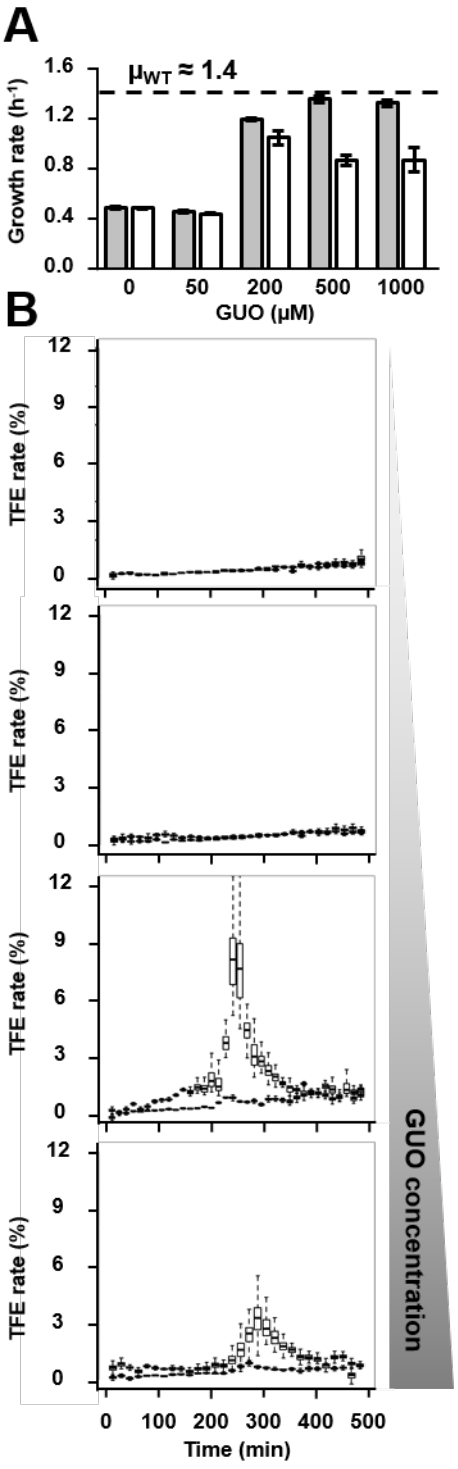


Figure 5.7: Effects of varying the GTP level on the TFE rate in rich medium (CH). The RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains grown in CH medium with different guanosine (GUO) concentrations: 50 μM, 200 μM, 500 μM and 1000 μM. A. Steady-state growth rate of the RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains grown in CH with different GUO concentrations (50, 200, 500 and 1000 μM). B. The panels represent the TFE rate as function of time for the RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains grown in CH medium with different GUO concentrations (50, 200, 500 and 1000 μM). The TFE rate was calculated using the second computation method and estimated by the bootstrap method (represented by boxplots).

5.6 Containing TFEs only via a regulation operating at the level of GTP biosynthesis in (p)ppGpp^0 cells results in sub-optimal growth.

Since the TFE rate increased together with the level of GTP in (p)ppGpp^0 cells, we wondered whether precisely regulating GTP biosynthesis (as the (p)ppGpp usually does by inhibiting Gmk and GuaB activities) would be sufficient to prevent the TFE burst in the transition to stationary phase. TFEs were monitored using the GFP^{Tyr}_{fs+1} reporter system in the RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains grown in CH medium. As shown on Figure 5.8A, the growth rates of the two strains increased with increasing IPTG concentration until reaching a plateau at 1.3 h^{-1} for IPTG 50 μM and above (similar to the WT growth rate in CH medium). During steady-state growth, the TFE rates remained constant at ≈0.2% in both RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains for the IPTG concentrations tested so far (Figure 5.8B). For IPTG concentrations of 5 μM or below, the growth rate was sub-optimal with respect to the WT and no TFE peak was detected but TFE rates slowly increased from exponential to stationary growth phases. At 15 μM IPTG, the TFE rate peaked for the two strains and up to 3.5% in the transition to the stationary phase of the (p)ppGpp^0 strain while the growth rates, and presumably GTP abundance, were still sub-optimal with respect to the WT. For IPTG concentrations above 15 μM, the TFE rate peak was of ≈1.0% in the RelA$^+$ *guaB*^{inducible} strain (which was similar to the peak observed in the RelA$^+$ strain grown in CH). In the (p)ppGpp^0 *guaB*^{inducible} strain, the burst in TFE rate increased with increasing IPTG concentration until reaching ≈5% at 200 μM IPTG (Figure 5.8B). It is worth to note that this value was similar to that observed in the (p)ppGpp^0 strain grown in CH, which indicated that the 200 μM IPTG mediated induction of *guaB* expression in CH most likely mimicked GTP biosynthesis of a (p)ppGpp^0 strain. In the late stationary growth phase, the TFE rates of both strains converged to a value of 1.0% whatever the IPTG concentration was being used. A strict control over GTP biosynthesis leading to sub-optimal growth rates (i.e. lower than that of the WT strain) did not prevent TFEs in the transition to the stationary phase as observed when using IPTG 15 μM with both the RelA$^+$ and (p)ppGpp^0 genetic backgrounds. As a consequence, the feedback regulation of (p)ppGpp on the activity of enzymes involved in GTP biosynthesis did not appear sufficient to prevent the TFE burst in the transition to

stationary phase (as observed during optimal adaptation of WT cells).

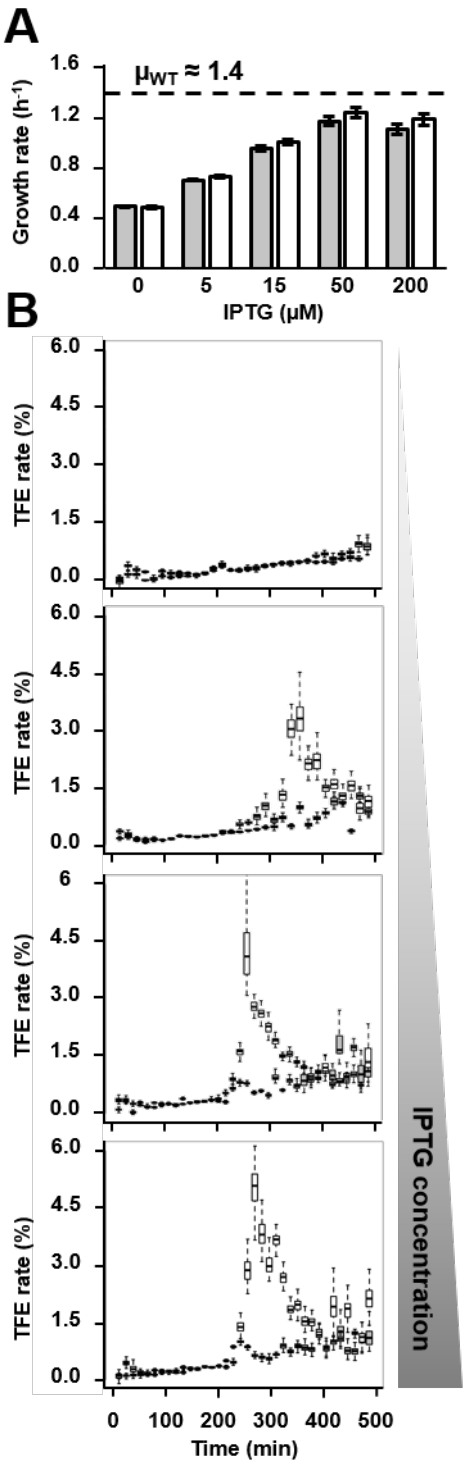


Figure 5.8: . Effects of regulating GTP biosynthesis on the growth rate and the TFE rate in rich medium (CH). The RelA$^+$ *guaB^{inducible}* and (p)ppGpp^0 *guaB^{inducible}* strains grown in CH medium with different IPTG concentrations: 5 μM, 15 μM, 50 μM and 200 μM. A. Steady-state growth rate of the RelA$^+$ *guaB^{inducible}* and (p)ppGpp^0 *guaB^{inducible}* strains grown in CH with different IPTG concentrations (5, 15, 50 and 200 μM). B. The panels represent the TFE rate as function of time for the RelA$^+$ *guaB^{inducible}* and (p)ppGpp^0 *guaB^{inducible}* strains grown in CH medium with different IPTG concentrations (5, 15, 50 and 200 μM). The TFE rate was calculated using the second computation method and estimated by the bootstrap method (represented by boxplots).

5.7 Active inhibition of translation initiation in (p)ppGpp^0 cells is sufficient to optimally prevent the burst in TFEs in the transition to the stationary phase.

GTP and (p)ppGpp respectively induces and inhibits translation initiation through competitive binding to IF2 [Milon et al., 2006] (Figure 5.5A). We wondered whether this competitive mode of action involving (p)ppGpp is sufficient to optimally prevent the TFE burst in the transition to the stationary phase. Our strategy to test this hypothesis was to make use of drugs known to inhibit translation initiation (i.e. linezolid, chloramphenicol and erythromycin [Wilson, 2014]) in order to functionally mimic the mode of action of (p)ppGpp on IF2. These drugs were injected on CH grown RelA$^+$ and (p)ppGpp^0 cells before the transition to stationary phase. Linezolid injection at concentrations of 1, 2 and 5 μM did not affect growth transition to the stationary phase (Figure 5.9A and 5.10A) but significantly reduced GFP production without stopping it (Figure 5.11). For both strains, the higher the linezolid concentrations were being used, the lower the TFE rate bursts were (Figure 5.9B, C, and D and Figure 5.10B, C, and D). When injecting linezolid concentrations of 2 μM or above on RelA$^+$ as well as on (p)ppGpp^0 cells, it turned out that the TFE bursts were completely abolished and that TFE rates smoothly increased to converge to about 1.0% in both strains (Figure 5.9E and 5.10E). It is remarkable that such behavior is identical to that observed during the adaptation of WT cells. Upon chloramphenicol and erythromycin injections, similar trends were observed for the RelA$^+$ and (p)ppGpp^0 strain grown in CH where the TFE rate was reduced to about 1% after drug injection and remained constant afterward (Figure 5.9F, G and H and 5.10F, G, H, I and J). Altogether, these results revealed that an active inhibition of the translation initiation was sufficient to prevent a burst in the TFE rate in both RelA$^+$ and (p)ppGpp^0 context. Hence, the inhibition of IF2 by (p)ppGpp during the stringent response when nutrients become limiting may account for the controlled transition in the TFE rate from exponential to stationary growth phases.

The next question to address was whether such a mechanism is robust enough to prevent TFE bursts induced by a rapid excess of GTP followed by nutrient depletion as can occur in a fluctuating environment. We therefore tested whether an active inhibition of translation initiation (with a 5 μM linezolid

injection) was still sufficient to prevent a TFE occurrence resulting from a rapid burst in GTP abundance (induced by a 200 μM GUO injection). After injection of 200 μM GUO, the TFE rate exhibited a sharper increase as compared to the control (DMSO injection) and reached a maximum value of $\approx$6% (Figure 5.9I), higher than that of the control of $\approx$4% (Figure 5.9B). Right after a simultaneous injection of 5 μM linezolid and 200 μM GUO, the TFE rate rapidly increased from $\approx$0.2% to $\approx$0.8% but did not peak (Figure 5.9J). Moreover, approximately 1h after the co-injection it significantly decreased. The GTP produced right after the injection of GUO may be responsible for the rapid TFE rate increase as compared to the control, but then the inhibition of translation initiation by linezolid prevented the burst in TFEs. Consequently, (p)ppGpp's inhibitory effect on IF2 is very likely to be responsible for rapidly preventing a high increase in the TFE rate when nutrients become limiting, even if the GTP level has not decreased yet.

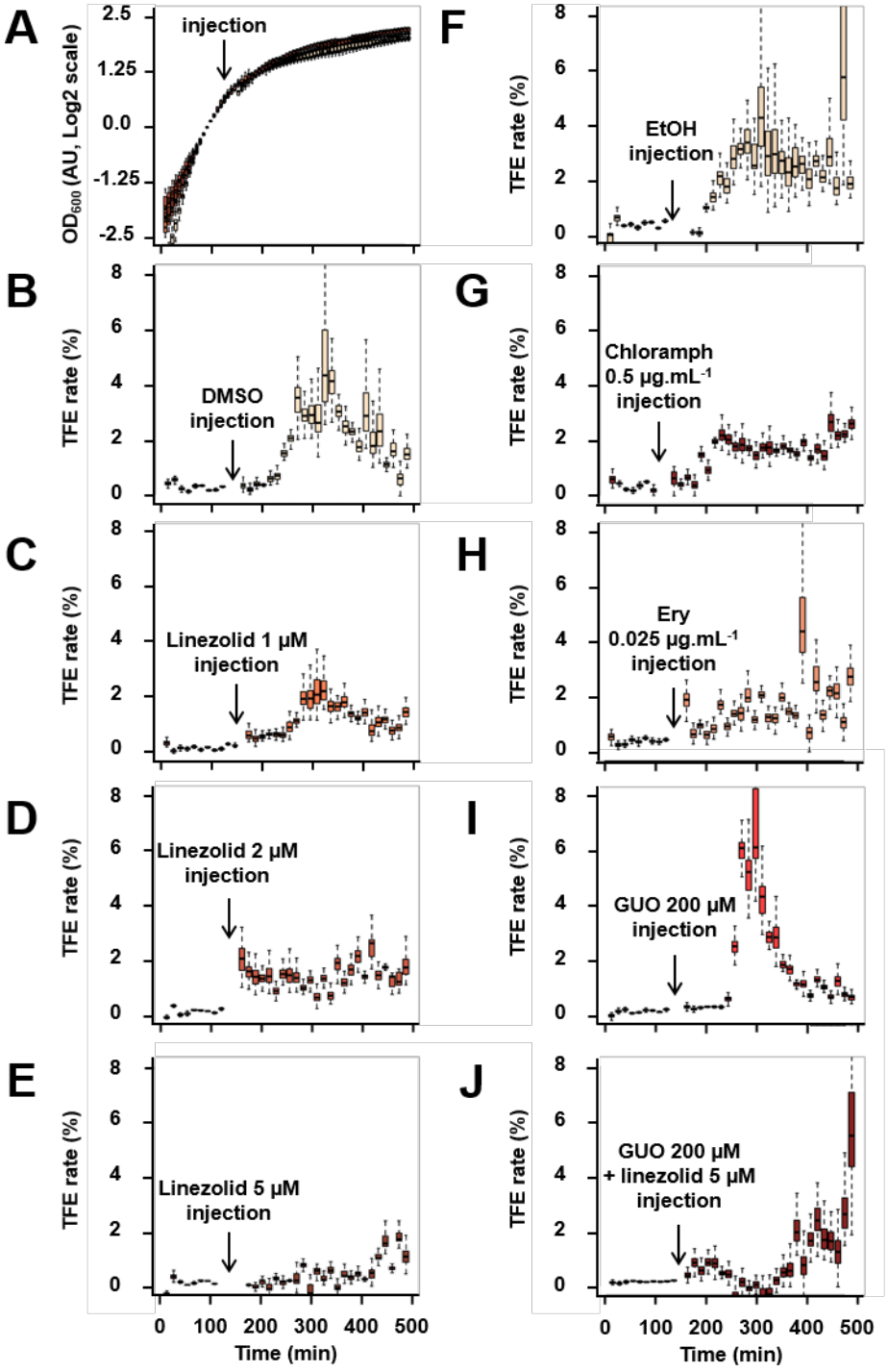

A
B
C
D
E
F
G
H
I
J
OD600 (AU, Log2 scale)
2.5
1.25
0.0
-1.25
-2.5
injection
TFE rate (%)
DMSO injection
Linezolid 1 µM injection
Linezolid 2 µM injection
Linezolid 5 µM injection
EtOH injection
Chloramph 0.5 µg.mL-1 injection
Ery 0.025 µg.mL-1 injection
GUO 200 µM injection
GUO 200 µM + linezolid 5 µM injection
0 100 200 300 400 500
Time (min)

Figure 5.9: Preventing the translational frameshift error peak appearance by inhibiting translation initiation. A. Growth curves of the (p)ppGpp^0 strain upon DMSO or linezolid injections at different final concentrations (1, 2 and 5 μM). B, C, D and E. TFE rate as function of time when the (p)ppGpp^0 and derivative strains were grown in CH upon either DMSO injection (B) or linezolid injection at 1 μM (C), 2 μM (D) and 5 μM (E) final concentrations. F, G and H. TFE rate as function of time when the (p)ppGpp^0 and derivative strains were grown in CH upon either ethanol injection (F); chloramphenicol injection at 0.5 μg.mL^{-1} (G) final concentration; and erythromycin injection at 0.25 μg.mL^{-1} (H) final concentrations. I. TFE rate as function of time when the (p)ppGpp^0 and derivative strains were grown in CH upon GUO injection at 200 μM final concentration. J. TFE rate as function of time when the (p)ppGpp^0 and derivative strains were grown in CH upon simultaneous injection of GUO at 200 μM and linezolid at 5 μM final concentrations. The TFE rate was calculated using the second computation method and estimated by the bootstrap method (represented by boxplots).

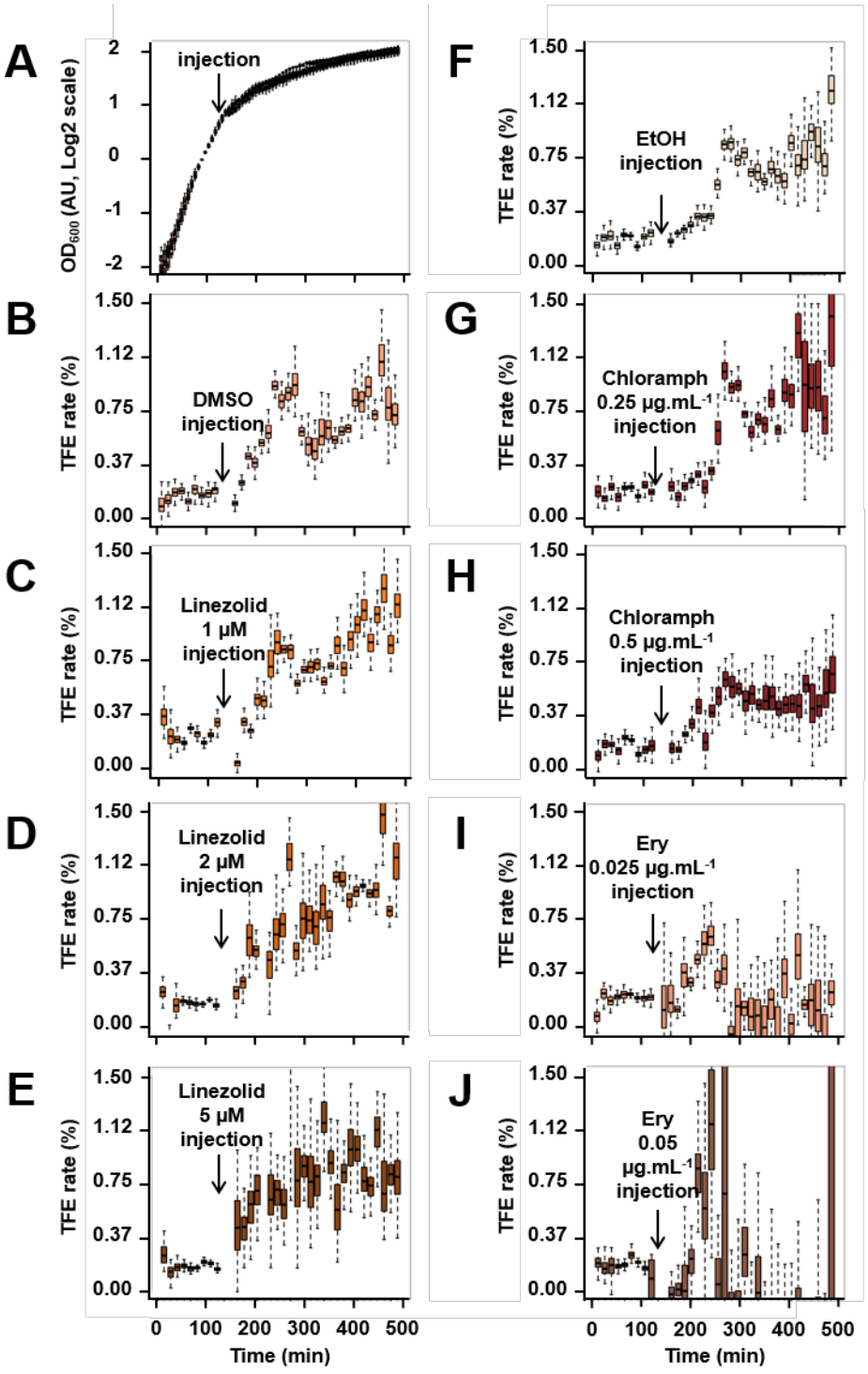

A
injection
OD600 (AU, Log2 scale)
B
DMSO injection
C
Linezolid 1 µM injection
D
Linezolid 2 µM injection
E
Linezolid 5 µM injection
F
EtOH injection
G
Chloramph 0.25 µg.mL-1 injection
H
Chloramph 0.5 µg.mL-1 injection
I
Ery 0.025 µg.mL-1 injection
J
Ery 0.05 µg.mL-1 injection
TFE rate (%)
Time (min)

Figure 5.10: Inhibiting translation initiation by injecting drugs during the growth of RelA$^+$ and derivative strains in CH. A. Growth curves of the RelA$^+$ strain upon DMSO or linezolid injections at different final concentrations (1, 2 and 5 μM). B, C, D and E. TFE rate as function of time when the RelA$^+$ and derivative strains were grown in CH upon either DMSO injection (B) or linezolid injection at 1 μM (C), 2 μM (D) and 5 μM (E) final concentrations. F, G, H, I and J. TFE rate as function of time when the RelA$^+$ and derivative strains were grown in CH upon either ethanol injection (F); chloramphenicol injection at 0.25 μg.mL^{-1} (G) and 0.5 μg.mL^{-1} (H) final concentration; and erythromycin injection at 0.25 μg.mL^{-1} (I) and 0.5 μg.mL^{-1} (J) final concentrations. The TFE rate was calculated using the second computation method and estimated by the bootstrap method (represented by boxplots).

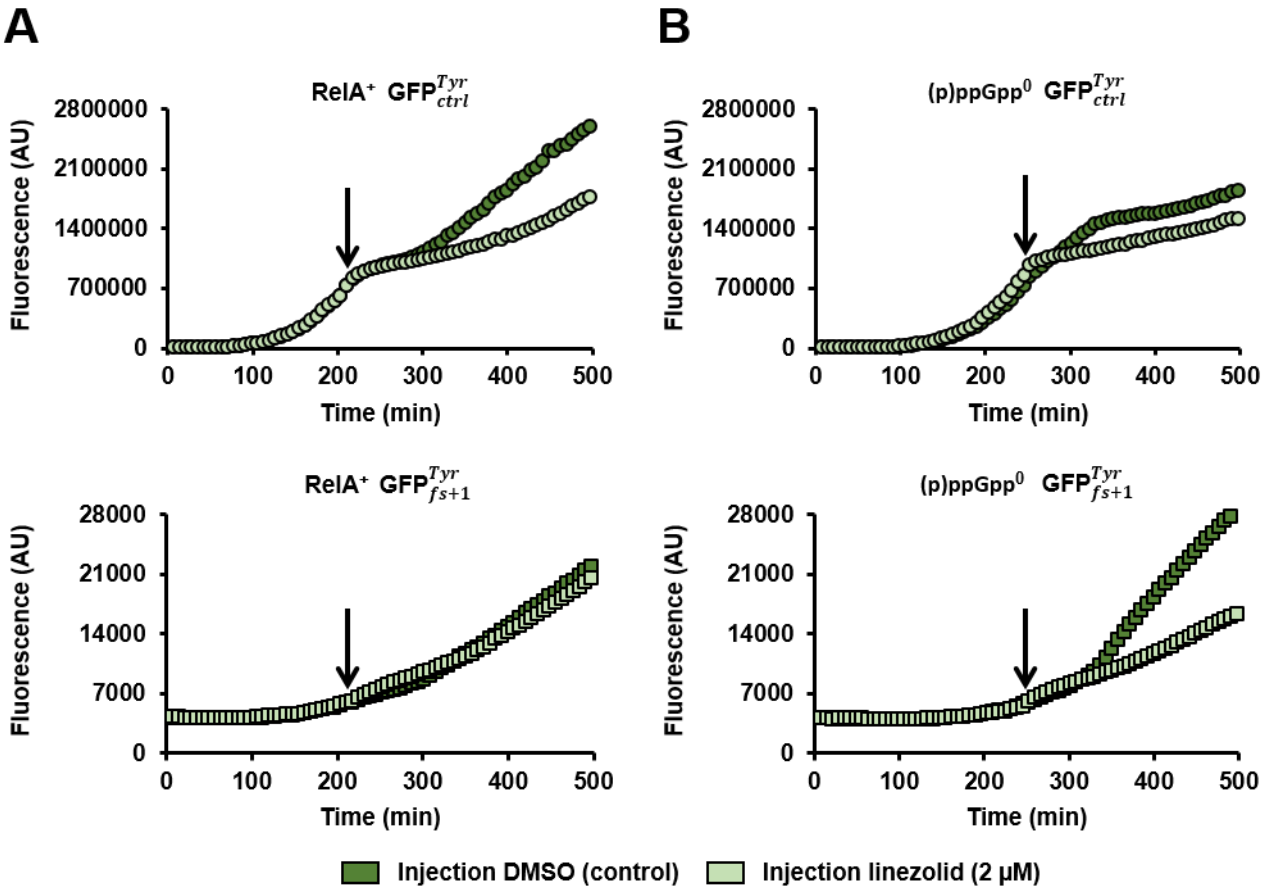


Figure 5.11: Effects of the translation initiation inhibition by linezolid injection on the production of GFP. A and B. The graph at the top represents the fluorescence of GFP^{Tyr}_{ctrl} as a function of time and the bottom graph represents the fluorescence of GFP^{Tyr}_{fs+1} as a function of time. The down arrow indicates when the DMSO (control) or linezolid (2 μM final concentration) were injected. A. Effects of the injection of linezolid on the GFP level of the RelA$^+$ strain grown in CH medium. B. Effects of the injection of linezolid on the GFP level of the (p)ppGpp^0 strain grown in CH medium.

Chapter 6

Effects of GTP and (p)ppGpp on the transcription process and the ribosome synthesis

6.1 How does the GTP/(p)ppGpp feedback loops affect transcription according to the nature of the TSS

6.1.1 The TSS influences the mRNA abundance evolution with growth rate

To decipher how the TSS influences the mRNA abundance evolution with growth ate, we first built reporter systems comprising a constitutive promoter (P_{fbaA}^{+1GG}) followed by the native translation initiation region (TIR) of the *fbaA* gene fused to a *gfpmut3* gene (figure 8.9). In the case of *fbaA*, the native TIR starts with a TSS consisting in two guanines (referred to as +1GG), and we mutated this region by replacing the two guanines by two adenines (referred to as +1AA), which led to the two genetic constructs P_{fbaA}^{+1GG} and P_{fbaA}^{+1AA}, respectively. We implemented P_{fbaA}^{+1GG} and P_{fbaA}^{+1AA} in the WT strain (Table 8.9, strains CLB004 and CLB001) and grew cells in five different media leading to five distinct growth rates: M9 with isoleucine, methionine, leucine and valine (M9IMLV) complemented with either 0.8 % (w/v %) pyruvate (M9IMLVPyr; μ=0.28±0.03h^{-1}) or 0.5 % malate and 0.3 % glucose (M9IMLVMalGlc; μ=0.92±0.04 h^{-1}), M9 with 0.5 % succinate and 0.5 % glutamate (M9SE; μ=0.51±0.02 h^{-1}), CH (μ=1.55±0.03h^{-1}) or CH with 0.5 % glucose (CHG; μ=2.29±0.20 h^{-1}). We measured the Optical Density (OD) at 600 nm (OD_{600}) as well as the fluorescence level which allowed us to deduce the GFP protein abundance $[GFP]$.

We can deduce the mRNA concentration from the GFP protein abundance $[GFP]$ through the following relationship established for steady-state growth by Borkowski *et al.* (2016) and described in section 2.2.1.3:

$$[GFP] = \frac{[mRNA]}{\mu} \times \frac{K_1[R_{free}]}{K_2 + [R_{free}]} \Leftrightarrow [GFP] \times \mu = [mRNA] \times \frac{K_1[R_{free}]}{K_2 + [R_{free}]}$$

The $\frac{K_1[R_{free}]}{K_2+[R_{free}]}$ ratio is assumed to be the same for both P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* (i.e. translation of each transcript is supposed not to be altered by the two-base replacement as a function of the growth rate), therefore the GFP productivity $P = [GFP] \times \mu$ is a good proxy to evaluate the GFP to be produced to compensate for the dilution rate (i.e. growth rate). Moreover, the ratio of GFP abundances of P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp*, i.e. $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$, is actually identical to the ratio of the mRNA concentrations of P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* $\frac{[mRNA]_{+1GG}}{[mRNA]_{+1AA}}$. Indeed, in a given growth medium, the growth rates (μ) and the $Rfree$ concentrations are supposed to be equal in both P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* strains such that:

$$\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}} = \frac{\frac{[mRNA]_{+1GG}}{\mu}\frac{K_1[R_{free}]}{K_2+[R_{free}]}}{\frac{[mRNA]_{+1AA}}{\mu}\frac{K_1[R_{free}]}{K_2+[R_{free}]}} \approx \frac{[mRNA]_{+1GG}}{[mRNA]_{+1AA}}$$

Hence, the productivity of P_{fbaA}^{+1GG}*gfp* increased more importantly with growth rate than did the productivity of P_{fbaA}^{+1AA}*gfp* during the steady-state phase of growth (figure 6.1 A). Consistently, the GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}} \approx \frac{[mRNA]_{+1GG}}{[mRNA]_{+1AA}}$ linearly increased with the growth rate (figure 6.1 B). This could be expected since the composition of the TSS has been shown for certain genes to change their level of transcription upon amino acid starvation and in vitro as a function of the variations in the GTP/ATP abundance ratio [Krásnỳ et al., 2008].

6.1.2 In the absence of (p)ppGpp the mRNA abundance is higher for genes possessing guanines as TSS

We next wondered whether the abundance of P_{fbaA}^{+1GG}*gfp* mRNAs would be higher than the abundance of P_{fbaA}^{+1AA}*gfp* mRNAs in a (p)ppGpp^0 strain across different growth conditions (as the GTP synthesis is not inhibited by (p)ppGpp in a (p)ppGpp^0 strain). We implemented the P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* constructs in the RelA$^+$ and (p)ppGpp^0 strains (table 8.9, strains CLB305, CLB307, CLB313 and CLB315). For strains grown in poor media (M9M and M9G), the growth rate was lower for the (p)ppGpp^0 strain (≈0.28 h$^-$1 in M9M and ≈0.35 h$^-$1 in M9G) than for the RelA$^+$ strain (≈0.62 h$^-$1 in M9M and ≈0.68 h$^-$1 in M9G) as previously observed (figure 6.2 A). In rich medium, the growth rate was slightly higher for the (p)ppGpp^0 strain than for the RelA$^+$ strain. In all growth media, the GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$was higher for the (p)ppGpp^0 strain (≈1.24 in M9M, ≈1.17 in M9G and ≈1.49 in CH) than for the RelA$^+$ strain (≈1.08 in M9M, ≈1.04 in M9G and ≈1.38 in CH) (figure 6.2 B). Also, the $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ ratio was higher in rich medium than in poor medium for both strains, which is consistent with the fact that the $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ ratio increased with growth rate in a WT strain (figure 6.1). In conclusion, these results suggest that the higher the GTP is, the more the genes possessing a guanine as TSS are favored for initiating transcription as compared to genes which have an adenine as TSS. Therefore, this may indicate that the GTP/(p)ppGpp feedback loops control resource allocation through a precise induction of the transcription of specific genes (i.e. genes with TSS enriched in guanine).

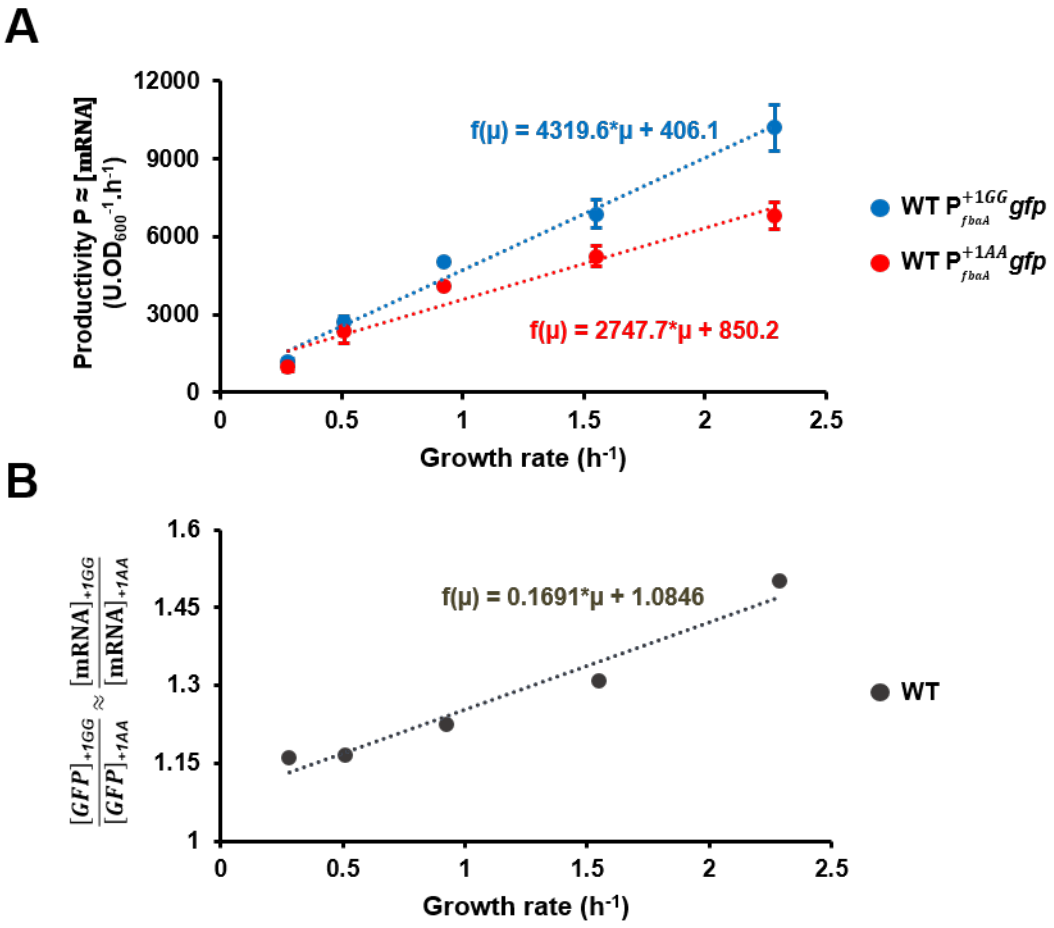


Figure 6.1: Effects of the TSS on GFP productivity and GFP abundance ratio for WT strains grown in different media A and B. The WT $P_{fbaA}^{+1GG}gfp$ and WT $P_{fbaA}^{+1AA}gfp$ strains were grown in five different media: M9IMLVPyr, M9IMLVMalGlc, M9SE, CH and CHG (see section 8.4.1). A. GFP productivity of theWT $P_{fbaA}^{+1GG}gfp$ and WT $P_{fbaA}^{+1AA}gfp$ strains as function of the growth rate. The productivity is expressed as $U.OD_{600}^{-1}.h^{-1}$ since the fluorescence is expressed in fluorescence arbitrary unit (U), the OD_{600} as absorbance unit (referred to as OD_{600}) and the growth rate is expressed in hour^{-1} (h^{-1}). The error bars correspond to the standard deviation of the GFP productivity calculated using multiple replicates. The blue and red dashed lines correspond to the fitting of the linear dependence of the GFP productivity of the WT $P_{fbaA}^{+1GG}gfp$ and WT $P_{fbaA}^{+1AA}gfp$ strains with the growth rate, respectively. The equations next to them correspond to the linear function of the fitting where μ designates the growth rate. B. GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ of the WT strain as function of the growth rate. The dashed line corresponds to the fitting of the linear dependence of the WT GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ with the growth rate; and the equation next to it corresponds to the linear function of the fitting where μ designates the growth rate.

6.1.3 Impact of the control of intracellular GTP level on gene expression according to the TSS composition in poor medium

From the results we just presented, we wondered whether controlling the intracellular GTP level would restore a similar GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ (and so mRNA abundance ratio $\frac{[mRNA]_{+1GG}}{[mRNA]_{+1AA}}$

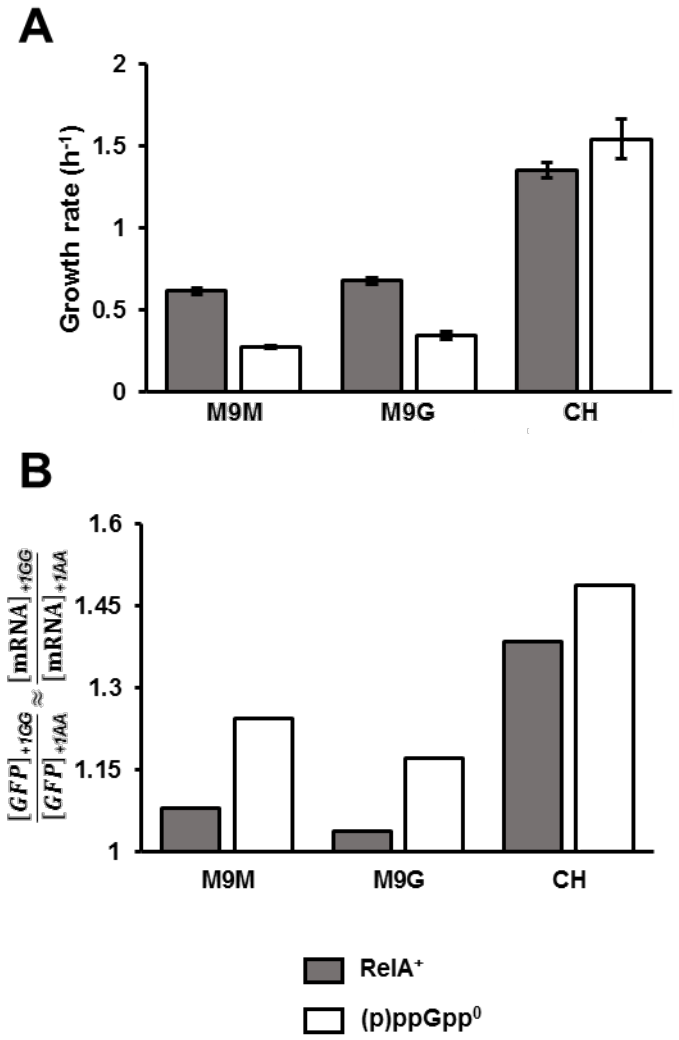


Figure 6.2: Impact of the absence of (p)ppGpp on the GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ for cells grown in different media. A. Growth rate of the RelA$^+$ and (p)ppGpp^0 strains grown in M9M, M9G and CH (see section 8.4.1). The error bars correspond to the standard deviation of the growth rate calculated using multiple replicates. B. GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ of the RelA$^+$ and (p)ppGpp^0 strains grown in M9M, M9G and CH.

) between the (p)ppGpp^0 and the RelA$^+$ strains. To answer this question, we implemented the P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* constructs in the RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains (Table strains, strains CLB309, CLB311, CLB317 and CLB319). To allow correct growth, the LB pre-cultures contained 100 μM IPTG which led to a residual IPTG concentration of 0.25 μM in the cultures. For IPTG concentrations equal or inferior to 15.25 μM, the growth rate as well as the GFP productivity of P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* increased with IPTG concentrations for both the RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains (figure 6.3 A, B and figure 6.4 A, B). The values of the GFP productivity of P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* reached a maximum value of approximately 1000 U.OD$_{600}$.h^{-1} and 900 U.OD$_{600}$.h^{-1} in M9G and M9M media, respectively. When the IPTG concentration was above 15 μM, the growth rate and the GFP productivity of P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* remained identical

for the RelA$^+$ *guaB*^{inducible} strain while they decreased for the (p)ppGpp^0 *guaB*^{inducible} strain. We observed for the (p)ppGpp^0 *guaB*^{inducible} strain that similar low growth rates were obtained at IPTG concentrations of 0.25, 50.25 and 200.25 μM, however the GFP productivity was different, which may indicate that a too high production of GTP in poor medium globally disturbed protein synthesis.
To see whether there was a difference in the transcription rate depending on the TSS composition, we plotted as a function of the IPTG concentration the GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ for the RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains; but we did not observe significant differences between the two strains (figure 6.3 C and figure 6.4 C). The same tendency was observed for both strains when grown in M9G: at low IPTG concentration (inferior to 15 μM) the gfp gene was more transcribed for P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* (ratio of ≈0.75-0.85); while at higher IPTG concentrations (15.25 μM and above), the transcription level was similar for both constructs (ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ of ≈1). When grown in M9M, a similar trend was observed but at IPTG concentration of 2.75 μM, the gfp genes of the P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* constructs were already transcribed at a similar level (ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ of ≈0.95 up to 1.1); while the gfp gene of the P_{fbaA}^{+1AA}*gfp* construct was more transcribed when there was no IPTG in the medium (ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ of ≈0.8).

6.1.4 Impact of variations in intracellular GTP on gene expression as a function of the TSS composition in rich medium

Given that the high rate of GTP synthesis seemed to affect protein synthesis in poor media, we tested whether we could see a difference in the transcription rate of constructs possessing either guanines or adenines as TSS when strains were grown in rich medium. We grew the RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains in CH medium with different IPTG concentrations. As previously observed (see part 5), the growth rate increased similarly for both strains when the IPTG concentration was increased (figure 6.5 A). In both strains, the evolution of the GFP productivity as a function of the growth rate was similar for both P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* (figure 6.5 B). The GFP productivity of P_{fbaA}^{+1GG}*gfp* increased linearly and importantly (from ≈1000 to ≈1500 U.OD$_{600}$.h^{-1}) with growth rate, while it just slightly increased for P_{fbaA}^{+1AA}*gfp* (from ≈950 to ≈1150 U.OD$_{600}$.h^{-1}). We compared these results with the observations made for the WT strain possessing either the P_{fbaA}^{+1GG}*gfp* or P_{fbaA}^{+1AA}*gfp* genetic constructs when grown in different growth media (Figure 6.1 A): we observed in both cases that the GFP productivity increased more rapidly with growth rate for P_{fbaA}^{+1GG}*gfp* than for P_{fbaA}^{+1AA}*gfp*. When comparing the GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ for the RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains, we found that it was similar for the different IPTG concentrations (figure 6.5 C top panel). Nevertheless, we clearly observed that without addition of IPTG, the transcription level of P_{fbaA}^{+1GG}*gfp* was slightly higher than that of P_{fbaA}^{+1AA}*gfp*. The opposite occurred when IPTG was added to the growth medium leading to almost twice more mRNA abundance for P_{fbaA}^{+1GG}*gfp* at the maximal IPTG concentration (200.25 μM). These data agreed with the evolution observed for the GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ when strains were grown in different growth media (figure 6.1): the ratio increased linearly with growth rate even if the slopes are different (figure 6.5 C bottom panel). Thus,

it is very likely that when the GTP level is increased, the genes which possess guanines as TSS are preferentially transcribed than the genes which have adenines as TSS; and this difference is enhanced by the increase of the GTP level.

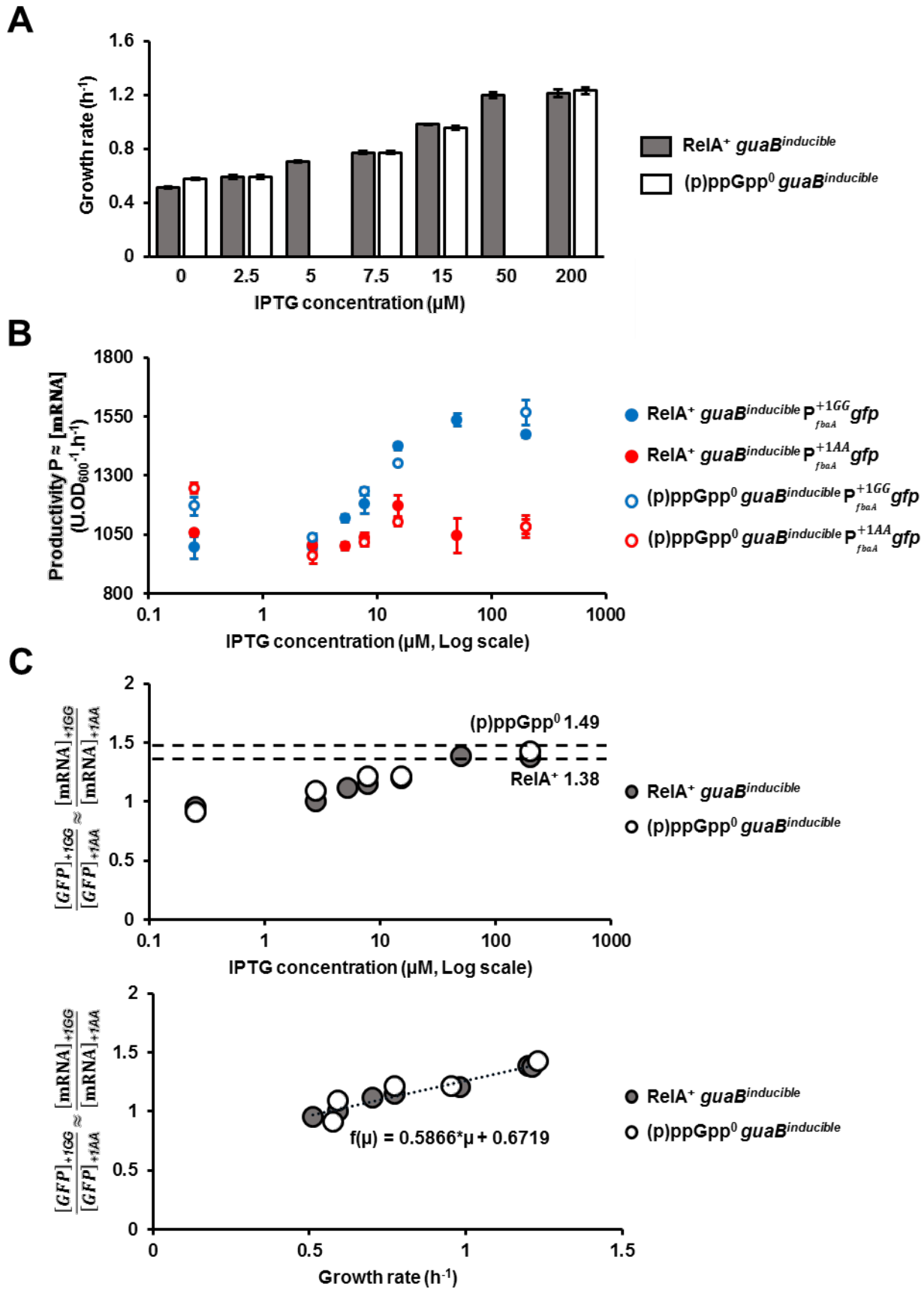

A
Growth rate (h-1)
1.6
1.2
0.8
0.4
0
0
2.5
5
7.5
15
50
200
IPTG concentration (µM)
RelA+ guaBinducible
(p)ppGpp0 guaBinducible
B
Productivity P ≈ [mRNA] (U.OD600-1.h-1)
1800
1550
1300
1050
800
0.1
1
10
100
1000
IPTG concentration (µM, Log scale)
RelA+ guaBinducible P+1GGfbaA gfp
RelA+ guaBinducible P+1AAfbaA gfp
(p)ppGpp0 guaBinducible P+1GGfbaA gfp
(p)ppGpp0 guaBinducible P+1AAfbaA gfp
C
[GFP]+1GG / [GFP]+1AA ≈ [mRNA]+1GG / [mRNA]+1AA
2
1.5
1
0.5
0
(p)ppGpp0 1.49
RelA+ 1.38
0.1
1
10
100
1000
IPTG concentration (µM, Log scale)
RelA+ guaBinducible
(p)ppGpp0 guaBinducible
[GFP]+1GG / [GFP]+1AA ≈ [mRNA]+1GG / [mRNA]+1AA
2
1.5
1
0.5
0
f(µ) = 0.5866*µ + 0.6719
0
0.5
1
1.5
Growth rate (h-1)
RelA+ guaBinducible
(p)ppGpp0 guaBinducible

Figure 6.5: Impact of the control of intracellular GTP level on gene expression according to the TSS composition for strains grown in CH with IPTG added. A. Mean of the growth rate of the RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains grown in CH with different concentrations of IPTG. The error bars correspond to the standard deviation of the growth rate calculated using multiple replicates. B. GFP productivity of the RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains carrying the P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* constructs as function of the IPTG concentration. The error bars correspond to the standard deviation of the GFP productivity calculated using multiple replicates. C. (top) GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ of the RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains as function of the IPTG concentration. The black dashed lines correspond to the value of this ratio for the RelA$^+$ and (p)ppGpp^0 strains grown in CH. (bottom) GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ of the RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains as function of the growth rate. The dashed line corresponds to the fitting of the linear dependence of the RelA$^+$ *guaB*^{inducible} GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ with the growth rate; and the equation next to it corresponds to the linear function of the fitting where $\mu*$) designates the growth rate.

We next tested whether we would observe similar results when adding guanosine (GUO) to the growth medium since the GTP level is higher in the (p)ppGpp^0 strain than in the WT strain upon GUO addition [Kriel et al., 2012]. The RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains were grown in CH medium where GUO was added at different concentrations. For the RelA$^+$ *guaB*^{inducible} strain, upon addition of 50 μM GUO and more, the maximal growth rate was reached as well as the maximal GFP productivity for both P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* (figure 6.6 A and B). This differs from observations in chapter 1 when we grew the RelA$^+$ *guaB*^{inducible} strain in CH upon different concentrations of GUO added. This difference may be due to an experimental problem where higher levels of IPTG were added to the pre-culture which already contributed to GTP synthesis. However, this does not affect the following obersvations and conclusions. Concerning the (p)ppGpp^0 *guaB*^{inducible} strain, the growth rate increased with GUO but it reached a maximal value for GUO concentrations equal or above 200 μM, which is inferior to the one found for the (p)ppGpp^0 strain in CH. Concerning GFP productivity, it decreased while the GUO concentration was increased, and this decrease was more important for P_{fbaA}^{+1AA}*gfp* than for P_{fbaA}^{+1GG}*gfp*. Nevertheless, when we plotted the GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$, we observed that for the RelA$^+$ *guaB*^{inducible} strain, the ratio increased and then reached a plateau ($\approx$1.8) at GUO concentrations above 200 μM, while for the (p)ppGpp^0 *guaB*^{inducible} strain it linearly increased with the GUO concentration until it reached a value of $\approx$3.5 (figure 6.6 C). Hence, these results confirm that the GTP level stimulated the transcription of genes which possessed a guanine as TSS at the expense of the transcription of genes which possessed an adenine as TSS. This might explain why at GUO concentrations above 200 μM the growth rate was affected: certain proteins required for growth might be overexpressed at the expense of essential proteins which overall resulted in a non-optimal resource allocation between the different cellular processes.

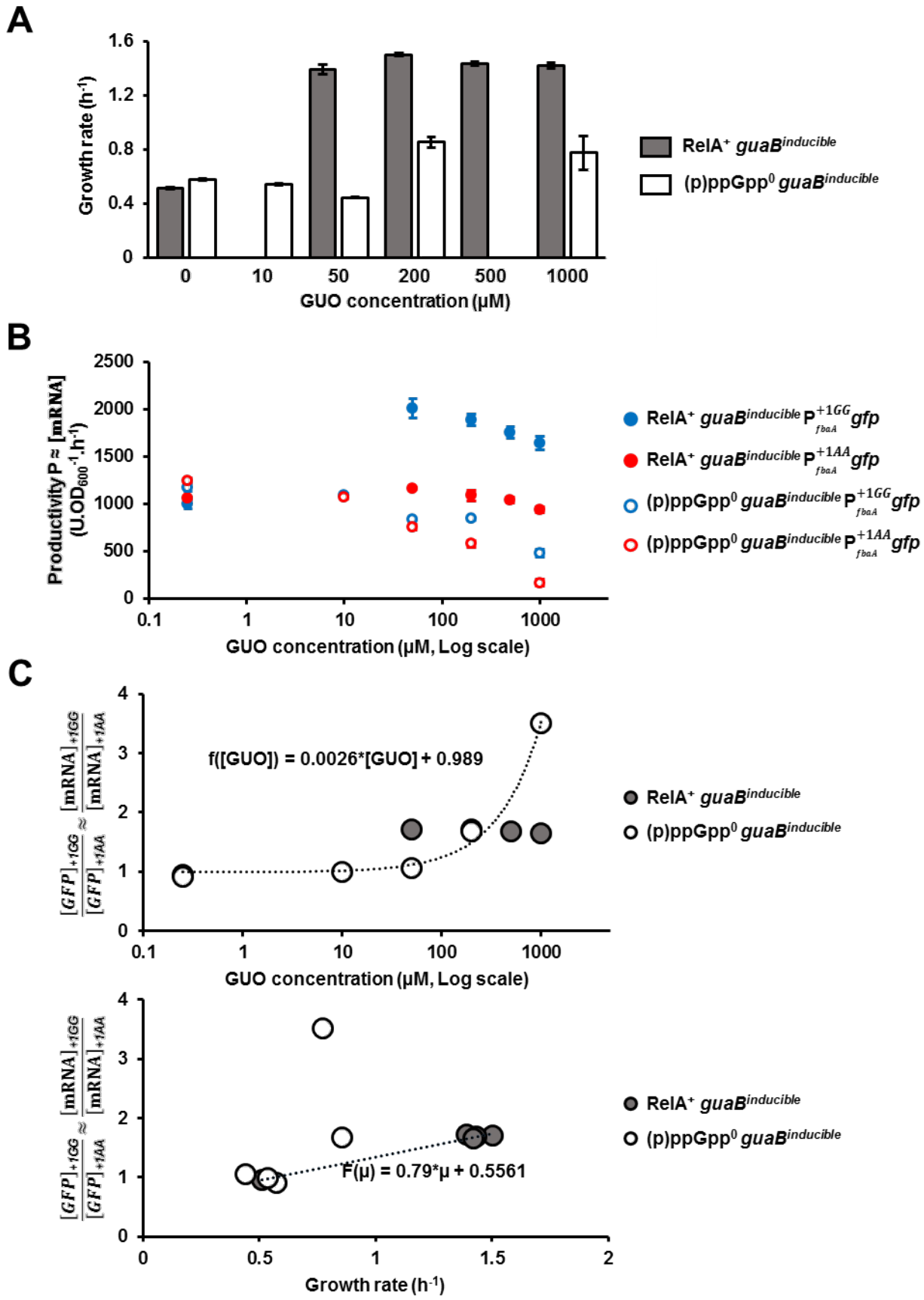

A
Growth rate (h-1)
1.6
1.2
0.8
0.4
0
0
10
50
200
500
1000
GUO concentration (µM)
RelA+ guaBinducible
(p)ppGpp0 guaBinducible
B
Productivity P ≈ [mRNA] (U.OD600-1.h-1)
2500
2000
1500
1000
500
0
0.1
1
10
100
1000
GUO concentration (µM, Log scale)
RelA+ guaBinducible P+1GGfbaA gfp
RelA+ guaBinducible P+1AAfbaA gfp
(p)ppGpp0 guaBinducible P+1GGfbaA gfp
(p)ppGpp0 guaBinducible P+1AAfbaA gfp
C
[GFP]+1GG / [GFP]+1AA ≈ [mRNA]+1GG / [mRNA]+1AA
f([GUO]) = 0.0026*[GUO] + 0.989
4
3
2
1
0
0.1
1
10
100
1000
GUO concentration (µM, Log scale)
RelA+ guaBinducible
(p)ppGpp0 guaBinducible
[GFP]+1GG / [GFP]+1AA ≈ [mRNA]+1GG / [mRNA]+1AA
F(µ) = 0.79*µ + 0.5561
4
3
2
1
0
0
0.5
1
1.5
2
Growth rate (h-1)
RelA+ guaBinducible
(p)ppGpp0 guaBinducible

Figure 6.6: Impact of the control of intracellular GTP level on gene expression according to the TSS composition for strains grown in CH with GUO added. A. Mean of the growth rate of the RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains grown in CH with different concentrations of IPTG. The error bars correspond to the standard deviation of the growth rate calculated using multiple replicates. B. GFP productivity of the RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains carrying the P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* constructs as function of the GUO concentration. The error bars correspond to the standard deviation of the GFP productivity calculated using multiple replicates. C. (top) GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ of the RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains as function of the IPTG concentration. The black dashed line corresponds to the fitting of the linear dependence of the (p)ppGpp^0 *guaB*^{inducible} GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ with the GUO concentration (remark: the fit curve appears as exponential since the GUO concentration is represented in Log scale); and the equation next to it corresponds to the linear function of the fitting where [GUO] designates the GUO concentration. (bottom) GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ of the RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains as function of the growth rate. The dashed line corresponds to the fitting of the linear dependence of the RelA$^+$ *guaB*^{inducible} GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ with the growth rate; and the equation next to it corresponds to the linear function of the fitting where μ designates the growth rate.

6.2 Impact of the deregulation of the GTP/(p)ppGpp feedback loops on ribosome synthesis

6.2.1 *rrn* expression in different growth media in the presence or absence of (p)ppGpp

As GTP/ATP variations seemed to affect the overall resource allocation, we wondered how *rrn* expression is affected since RNA synthesis is the rate-limiting step in ribosome synthesis [Henkin and Yanofsky, 2002][Paul et al., 2004b]. We measured the expression level of two *rrn* promoters (P_{rrnJ} and P_{rrnO}) by fusing the PrrnJ and PrrnO to the gfp gene, leading to the genetic constructs P_{rrnJ}*sfGFP* and P_{rrnO}*sfGFP*, respectively (figure 8.7). We inserted the P_{rrnJ}*sfGFP* and P_{rrnO}*sfGFP* into the RelA$^+$ and (p)ppGpp^0 strains and measured the level of fluorescence of these strains cultivated in different media. We observed during steady-state growth in rich medium (CH) that in both strains the growth rate was similar while the GFP productivity was slightly higher in the (p)ppGpp^0 strain than in the RelA$^+$ strain for both P_{rrnJ}*sfGFP* and P_{rrnO}*sfGFP* constructs (figure 6.7). When strains were grown in poor media (M9M and M9G), the steady-state growth rate as well as the GFP productivity of both P_{rrnJ}*sfGFP* and P_{rrnO}*sfGFP* were lower for the (p)ppGpp^0 strain than for the RelA$^+$ strain. This is unexpected since in the (p)ppGpp^0 strain the GTP level is supposed to be higher which would lead to higher expression level of *rrn* as observed by Krasny and Gourse (2004).

6.2.2 In the absence of (p)ppGpp, an excess of GTP disturbs *rrn* expression

We tested whether modulating the GTP level in the absence of (p)ppGpp could prevent a disequilibrium in *rrn* expression in poor media. We implemented P_{rrnJ}*sfGFP* and P_{rrnO}*sfGFP* into the RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains and then cultivated cells in either M9G or M9M media with different concentrations of IPTG. As previously observed in part 5, the growth rate increased for both strains for IPTG concentration equal or inferior to 15 μM; but for concentrations above 15 μM the growth rate remained the same in the RelA$^+$ *guaB*^{inducible} strain while it decreased in the (p)ppGpp^0 *guaB*^{inducible} strain in both M9G and M9M (figures 6.8 A and 6.9 A). Concerning the GFP productivity of P_{rrnJ}*sfGFP* and P_{rrnO}*sfGFP* in M9G, for the RelA$^+$ *guaB*^{inducible} strain it increased for IPTG concentrations up to 15 μM and for higher IPTG concentrations it remained approximately the same ($\approx$35 000 U.OD$_{600}$.h^{-1} for P_{rrnJ}*sfGFP* and $\approx$75 000 for U.OD$_{600}$.h^{-1} P_{rrnO}*sfGFP*) (figure 6.8 B and C). However, for the (p)ppGpp^0 *guaB*^{inducible} strain, the GFP productivity increased with IPTG concentrations (until 15 μM) up to a value of $\approx$35 000 U.OD$_{600}$.h^{-1} for P_{rrnJ}*sfGFP* and $\approx$75 000 U.OD$_{600}$.h^{-1} for P_{rrnO}*sfGFP*; and then it importantly decreased to reach a value of $\approx$5 600 U.OD$_{600}$.h^{-1} for P_{rrnJ}*sfGFP* and $\approx$3 900 U.OD$_{600}$.h^{-1} for P_{rrnO}*sfGFP* at IPTG 200 μM. Similar tendencies were observed for the GFP productivity of P_{rrnJ}*sfGFP* and P_{rrnO}*sfGFP* in M9M, but the highest value was reached at IPTG 5 μM where it was equal to $\approx$28 000 U.OD$_{600}$.h^{-1} for P_{rrnJ}*sfGFP* and $\approx$60 000 U.OD$_{600}$.h^{-1} for P_{rrnO}*sfGFP* in both RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains (figure 6.9 B and C). For IPTG concentrations equal or above 15 μM, the GFP productivity remained approximately the same for the RelA$^+$ *guaB*^{inducible} strain while it decreased for the (p)ppGpp^0 *guaB*^{inducible} strain to reach a value of $\approx$8 900 U.OD$_{600}$.h^{-1} for P_{rrnJ}*sfGFP* and $\approx$9 800 U.OD$_{600}$.h^{-1} for P_{rrnO}*sfGFP* at IPTG 200 μM.

Furthermore, we observed for similar growth rates obtained with different IPTG concentrations that the GFP productivity of P_{rrnJ}*sfGFP* and P_{rrnO}*sfGFP* was actually different. For instance, in M9G, the growth rate obtained at IPTG 2.5, 5 and 30 μM were in the range 0.35-0.45 h^{-1} but their respective GFP productivity were of $\approx$12 000, $\approx$12 000 and $\approx$18 000 U.OD$_{600}$.h^{-1} for P_{rrnJ}*sfGFP*and $\approx$49 000, $\approx$45 000 and $\approx$25 000 U.OD$_{600}$.h^{-1} for P_{rrnO}*sfGFP*. Similar trends were observed in M9M: the growth rate obtained at IPTG 5, 30 and 50 μM were in the range 0.24-0.28 h^{-1} but their respective GFP productivity were of $\approx$5 500, $\approx$13 000 and $\approx$11 500 U.OD$_{600}$.h^{-1} for P_{rrnJ}*sfGFP* and $\approx$23 000, $\approx$18 500 and $\approx$13 000 U.OD$_{600}$.h^{-1} for P_{rrnO}*sfGFP*.

These results suggest that in the absence of (p)ppGpp, *rrn* expression tend to be deregulated. It must be noted that since perturbing the GTP/(p)ppGpp feedback loops affects the transcription and translation processes, it is very likely that the GFP level does not completely reflect the expression level of *rrnJ* and *rrnO* (i.e. resources are expected to be mainly directed toward ribosome synthesis which should impacts the synthesis of other proteins).

6.2.3 In the absence of intracellular (p)ppGpp, the *rrn* expression rises upon nutrient downshift

Since *rrn* expression level are disturbed upon high intracellular levels of GTP for strains grown in minimal media, we wondered how the absence of the (p)ppGpp negative feedback regulations would affect the *rrn* expression during nutrient downshift. We thus plotted the overall OD and fluorescence levels as a function of time for the RelA$^+$ and (p)ppGpp^0 strains carrying the P$_{rrnJ}$*sfGFP* and P$_{rrnO}$*sfGFP* constructs grown in CH (figure 6.10 A). The expression level was similar between both strains during the exponential phase, but in the transition to the stationary phase, *rrn* promoters continued to be strongly expressed in the (p)ppGpp^0 strain while their expression was strongly decreased in the RelA$^+$ strain (figure 6.10 B and C). Then, the expression level became almost null (i.e. the fluorescence level was approximately constant) for both strains during the rest of the stationary phase. Thus, upon nutrient downshift and in the absence of (p)ppGpp, the *rrn* were more transcribed than they should be which may indicate that the GTP level were still high. We therefore anticipate that upon nutrient upshift the GTP level should rise again and lead to an increase in *rrn* transcription.
To test this hypothesis, we added fresh CH medium after the cells grown in CH medium had entered the transition phase. We observed in both strains that the growth rate as well as the *rrn* expression increased again after CH addition as compared to strains grown without the addition of fresh medium (figure 6.11). Interestingly, we observed that upon addition of CH the GFP concentration in the RelA$^+$ strains reached similar values as the ones found for the (p)ppGpp^0 strain grown without an addition of fresh CH. This results showed that in absence of (p)ppGpp, the transcription level of *rrn* continued to evolve as if there was no nutrient downshift. Nevertheless, the nutrients eventually became depleted which in the end impacted the concentration of GTP precursors leading to an arrest in *rrn* transcription.

6.2.4 During the transition to stationary phase, high GTP abundance results in high *rrn* expression levels according to the *rrn* promoter

To confirm that excess of GTP disturbs *rrn* transcription during the transition to the stationary phase, we cultivated the RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains carrying the P$_{rrnJ}$*sfGFP* and P$_{rrnO}$*sfGFP* constructs in CH with different concentrations of IPTG. To get a better insight into the *rrn* expression level during the transition to the stationary phase, we also defined the GFP abundance mean ratio $\frac{\overline{[GFP]}_{(p)ppGpp^0}}{\overline{[GFP]}_{RelA^+}}$ which is the ratio of the mean of the GFP abundance of the (p)ppGpp^0 *guaB*^{inducible} strain on the mean of the GFP abundance of the RelA$^+$ *guaB*^{inducible} strain for either the P$_{rrnJ}$*sfGFP* and P$_{rrnO}$*sfGFP* constructs from approximately 300 to 450 minutes which corresponds to the duration of the TFER peak observed for the (p)ppGpp^0 strain grown in CH (part 5). For IPTG concentrations equal or under 2.5 μM, the GFP abundance mean ratio $\frac{\overline{[GFP]}_{(p)ppGpp^0}}{\overline{[GFP]}_{RelA^+}}$ was near 1 for both *rrnJ* and *rrnO*, which could means that the *rrn* abundance was the same in the RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains (table 6.1). For IPTG concentration equal or above 15 μM, the GFP abundance mean ratio $\frac{\overline{[GFP]}_{(p)ppGpp^0}}{\overline{[GFP]}_{RelA^+}}$ increased with IPTG concentration to

reach a value of ≈1.2 for both *rrnJ* and *rrnO*. This could mean that upon high IPTG concentration, the *rrn* abundance was nearly 20% higher in the (p)ppGpp^0 *guaB*^{inducible} strain than in the RelA$^+$ *guaB*^{inducible} strain during the transition to the stationary phase.

Table 6.1: Table of the GFP abundance mean ratio $\frac{\overline{[GFP]_{(p)ppGpp^0}}}{\overline{[GFP]_{RelA^+}}}$ of the P$_{rrnJ}$*sfGFP* and P$_{rrnO}$*sfGFP* constructs carried by the RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains grown in CH with different IPTG concentrations.

IPTG concentration (μM)	0	2.5	15	50	200
$\frac{\overline{[GFP]_{(p)ppGpp^0}}}{\overline{[GFP]_{RelA^+}}}$ of P$_{rrnJ}$*sfGFP*	1.02	1.01	1.13	1.21	1.2
$\frac{\overline{[GFP]_{(p)ppGpp^0}}}{\overline{[GFP]_{RelA^+}}}$ of P$_{rrnO}$*sfGFP*	1.04	1.03	1.19	1.21	1.2

We also tested whether increasing GTP at high levels by adding GUO in CH medium at different concentrations would also affect the expression level of the *rrn*. We cultivated the RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains carrying the P$_{rrnJ}$*sfGFP* and P$_{rrnO}$*sfGFP* constructs in CH with different concentrations of GUO. For P$_{rrnJ}$*sfGFP*, the GFP abundance mean ratio $\frac{\overline{[GFP]_{(p)ppGpp^0}}}{\overline{[GFP]_{RelA^+}}}$ was near 1 at GUO concentrations of 0, 50 and 200 μM; and it increased up to 1.38 at GUO 500 μM and 1.23 at GUO 1000 μM (table 6.2). For P$_{rrnO}$*sfGFP*, the GFP abundance mean ratio $\frac{\overline{[GFP]_{(p)ppGpp^0}}}{\overline{[GFP]_{RelA^+}}}$ was near 1 in absence of extracellular GUO while it increased with GUO concentration to reach a maximal value of 2.16 at GUO 500 μM. Thus, the *rrnO* abundance could be more than twice higher in the (p)ppGpp^0 *guaB*^{inducible} strain than in the RelA$^+$ *guaB*^{inducible} strain at GUO 500 μM. Hence, these results suggest that high intracellular GTP levels enhanced the transcription of *rrn* in a significant manner and that the sensitivity is dependent on the *rrn* promoter sequence since the observed effect was more pronounced for P$_{rrnO}$*sfGFP* as compared to P$_{rrnJ}$*sfGFP*.

Table 6.2: Table of the GFP abundance mean ratio $\frac{\overline{[GFP]_{(p)ppGpp^0}}}{\overline{[GFP]_{RelA^+}}}$ of the P$_{rrnJ}$*sfGFP* and P$_{rrnO}$*sfGFP* constructs carried by the RelA$^+$ *guaB*^{inducible} and (p)ppGpp^0 *guaB*^{inducible} strains grown in CH with different GUO concentrations.

GUO concentration (μM)	0	50	200	500	1000
$\frac{\overline{[GFP]_{(p)ppGpp^0}}}{\overline{[GFP]_{RelA^+}}}$ of P$_{rrnJ}$*sfGFP*	1.02	1.04	1.05	1.38	1.23
$\frac{\overline{[GFP]_{(p)ppGpp^0}}}{\overline{[GFP]_{RelA^+}}}$ of P$_{rrnO}$*sfGFP*	1.04	1.28	1.72	2.16	1.78

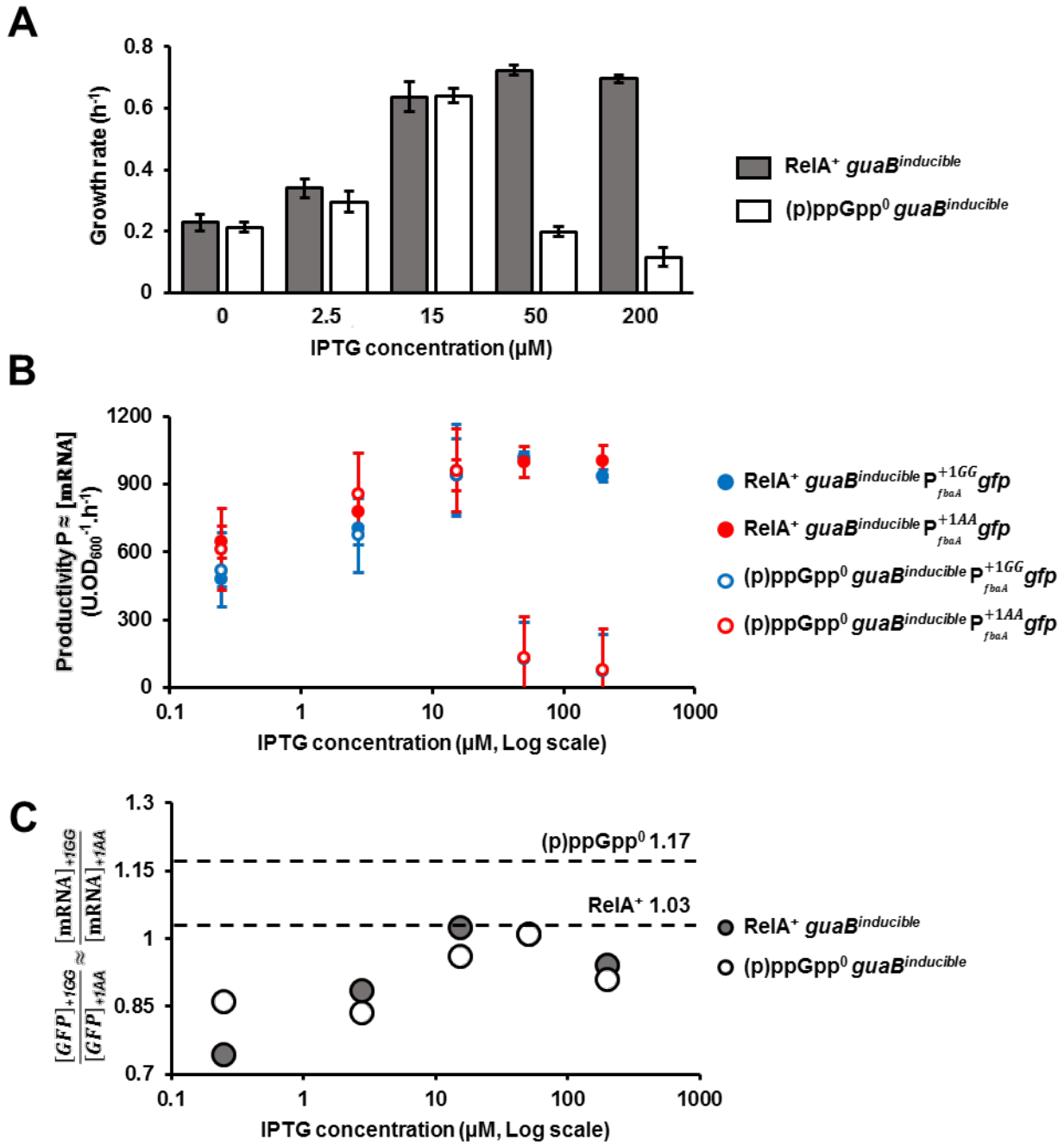


Figure 6.3: Impact of the control of intracellular GTP level on gene expression according to the TSS composition for strains grown in M9G with IPTG added. A. Mean of the growth rate of the RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains grown in M9G with different concentrations of IPTG. The error bars correspond to the standard deviation of the growth rate calculated using multiple replicates. B. GFP productivity of the RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains carrying the P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* constructs as function of the IPTG concentration. The error bars correspond to the standard deviation of the GFP productivity calculated using multiple replicates. C. GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ of the RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains as function of the IPTG concentration. The black dashed lines correspond to the value of this ratio for the RelA$^+$ and (p)ppGpp^0 strains grown in M9G.

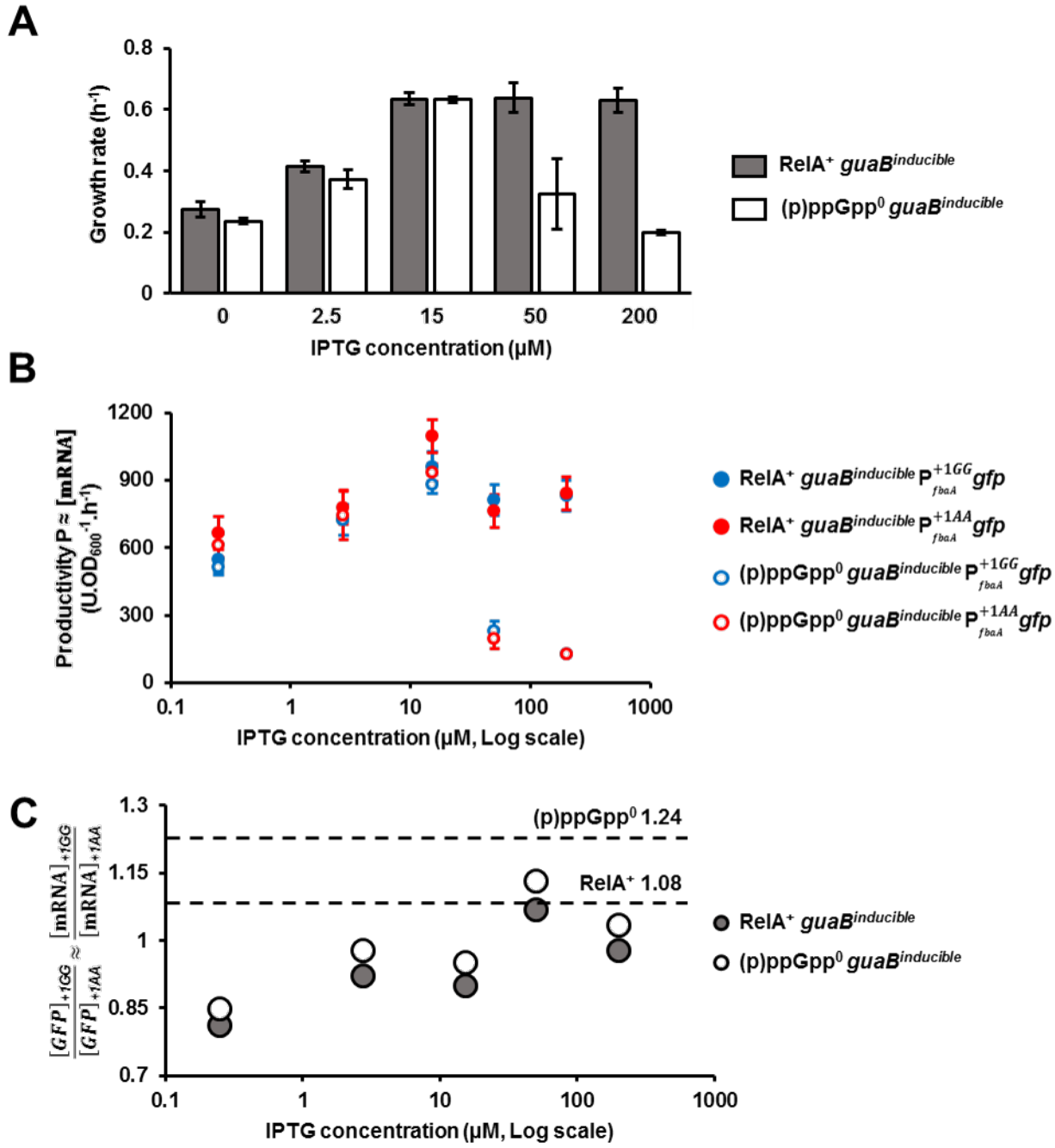


Figure 6.4: Impact of the control of intracellular GTP level on gene expression according to the TSS composition for strains grown in M9M with IPTG added. A. Mean of the growth rate of the RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains grown in M9M with different concentrations of IPTG. The error bars correspond to the standard deviation of the growth rate calculated using multiple replicates. B. GFP productivity of the RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains carrying the P_{fbaA}^{+1GG}*gfp* and P_{fbaA}^{+1AA}*gfp* constructs as function of the IPTG concentration. The error bars correspond to the standard deviation of the GFP productivity calculated using multiple replicates. C. GFP abundance ratio $\frac{[GFP]_{+1GG}}{[GFP]_{+1AA}}$ of the RelA$^+$ *guaB*^{inducible}and (p)ppGpp^0 *guaB*^{inducible} strains as function of the IPTG concentration. The black dashed lines correspond to the value of this ratio for the RelA$^+$ and (p)ppGpp^0 strains grown in M9M.

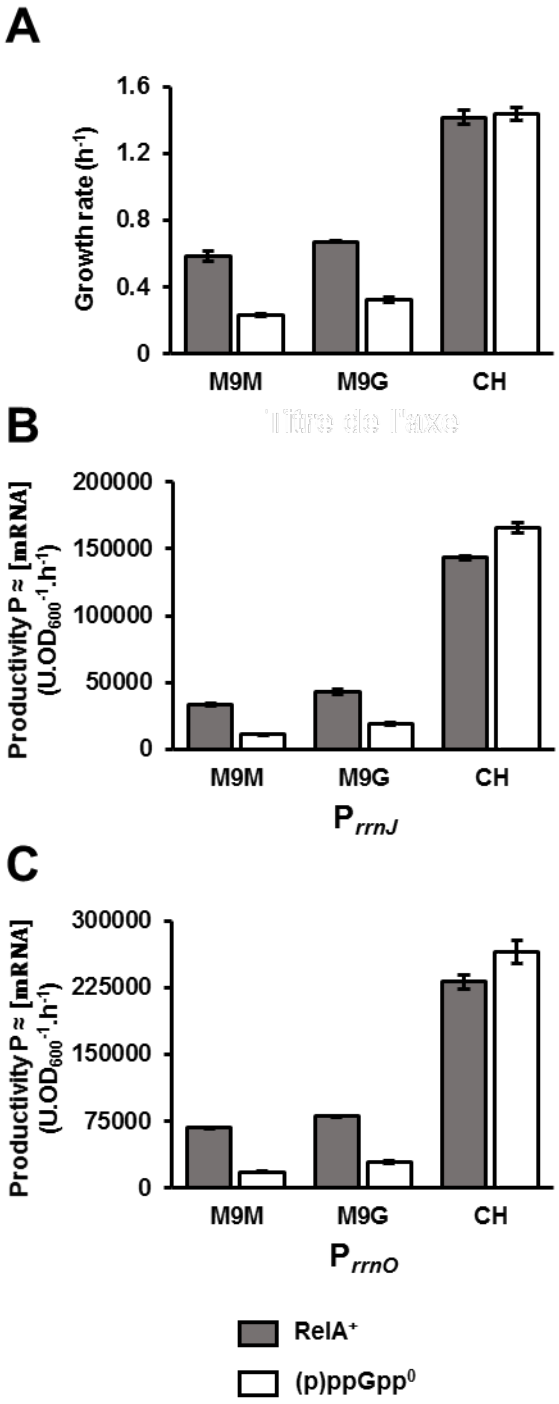


Figure 6.7: *rrn* expression in different growth media in the presence or absence of (p)ppGpp . A. Growth rate of the RelA$^+$ and (p)ppGpp^0 strains grown in M9M, M9G and CH (see 8.4.1). B. GFP productivity of the RelA$^+$ P_{rrnJ}*sfGFP* and (p)ppGpp^0 P_{rrnJ}*sfGFP* strains grown in M9M, M9G and CH. C. GFP productivity of the RelA$^+$ P_{rrnO}*sfGFP* and (p)ppGpp^0 P_{rrnO}*sfGFP* strains grown in M9M, M9G and CH. The error bars correspond to the standard deviation of the growth rate (A) or GFP productivity (B and C) calculated using multiple replicates.

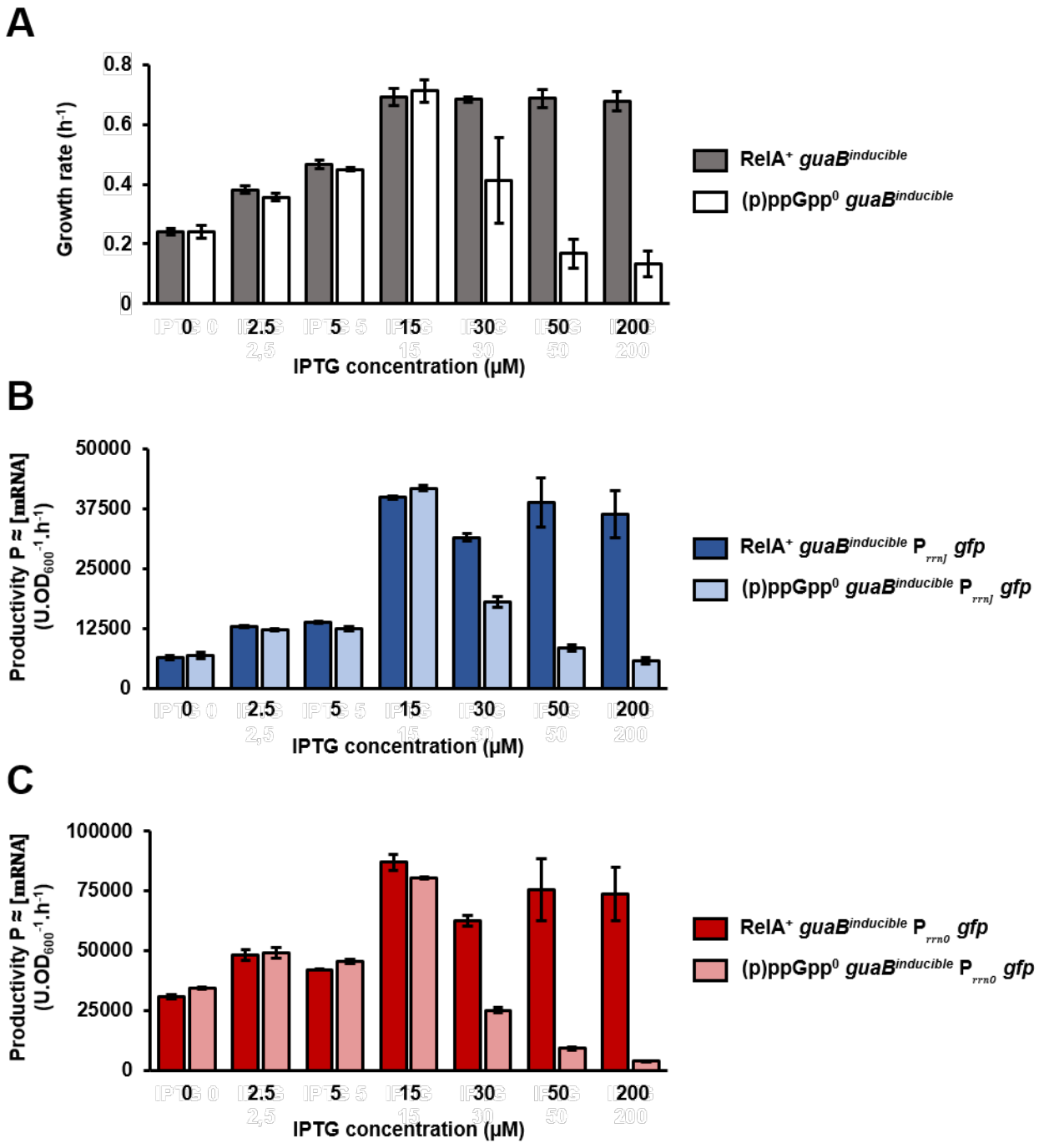


Figure 6.8: Impact of the control of intracellular GTP level on *rrns* expression for strains grown in M9G with IPTG added. A.Mean of the growth rate of the RelA$^+$ $guaB^{inducible}$ and (p)ppGpp^0 $guaB^{inducible}$ strains grown in M9G with different concentrations of IPTG. B. GFP productivity of the RelA$^+$ $guaB^{inducible}$ and (p)ppGpp^0 $guaB^{inducible}$ strains carrying the $P_{rrnJ}sfGFP$ construct for different IPTG concentrations. C. GFP productivity of the RelA$^+$ $guaB^{inducible}$ and (p)ppGpp^0 $guaB^{inducible}$ strains carrying the $P_{rrnO}sfGFP$ construct for different IPTG concentrations. The error bars correspond to the standard deviation of the growth rate (A) or GFP productivity (B and C) calculated using multiple replicates.

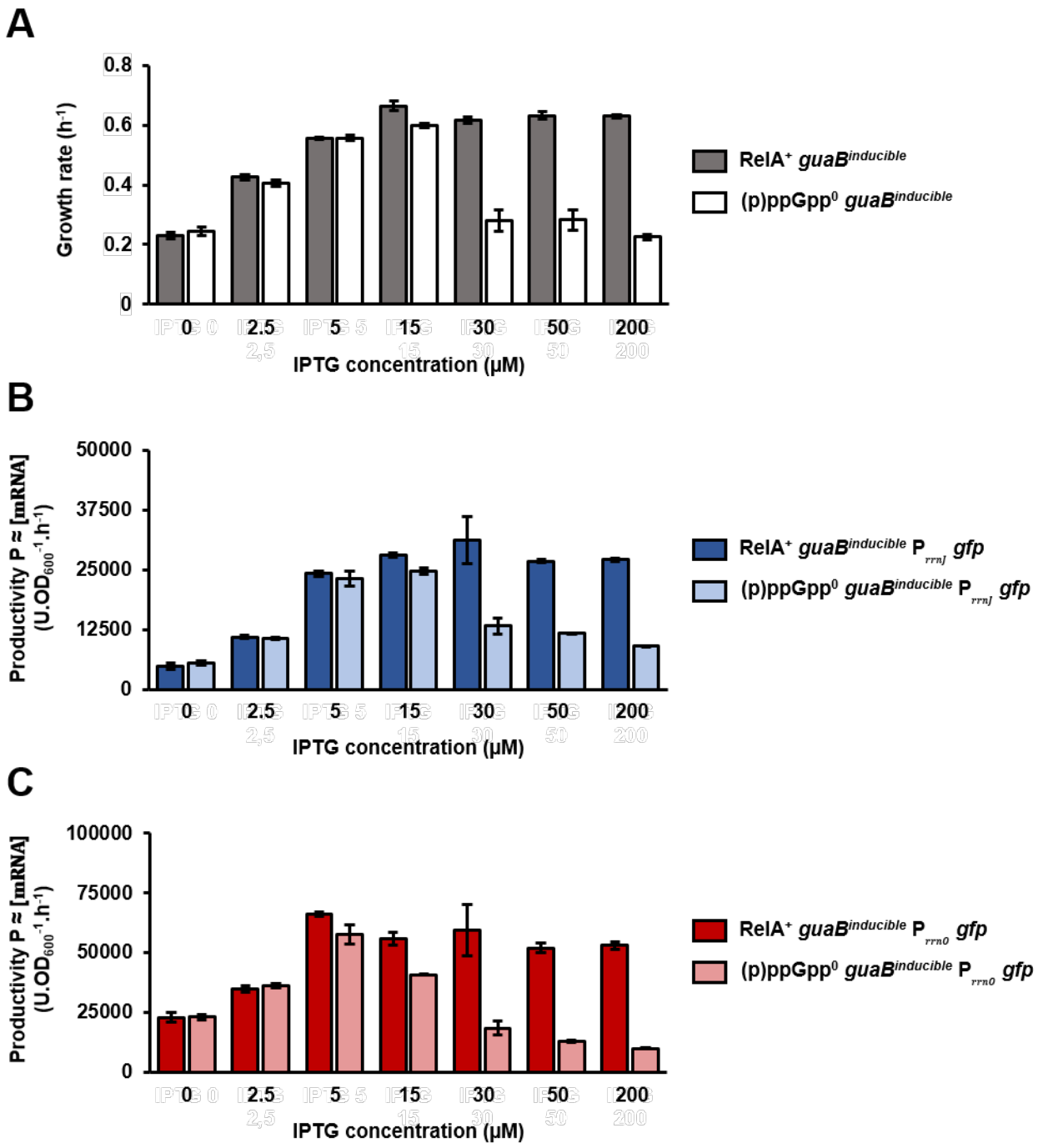


Figure 6.9: Impact of the control of intracellular GTP level on *rrns* expression for strains grown in M9M with IPTG added. A.Mean of the growth rate of the RelA$^+$ $guaB^{inducible}$ and (p)ppGpp^0 $guaB^{inducible}$ strains grown in M9M with different concentrations of IPTG. B. GFP productivity of the RelA$^+$ $guaB^{inducible}$ and (p)ppGpp^0 $guaB^{inducible}$ strains carrying the $P_{rrnJ}sfGFP$ construct for different IPTG concentrations. C. GFP productivity of the RelA$^+$ $guaB^{inducible}$ and (p)ppGpp^0 $guaB^{inducible}$ strains carrying the $P_{rrnO}sfGFP$ construct for different IPTG concentrations. The error bars correspond to the standard deviation of the growth rate (A) or GFP productivity (B and C) calculated using multiple replicates.

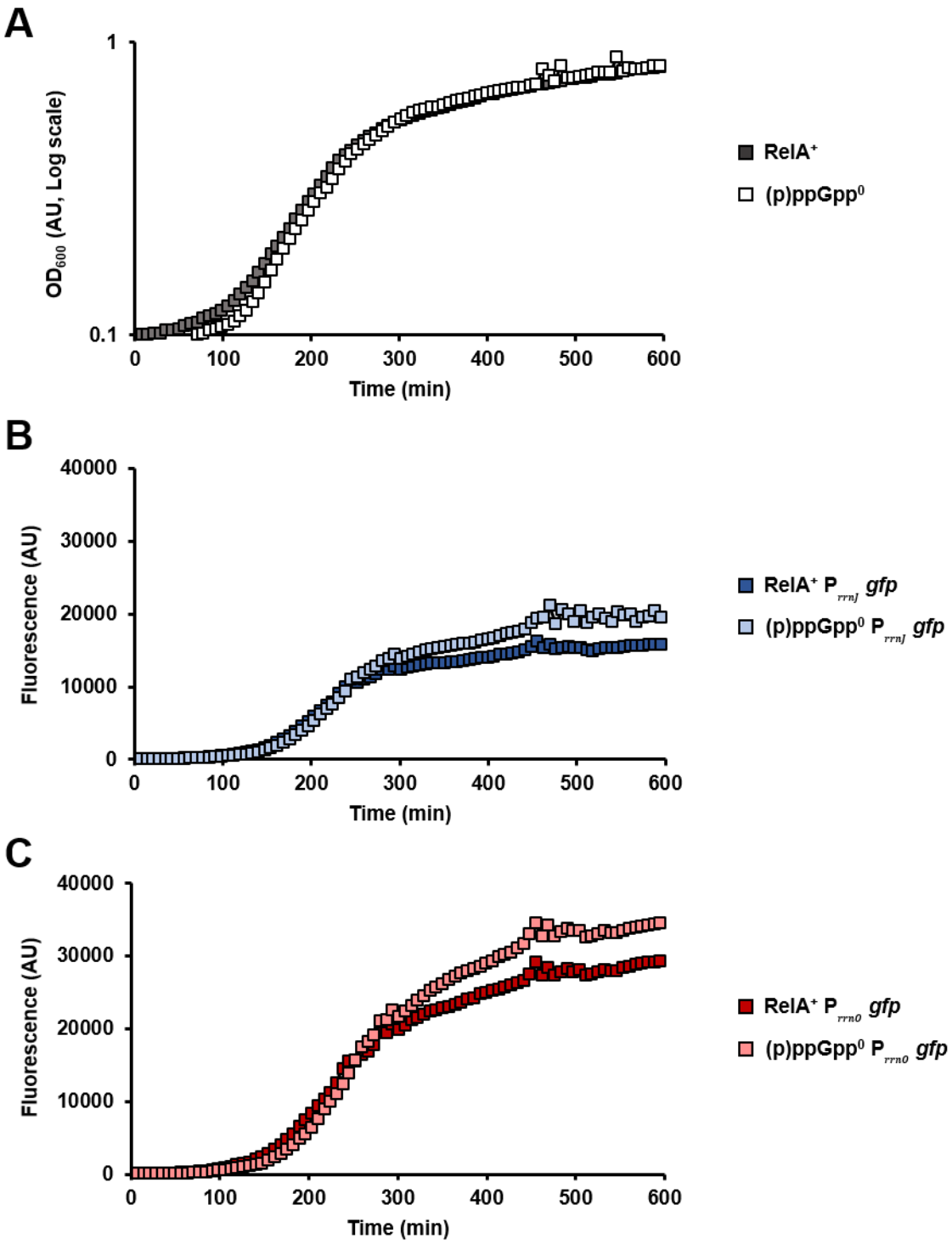


Figure 6.10: Effect of the absence of intracellular (p)ppGpp on *rrn* expression upon nutrient downshift. A. OD_{600} as function of time for the RelA$^+$ and (p)ppGpp^0 strains grown in CH. B. Fluorescence as function of time for the RelA$^+$ P$_{rrnJ}$*sfGFP* and (p)ppGpp^0 P$_{rrnJ}$*sfGFP* strains grown in CH. C. Fluorescence as function of time for the RelA$^+$ P$_{rrnO}$*sfGFP* and (p)ppGpp^0 P$_{rrnO}$*sfGFP* strains grown in CH.

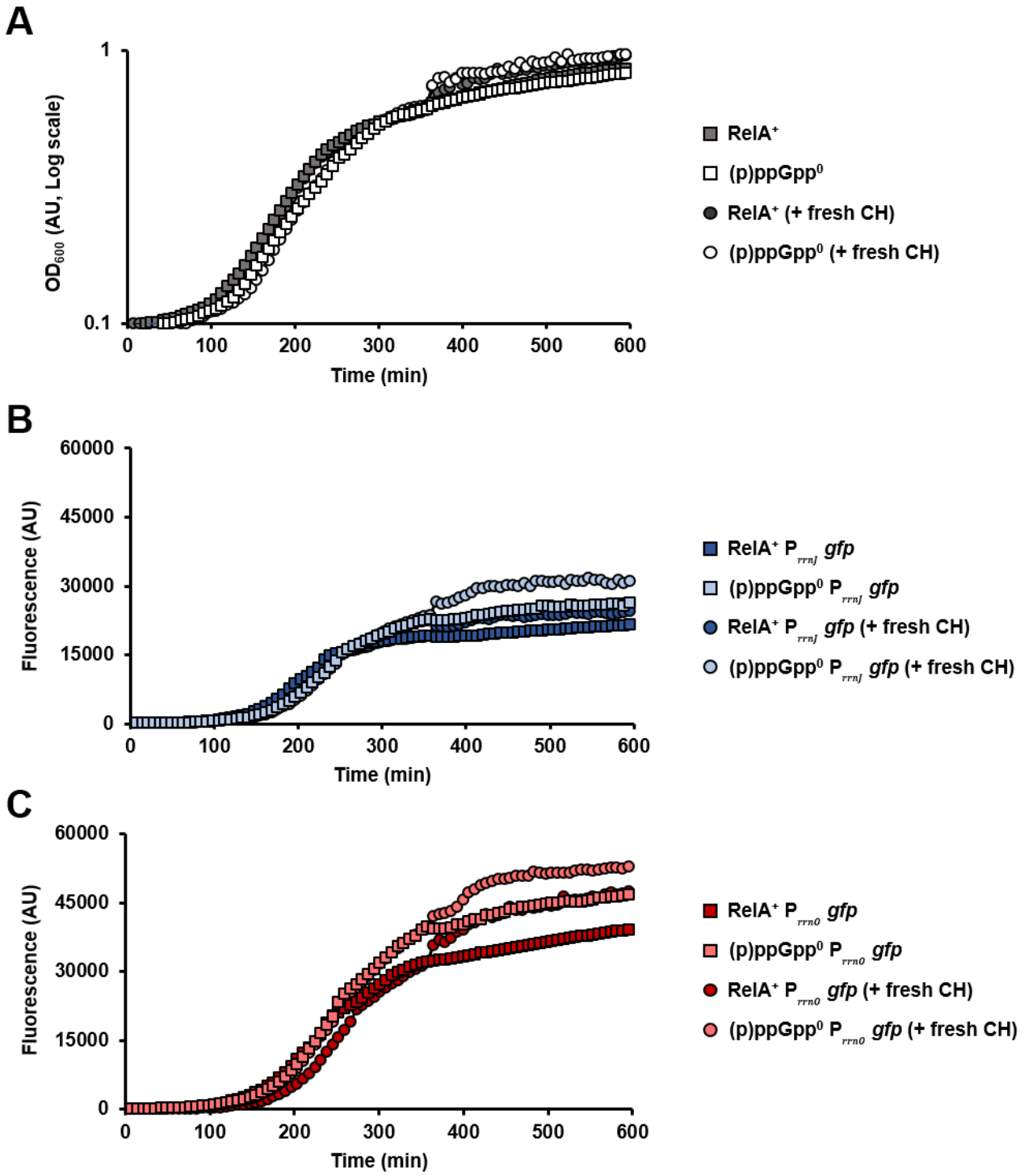


Figure 6.11: Effect of the presence or the absence of intracellular (p)ppGpp on *rrn* expression upon nutrient upshift. A. OD_{600} as function of time for the RelA$^+$ and (p)ppGpp^0 strains grown in CH (squares) or in CH with fresh CH added when cells had entered the transition phase (rounds). B. Fluorescence as function of time for the RelA$^+$ P_{rrnJ}*sfGFP* and (p)ppGpp^0 P_{rrnJ}*sfGFP* strains grown in CH (squares) or in CH with fresh CH added when cells had entered the transition phase (rounds). C. Fluorescence as function of time for the RelA$^+$ P_{rrnO}*sfGFP* and (p)ppGpp^0 P_{rrnO}*sfGFP* strains grown in CH (squares) or in CH with fresh CH added when cells had entered the transition phase (rounds).

Part III

Discussion

Chapter 7

Discussion and perspectives

7.1 Determining the absolute TFE rate

The values obtained with the different TFE reporter systems are unlikely to reflect the average level of translational errors which can occur for any protein in the cell. Indeed, this would mean that at every codon there is a 1-10% chance that a translational error occurs and consequently a protein of 100 amino acids or more would always be defective. Hence, the cell would produce only non-functional proteins. This 1-10% range for the TFE rate was already proposed by Meyerovich *et al.* (2010) but as explained it is not coherent with cell's viability. As observed, when we modified the GFP sequence to map the +1 frameshift, we did not detect fluorescence and thus TFEs anymore. It does not mean that TFE did not occur anymore but the TFE rate is strictly lower than the detection threshold of 0.1% ($<10^{-3}$ per codon) that we detected with our home-made TFE reporter systems. Indeed, it was reported that "spontaneous" frameshifts occur at a range of $5x10^{-5}$ to $3x10^{-3}$ per codon [Parker, 1989]. The range of 10^{-4} TFE per codon is unlikely to alter the cell's fitness but it might be too low to be detected through fluorescence monitoring. Indeed, the lowest level we detected was of 0.1% (10^{-3}) but our reporter systems still provides a good insight into how the TFE rate evolves according to the growth medium as well as the growth phase. Ideally, an interesting possibility would be to use mass spectrometry to sequence a subset of his-tagged proteins (such as a GFP or an essential protein of *B. subtilis* after purification) and identify whether misincorporations or frameshift event occurred after growing the strain containing the tagged proteins in selected media. This might provide a quite reliable order of magnitude of the average level of translational errors which can occur for any proteins in the cell according to growth conditions.

7.2 Comparison of the TFE rate between *B. subtilis* and other bacterial species

We showed that during steady-state growth in poor medium, the TFE rate is higher in the absence of intracellular (p)ppGpp. Similar results were obtained for *E. coli* strains grown in minimal media: the TFE level was higher in a *relA*- strain as compared to the *relA*+ strain [Hall and Gallant, 1972].

Moreover, we showed that the TFE rate peaks during a nutritional downshift when RelA as well as the secondary (p)ppGpp synthetases RelP and RelQ are deleted. Experiments during which *E. coli* strains were grown in minimal media and drugs were injected to mimic amino acid starvation also led to higher TFE rates in the *relA*- strain as compared to the *relA*+ strain [Masucci et al., 2002]. However, Masucci *et al.* (2002) used a reporter system which did not allow to monitor the TFE rate during the overall cell growth. We therefore expect that, similarly to *B. subtilis*, the TFE level also peaks during the transition to the stationary phase in *E. coli* in the absence of RelA to sense the charging level of tRNAs. Nevertheless, it might be more difficult to experimentally determine in *E. coli* how the (p)ppGpp/GTP regulatory feedback loops contribute to resource allocation and in fine to TFE modulation. Indeed, the *E. coli relA*- strain still possesses the SpoT enzyme which can synthesize (p)ppGpp even though amino acid limitation does not constitute a trigger for its activity. In addition, (p)ppGpp interacts directly with the RNAP in *E. coli* to redirect transcription of a broad range of promoters [Chatterji et al., 1998][Toulokhonov et al., 2001] which is not the case in *B. subtilis*. Hence this complicates the study of the (p)ppGpp/GTP feedback loops as compared to *B. subtilis*.

Furthermore, during the stationary phase, the TFE rate was higher than during the exponential phase in CH medium, whether the cell still possessed the *relA* gene or not. This observation is consistent with previous works where TFEs were found to be increased during stationary phase in *B. subtilis* [Meyerovich et al., 2010] as well as in mycobacteria [Javid et al., 2014] which conferred fitness advantages to cells upon stressful conditions such as antibiotic treatments [Meyerovich et al., 2010, Javid et al., 2014].

7.3 Role of RelP and RelQ during nutritional downshift

The TFE rate of the WT and RelA$^+$ strains was the same during exponential growth in poor medium. However, the TFE rate peaked during the transition to the stationary phase in the absence of the *relP* and *relQ* genes (RelA$^+$ strain) but not in the WT strain. Hence, this suggests that these enzymes are more active when *B. subtilis* faces environmental changes than during its steady-state growth. Indeed, *relP* transcripts are only present during early stationary phase and under specific treatments such as the addition of antibiotics, ethanol, high salt and acidic or alkalic pH stress conditions [Geiger et al., 2014][Thackray and Moir, 2003][Zweers et al., 2012]. Concerning *relQ*, it is predominantly transcribed during exponential growth [Nanamiya et al., 2008] but the relQ enzyme can switch from a "passive state" (low (p)ppGpp synthetase activity) to an "active state" (high (p)ppGpp synthetase activity) [Steinchen et al., 2018]. Moreover, it has been suggested that its activity is stimulated by amino acid starvation and also by the cell energy imbalance (i.e. if there is a great excess of GDP) [Arenz et al., 2016] and its activity could be intensively coupled to RelA activity [Steinchen et al., 2018]. Hence, even though RelQ is already present during the exponential phase, it is likely that it is in a "passive state" ready to switch to an "active state" during the transition to stationary phase in order to rapidly assist RelA in producing (p)ppGpp when nutrients are depleted and thus prevent

the TFE rate to peak. RelP could also enhance (p)ppGpp production to help the cell to cope with a nutritional downshift but it might be more needed when the cell is already in the stationary phase. It could be interesting to see whether RelQ activity is more important than RelP activity to prevent TFEs burst during transition to stationary phase by growing mutants with either *relP* or *relQ* deleted in CH medium, and then compare their respective TFE rates.

7.4 *B. subtilis* makes use of the TSS as a way to allocate resources according to the GTP/ATP ratio which is growth rate dependent

We showed that the transcription level of genes which possess guanines as TSS increased more importantly with growth rate than for genes which possess adenines. These observations can be correlated to the GTP/ATP ratio which also increases with growth rate [Bittner et al., 2014]. Krasny *et al.* (2008) observed that during the stringent response the modified *ilv* promoter (guanine as TSS) was downregulated while the native *ilv* promoter was upregulated (adenine as TSS). In vitro experiments also showed that the TSS of a promoter partly determined its sensitivity to the GTP/ATP ratio [Krásnỳ and Gourse, 2004][Krásnỳ et al., 2008]. In *B. subtilis*, promoters of genes encoding proteins involved in growth enhancement (ribosomal proteins RpsL and RpsG, enzymes of the glucose and pyruvate metabolism) possess a guanine as TSS and their transcription positively correlates with GTP abundance [Krásnỳ et al., 2008][Tojo et al., 2010]. The enhancement of the transcription of genes possessing a guanine as TSS when the GTP/ATP ratio is high occurs at the expense of the transcription of genes possessing an adenine as TSS (i.e. the RNAP is allocated to different promoters). Part of these genes encode enzymes involved in cellular processes mainly turned on during nutrient depletion such as amino acid production [Krásnỳ et al., 2008][Kriel et al., 2014] or metabolic pathways producing important carbohydrates like pyruvate [Tojo et al., 2008][Tojo et al., 2010]. However, when the GTP/ATP ratio becomes lower under starvation conditions, these genes are upregulated. Furthermore, given that rRNA synthesis is the rate-limiting step in ribosome synthesis [Henkin, 2002][Paul et al., 2004b] and that all rRNA promoters initiate with GTP [Krásnỳ and Gourse, 2004], this makes GTP abundance the driving force for ribosome production and in fine growth [Bittner et al., 2014]. Together, these results show that *B. subtilis* uses the TSS of promoters as a way to direct the RNAP towards the promoters of genes involved in strategical cellular processes according to the GTP/ATP ratio, which is growth rate dependent. These observations can be paralleled with the *E. coli* transcription machinery abundance which specifically influences the expression of each gene according to their promoter sequence in a growth-rate dependent manner [Gerosa et al., 2013]. Consequently, a disequilibrium of the GTP/ATP ratio through disturbance of the GTP/(p)ppGpp feedback regulatory loops could lead to non-optimal distribution of the RNAPs across the promoters and thus affect the growth rate. Indeed, we observed for poor and rich media that in the absence of (p)ppGpp the transcription is partly directed towards genes possessing guanine as TSS at the expense of genes possessing adenine as

TSS. For the (p)ppGpp^0 strain grown in CH, the difference between the transcription level of the two TSS types increased with the addition of GUO and so with the GTP/ATP ratio [Bittner et al., 2014]. The growth rate obtained at GUO concentration of 500 and 1000 μM was lower for the (p)ppGpp^0 strain than the growth rate normally obtained during growth in CH, even though a high GTP/ATP ratio is supposed to be favorable for growth. A possible explanation is that certain proteins required for growth might be overexpressed at the expense of proteins essential for bacterial fitness, which eventually resulted in a non-optimal resource allocation between the different cellular processes at the expense of growth. This is consistent with the current model of growth (section 2.2.2).

7.5 *Rrn* transcription and in fine ribosome synthesis positively correlate with the GTP/ATP ratio

We showed that the *rrn* expression levels were higher in rich medium than poor medium and that it increased with growth rate and thus with the GTP/ATP ratio. This agrees with what has been discussed previously given that all rRNA promoters have a guanine as TSS [Krásnỳ and Gourse, 2004]. However, we could not directly compare the *rrn* expression level between the RelA$^+$ and (p)ppGpp^0 strain since the GFP productivity used to monitor transcription was also dependent on translation. As suggested previously, disturbing the GTP/(p)ppGpp feedback regulatory loops also alters the transcription and translation processes which potentialy leads to conclusions that could seem contradictory (i.e. the GFP productivity reporting *rrn* expression is lower for the (p)ppGpp^0 strain than for the RelA$^+$ strain even though the GTP/ATP ratio is higher). To circumvent this difficulties, it would be interesting to redo the same experiments (i.e. IPTG gradient in poor and rich media) but instead of monitoring the GFP fluorescence, perform a ribosome profiling of the RelA$^+$ $guaB^{inducible}$ and (p)ppGpp^0 $guaB^{inducible}$ strains. In addition, to get a better insight into how the ribosome concentration evolves during steady-state growth, it would be interesting to implement in the RelA$^+$ strain a construct where a fluorescent protein is fused to an essential ribosomal protein.

Moreover, we observed for a RelA$^+$ strain grown in CH that *rrn* transcription increased during exponential growth and then slowed down during the transition to the stationary phase. Similar observations were made by Rosenberg *et al.* (2012) who concluded that *rrn* expression is growth phase dependent in rich medium. However, in the absence of (p)ppGpp, we observed that the transcription level of *rrn* continued to increase even though there was a nutrient downshift. A possible explanation is that the GTP/ATP ratio is still high in the (p)ppGpp^0 strain even after nutrient depletion. Indeed, Kriel *et al.* (2014) observed that after RHX treatment (drug used to simulate amino acid starvation), the GTP/ATP ratio is higher than before injection. Since *rrn* transcription is the limiting factor for ribosome production [Henkin, 2002][Paul et al., 2004b], these results suggest that ribosomes are still overproduced when the GTP/ATP ratio is important. In conclusion, during a nutrient downshift, when the GTP/ATP ratio is not controlled by (p)ppGpp, the translation machinery is in high demand

for resources such as amino acids needed for translation. However, these resources cannot be provided since the medium is depleted in nutrients and the enzymes required to synthetize amino acids are not produced (i.e. CodY represses the transcription of the genes which encode these enzymes [Sonenshein, 2007]) at high GTP/ATP ratio [Kriel et al., 2014]. Nevertheless, *rrn* transcription eventually slowed down which could be due to the lack of nutrients required for the biosynthesis of GTP precursors and thus GTP.

7.6 Non-optimal GTP/ATP ratio is responsible for high TFEs during steady-state growth and contribute to TFE burst during nutritional downshift

During steady-state growth in poor medium, we showed that beyond the basal TFE level GTP itself may be the TFE trigger factor, and that (p)ppGpp was not strictly required to reduce TFE occurrence but most likely maintained lower TFE rates in WT cells via its feedback regulation on GTP biosynthesis. In addition, we observed that high GTP levels during exponential growth (i.e. extracellular GUO added) in rich medium also led to higher TFE rates in the absence of (p)ppGpp. During exponential growth in CH with GUO added at a concentration of 1000 μM, mRNAs possessing GG as TSS were about 3.5 times more produced than the mRNAs possessing AA as TSS; while the TFE rate in the (p)ppGpp^0 *guaB*^{inducible} strain was about 3 times higher than in the RelA$^+$ *guaB*^{inducible} strain. Such difference could be attributed to the fact that a high GTP/ATP ratio leads to an enhanced translation machinery as previously discussed.

Moreover, during transition to the stationary phase in CH, we observed that the TFE peaks in the (p)ppGpp^0 strain correlates to an increase in the *rrn* expression level and so in ribosome synthesis. When increasing the GTP/ATP ratio by adding extracellular guanosine to the CH growth medium, we observed that the *rrn* were more expressed in the (p)ppGpp^0 strain than the RelA$^+$ strain during the peak duration. The higher was the GTP/ATP ratio, the higher was the difference in the *rrn* expression between the two strains as well as the peak height. However, controlling the intracellular GTP level during growth in minimal media resulted in a similar distribution of the RNAP between the promoters according to their TSS in the (p)ppGpp^0 and RelA$^+$ strains. This GTP control also restored a similar *rrn* expression level as well as a similar TFE rate between the (p)ppGpp^0 and RelA$^+$ strains.

Overall, these results suggest that disturbing the GTP/(p)ppGpp regulatory feedback loops and thus increasing the GTP/ATP ratio leads to resource reallocation towards the translation machinery; which correlates with a higher TFE rate during steady-state growth phase and the transition to the stationary phase. Nevertheless, restoring an optimal resource allocation through control of the intracellular GTP level in a (p)ppGpp^0 strain restores its TFE rate to a level similar to the WT TFE rate during

exponential growth. This means that production of (p)ppGpp by RelA does not only take place during environmental changes (nutrient depletion or other cellular stresses) as assumed in the literature [Potrykus and Cashel, 2008][Hauryliuk et al., 2015] until recently but also during steady-state growth (i.e. exponential growth). This is in line with what Bittner *et al.* (2014) suggested, that overloaded GTP levels (i.e. not controlled by (p)ppGpp) can induce stress and inhibit growth instead of enhancing growth even if all amino acids are present in the growth medium. Hence, RelA is always sensing the charging levels of tRNAs and so constantly produces either GDP or (p)ppGpp according to the resources available, which makes the (p)ppGpp abundance also determinant for growth rate adjustment as already suggested by Marr [1991].

7.7 Consequences of disturbing the GTP/(p)ppGpp regulatory feedback loops on tRNA charging levels and TFEs

As seen previously, disturbing the GTP/(p)ppGpp regulatory feedback loops led to a high intracellular GTP/ATP ratio as well as a high ribosome production. Consequently, the translation machinery is in high demand for charged tRNAs that cannot be provided during a nutritional downshift; which should lead to higher translational error levels in a strain which cannot sense the tRNA charging level by RelA ((p)ppGpp^0 in *B. subtilis*) as suggested by Sorensen (2001) (see Problematic). Indeed, we observed during growth in rich medium that the TFE rate peaked in the transition to the stationary phase in the absence of (p)ppGpp. In *E. coli*, an increase in TFEs level was also observed after the injection of a drug which mimics starvation of the amino acid corresponding to the "hungry" codon of the frameshift prone sequence [Masucci et al., 2002]. We decided to reproduce a similar experiment by injecting drugs which simulate a leucine or isoleucine starvation using the GFP^{Ile}_{fs-1} and GFP^{Leu}_{fs-1} reporter systems. We tested several drug concentrations but either we did not see any difference with the TFE rate measured in rich and poor medium or the cells died just after the injection. Hence, the abundance of the "cognate" loaded tRNA may not be by itself triggering the pause at the "hungry codon" site as it has been proposed [Barak et al., 1996]. A recent work could explain such possibility: they found that the majority of tRNAs are degraded upon amino acid starvation, including both the cognate and non-cognate tRNAs of the amino acid *E. coli* was starved for; and this was observed for both *relA*+ and *relA*- strains [Svenningsen et al., 2016]. This means that the starvation induced tRNA degradation is independent of the (p)ppGpp-mediated stringent response. Inhibiting mRNA synthesis also resulted in tRNA degradation, which led to the conclusion that tRNA degradation is always triggered as response to amino acid starvation [Svenningsen et al., 2016]. This means that, even if the cell is starved for an amino acid not coded by the "hungry" codon of the frameshift prone sequence, the uncharged tRNA corresponding to this "hungry" codon will still be degraded and in fine the abundance of its charged form will decrease. Consequently, the TFE rate will still rise if the translation machinery is not repressed in order to adapt the demand for amino acids in accordance to the availability of charged tRNAs. Our results agree with such point of view since the burst in

TFEs occurred during a global nutrient downshift (i.e. the cell is not starved for a particular amino acid) for different frameshift prone sequences which possessed different "hungry" codons. Hence, the supply and demand in resources appear to play a crucial role in the TFEs, whether the cell is starved for the amino acid corresponding to the frameshift prone sequence "hungry" codon or another codon. To confirm this hypothesis, it would be interesting to insert into the (p)ppGpp^0 and RelA$^+$ strains containing the GFP^{Tyr}_{fs+1} reporter system a second TFE reporter system using the mKate2 fluorescent protein (red color emitted) containing the frameshift prone sequence possessing leucine as "hungry" codon. We could thus check at the population and single-cell level if the TFE rates of these two reporter systems evolve similarly (i.e. if in CH the TFE rate starts to peak simultaneously or not).

7.8 The cell uses a fast and a slow regulatory feedback loop to control its TFE rate

The control of intracellular GTP levels can prevent a too high intracellular GTP/ATP ratio as well as a too high *rrn* expression level which should lead to an optimal ribosome concentration and in fine TFE rate. However, the control of the GTP/ATP ratio as well as the *rrn* synthesis rate by induction of *guaB* expression at IPTG concentration leading to sub-optimal growth rates did not prevent a burst in the TFE rate in the transition to the stationary phase. Hence, reducing the ribosome concentration might not be enough to reduce the demand in charged tRNAs and thus to prevent a TFE burst when the cell faces a nutritional downshift. Nevertheless, this demand can be importantly reduced by the action of the (p)ppGpp on IF2 since inhibiting translation initiation in the (p)ppGpp^0 strain prevented a TFE rate burst during the transition to the stationary phase and restored the level of TFEs observed in the RelA$^+$ strain. We propose that (p)ppGpp's action on IF2 is the key to rapidly optimize nutrients supply and demand so they match in order to avoid a burst in TFEs. Therefore, these results strongly suggest that: (i) (p)ppGpp's action on IF2 acts as a fast negative feedback loop to adapt the overall translation apparatus activity to the level of charged tRNAs and prevent a TFE burst; and (ii) the inhibitory action of (p)ppGpp on the activity of enzymes involved in GTP biosynthesis to make the GTP level drop constitutes a slow feedback loop whose goal is to decrease the ribosome concentration to match the level of charged tRNAs during the stationary phase to keep the TFE rate under a certain level (figure 7.1).

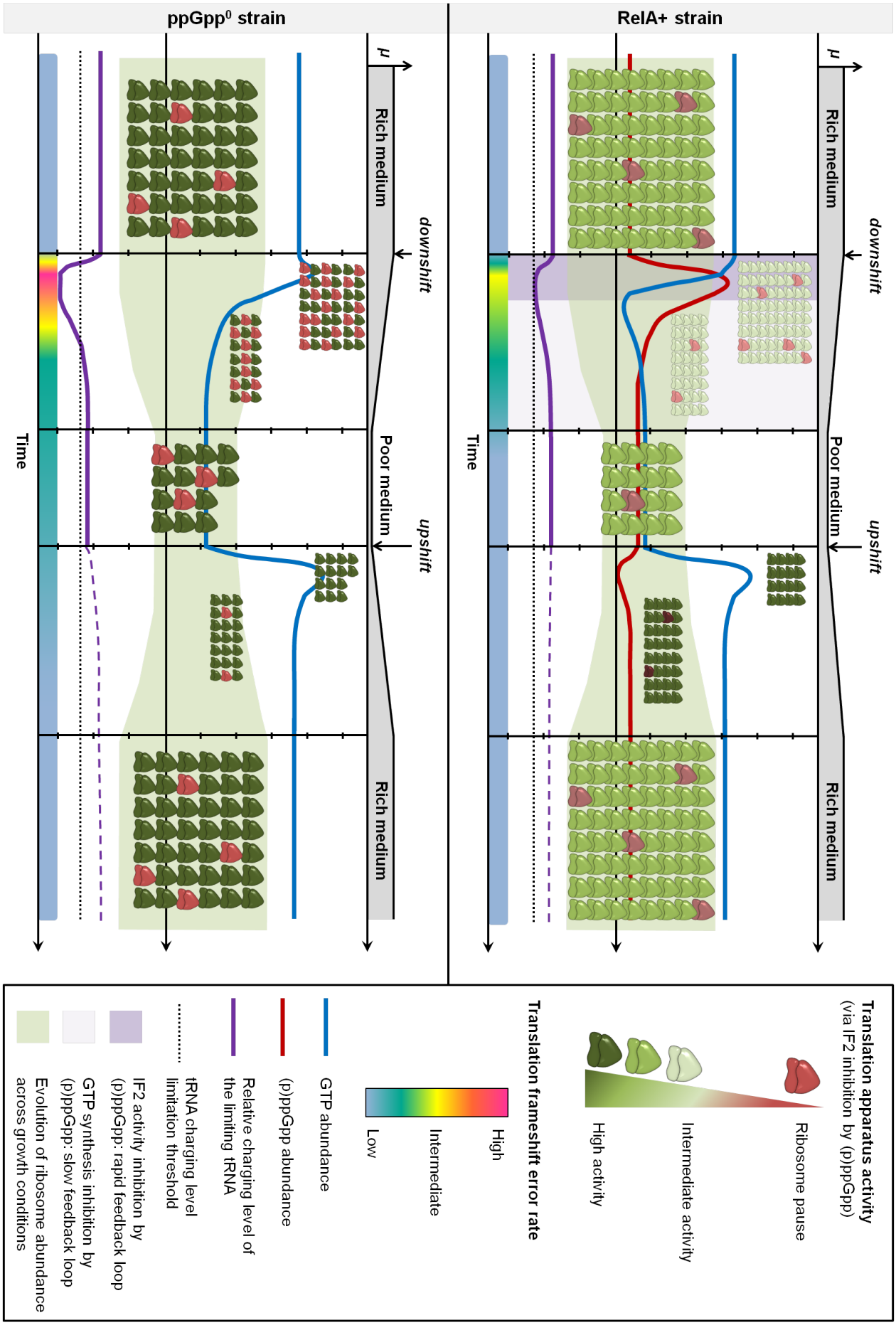

RelA+ strain
ppGpp0 strain
μ
Rich medium
Poor medium
Rich medium
downshift
upshift
Time
Translation apparatus activity
(via IF2 inhibition by (p)ppGpp)
Ribosome pause
Intermediate activity
High activity
Translation frameshift error rate
High
Intermediate
Low
GTP abundance
(p)ppGpp abundance
Relative charging level of the limiting tRNA
tRNA charging level limitation threshold
IF2 activity inhibition by (p)ppGpp: rapid feedback loop
GTP synthesis inhibition by (p)ppGpp: slow feedback loop
Evolution of ribosome abundance across growth conditions

Figure 7.1: Translation apparatus activity, ribosome content and translation error rates in response to nutritional perturbations. Profiles of GTP, (p)ppGpp and ribosome abundances across growth conditions were found in references from section 2 as well as in the present work. The relative charging level of the limiting tRNA was inferred from the translation frameshift error rate measured in the present work.

Part IV

Material and Methods

Chapter 8

Material and Methods

8.1 Molecular biology techniques

8.1.1 Polymerase Chain Reaction

The amplification of DNA fragments were achieved through different successive cycles of synthesis in a total volume of 50 μL Polymerase Chain Reaction (PCR) mix. Different DNA polymerases were used to amplify DNA fragments and were chosen according to different criteria. The Phusion^{TM}DNA polymerase (FINNZYMES®), the Q5^{TM}DNA polymerase (FINNZYMES®), the Thermo Scientific^{TM}Extensor Long PCR enzyme and the Invitrogen^{TM}Platinum^{TM}SuperFi^{TM}polymerase were used to amplify fragments for building genetic constructs for which a very accurate replication is required. The DyNAzyme^{TM}(FINNZYMES®) and the Thermo Scientific^{TM}DreamTaq polymerase were used to perform colony PCR for screening purposes. They were used according to the supplier recommendations.

8.1.2 DNA purification

The DNA fragments obtained by PCR or digestion were purified using the PCR Wizard SV Gel and PCR Clean-Up system (PROMEGA) according to the supplier recommendations.

8.1.3 DNA strand digestion

The DNA strands were digested by restriction enzymes following the instructions from the supplier.

8.1.4 DNA strand ligation

The DNA fragments ligation were done by the T4 DNA ligase (INVITROGEN) following the instructions from the supplier.

8.1.5 Ligation-Independent Cloning

Oligonucleotides used to amplify the PCR fragment to be inserted in the pBaSysBioII (pBSBII) vector were constructed in order to contain the extensions with a cleavage site for the *Sma*I restriction enzyme,

and the sequences CCGCGGGCTTTCCCAGC (forward strand) and GTTCCTCCTTCCCACC (reverse strand) which are compatible with the insertion into the pBSBII by the Ligation-Independent Cloning (LIC) method as previously described (figure 8.1) [Botella et al., 2010]. Then, the pBSBII plasmid was linearized by digestion with the restriction enzyme *Sma*I, then treated with T4 polymerase whose exonuclease activity 3'→5' creates the cohesive extremities in the presence of dATP. The same treatment was applied to the PCR fragment to be inserted in the pBSBII but in presence of dTTP. The two products were then hybridized at room temperature with each other to form a new pBSBII plasmid with the PCR fragment inserted upstream of the *gene X* sequence. The obtained plasmid is sufficiently stable for direct transformation into competent *E. coli* Top10 strains. The extracted product will then be inserted into *B. subtilis* chromosome by simple cross-over.

8.1.6 Gibson assembly method

This method was reported by Daniel Gibson [Gibson et al., 2009]. It consists in assembling multiple DNA fragments which contain ≈20-40 base pair overlap with adjacent DNA fragments (figure 8.2). They are mixed with a cocktail of three enzymes which possess exonuclease, DNA polymerase, and DNA ligase activities, along with other buffer components. The exonuclease chews back DNA from the 5' end. The resulting single-stranded regions on adjacent DNA fragments can anneal. The DNA polymerase incorporates nucleotides to fill in any gaps. The DNA ligase covalently joins the DNA of adjacent segments, thereby removing nicks in the DNA. The entire mixture is incubated at 50°C for up to one hour. The resulting product consists in different DNA fragments joined into one which is then transformed into competent *E. coli* DH5α strains. The NEB Gibson Assembly®Cloning kit was used following the supplier's recommendations. The NEBuilder®HiFi DNA Assembly Master Mix, which is also based on the Gibson Assembly method, was also used following the supplier's recommendations.

8.2 Methodologies specific to *E. coli* strains

8.2.1 Transformation

The transformations of *E. coli* were performed by heat shock on the TOP10 or DH5α chimiocompetent strains. Cells were mixed with the plasmid of interest on ice during 10 minutes, then transferred at 37°C for 1 minute and back on ice for 10 minutes. LB was added on the mix and prior to an incubation at 37°C for one hour. Then, cells were spread on petri dishes containing LB and ampicillin (100 μg/mL) and incubated overnight at 37°C. The pBSBII derived plasmids were transformed into TG1 *E. coli* strain in order to ensure the success of the transformation in *B. subtilis* when using the MGI and MGII transformation protocol (see part 8.3.1). Indeed, TG1 creates multimeric plasmids that enhance the recombination which ensures that the construct structure is preserved within *B. subtilis*.

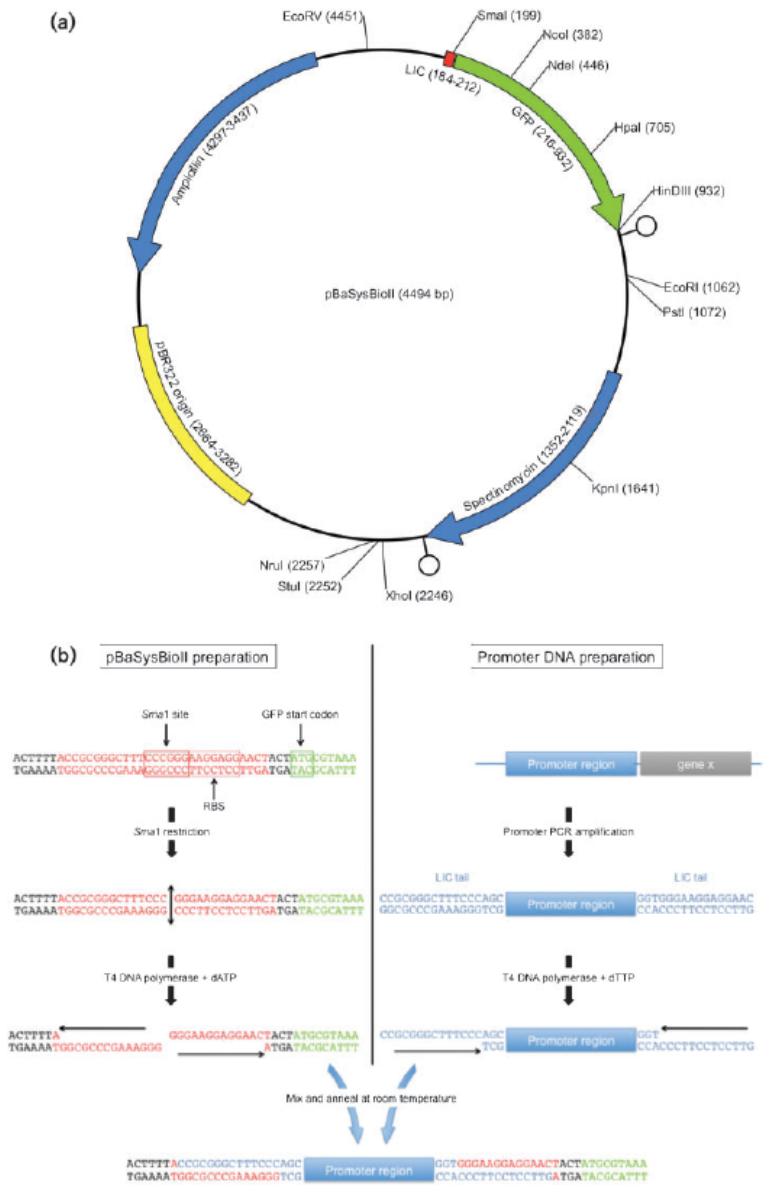


Figure 8.1: Schematic diagram of (a) pBaSysBioII and (b) the LIC system (from [Botella et al., 2010]). a) pBaSysBioII plasmid ORFs are indicated by thick colored arrows, with the direction of transcription indicated. The identity of each ORF is indicated adjacent to each box, followed by its plasmid coordinates. Transcriptional terminators are indicated by lollipops. Relevant restriction sites are also indicated. (b) Cleavage of pBaSysBioII with *Sma*I and treatment with the T4 DNA polymerase in the presence of dATP generates linearized vector DNA with 14- and 13-base 5'overhangs (red). Promoter fragments suitable for high-throughput cloning are generated by amplification of chromosomal DNA using forward and reverse primers with a LIC tail (blue) containing sequences complementary to the LIC sequences of the pBaSysBioII plasmid. Treatment of the PCR-amplified promoter fragments with T4 DNA polymerase in the presence of dTTP generates 5' single-stranded DNA overhangs that are perfectly complementary to those of the vector. Annealing of the treated vector and PCR fragments at room temperature produces a circular duplex species with staggered nicks on the two strands. The duplex is sufficiently stable for direct introduction into competent *E. coli* strains.

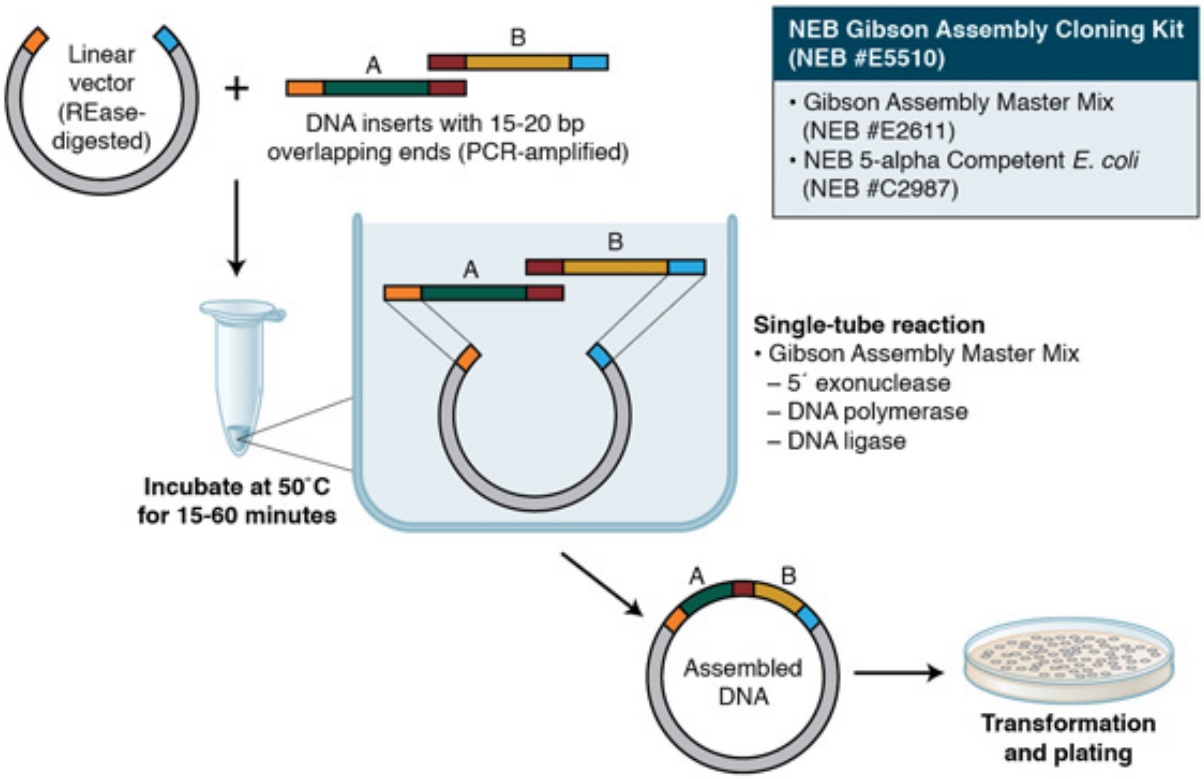


Figure 8.2: The Gibson assembly method (from NEB Gibson Assembly®Cloning kit protocol manual). Gibson Assembly employs three enzymatic activities in a single-tube reaction: 5' exonuclease, the 3' extension activity of a DNA polymerase and DNA ligase activity. The 5' exonuclease activity chews back the 5' end sequences and exposes the complementary sequence for annealing. The polymerase activity then fills in the gaps on the annealed regions. A DNA ligase then seals the nick and covalently links the DNA fragments together.

8.2.2 Extraction of plasmids

The plasmids were extracted from *E. coli* TOP10, DH5α or TG1 strains with the Wizard Plus SV Minipreps DNA Purification System (PROMEGA) kit which uses the alkaline lysis principle [Bimboim and Doly, 1979]. The extraction was performed following the supplier's instructions.

8.3 Methodologies specific to *B. subtilis* strains

8.3.1 Transformation

The transformation of the *B. subtilis* BSB168 strain [Buescher et al., 2012] makes use of natural competence at the beginning of the stationary phase. Cells become competent when grown on a minimal medium MGI malate followed by a dilution in a deficient medium MGII malate (see section 8.4.1). *B. subtilis*, at the beginning of the stationary phase, allows the entrance of the DNA found in the medium. This exogenous DNA can be used as a nutritional source or bring genetic information to the bacterium (acquisition of plasmids or insertion of DNA fragments into bacterial DNA). A one-step method [Harwood and Cutting, 1990] based on the same principle as just described but using MC medium (see section 8.4.1) was also used to transform *B. subtilis* derived strains. Transformants were then selected on LB plates supplemented with the required antibiotic.

8.3.2 Chromosomal DNA extraction

The genomic DNA of *B. subtilis* was extracted using the GenElute Bacterial Genomic DNA (Sigma-Aldrich) kit. The extraction was performed following the supplier's instructions.

8.3.3 Pop-in/Pop-out marker-less deletion technique

The pop-in/pop-out system [Fabret et al., 2002][Tanaka et al., 2012] allows marker-less deletions in the genome of *B. subtilis*. The deletion system is composed of the master strain in which all the deletions were introduced and of a cassette allowing the positive selection of deletions and the eviction of the markers (Figure 8.3). The master strain was derived from the BSB168 strain and obtained by replacing its *upp* gene with a neomycin-resistance gene under the control of the Lambda Pr promoter (λPr-*neo*) to give the master strain BSB168 λPr-*neo*::Δ*upp* (Table strain REF). All deletions were introduced in the master strain by homologous replacement of the targeted chromosome region by a DNA fragment called 'cassette *upp-phleo-cI*', carrying the phleomycin-resistance gene for positive selection of cassette integration and both the *upp* and P_{sak}-λcI genes for counterselection and cassette eviction (Figure 8.4). The transformations are made using the one-step protocol described in part 8.3.1. The transformants obtained after chromosomal insertion of the 'cassette *upp-phleo-cI*' are selected on LB plate containing phlemoycin and single-colonies are then streaked on a fresh plate. Colony PCR are performed to confirm the DNA region deletion and the correct mutants are selected to proceed to the pop-out step. It consists in first inoculating overnight the bacteria in LB cultures without antibiotics, and they are then spread on 15 μg/ml Neomycin-LB plates overnight. Then the

plate obtained with grown colonies is replicated using replica plating equipment onto both 4 μg/ml Phleomycin and 15 μg/ml Neomycin-LB plates overnight. The colonies which possess the desired genotype are sensitive to phleomycin and resistant to neomycin. The positive clones responding to this criteria are plated onto both 4 μg/ml Phleomycin and 15 μg/ml Neomycin-LB plates overnight, and then they are checked by colony PCR and sequencing.

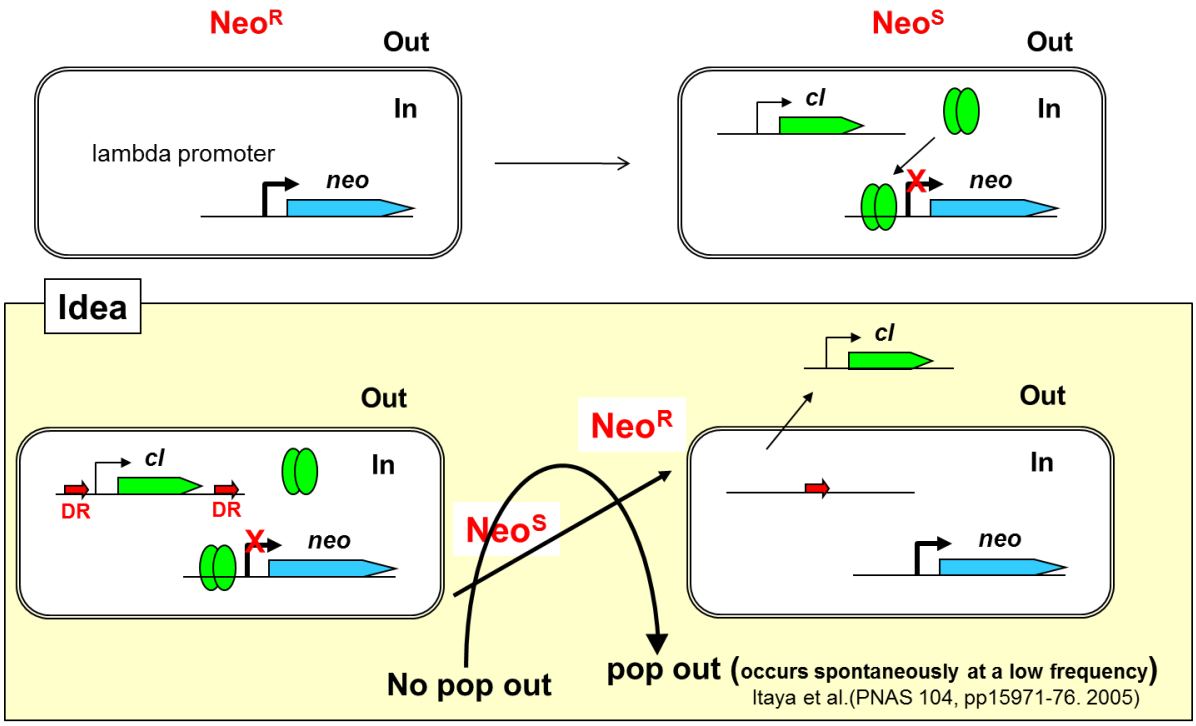


Figure 8.3: Principle of the marker-less deletion system. The master the strain is neomycin resistant (Neo^R) but when a chromomal region has been deleted by insertion of the DNA fragment called 'cassette *upp-phleo-cI*', the protein encoded by *cI* represses the lambda promoter regulating the gene conferring resistance to neomycin. Consequently, the strain becomes sensitive to neomycin (Neo^S) and resistant to phleomycin (Phleo^R). The eviction of the 'cassette *upp-phleo-cI*' spontaneously takes place at low frequency which makes the strain resistant to neomycin and sensitive to phleomycin. This ensures that the markerless deletion at the desired DNA region has occurred.

8.4 Growth medium and bacterial strains

8.4.1 Growth conditions

The *E. coli* strains were only grown in LB medium and *B. subtilis* was grown in LB for transformations and for (pre)-pre-cultures during the LCA experiments. For other experiments, *B. subtilis* was grown in CH (which contained 10% w/v casamino acid) [Partridge and Errington, 1993] or modified M9 medium complemented with different carbon sources and/or amino acids. The modified M9 minimal medium [Harwood and Cutting, 1990] consisted of the following components (per liter): 8.5 g $Na_2HPO_4 \bullet 2H_2O$, 3.0 g KH_2PO_4, 1.0 g NH_4Cl, 0.5 g NaCl. The following components were sterilized separately and

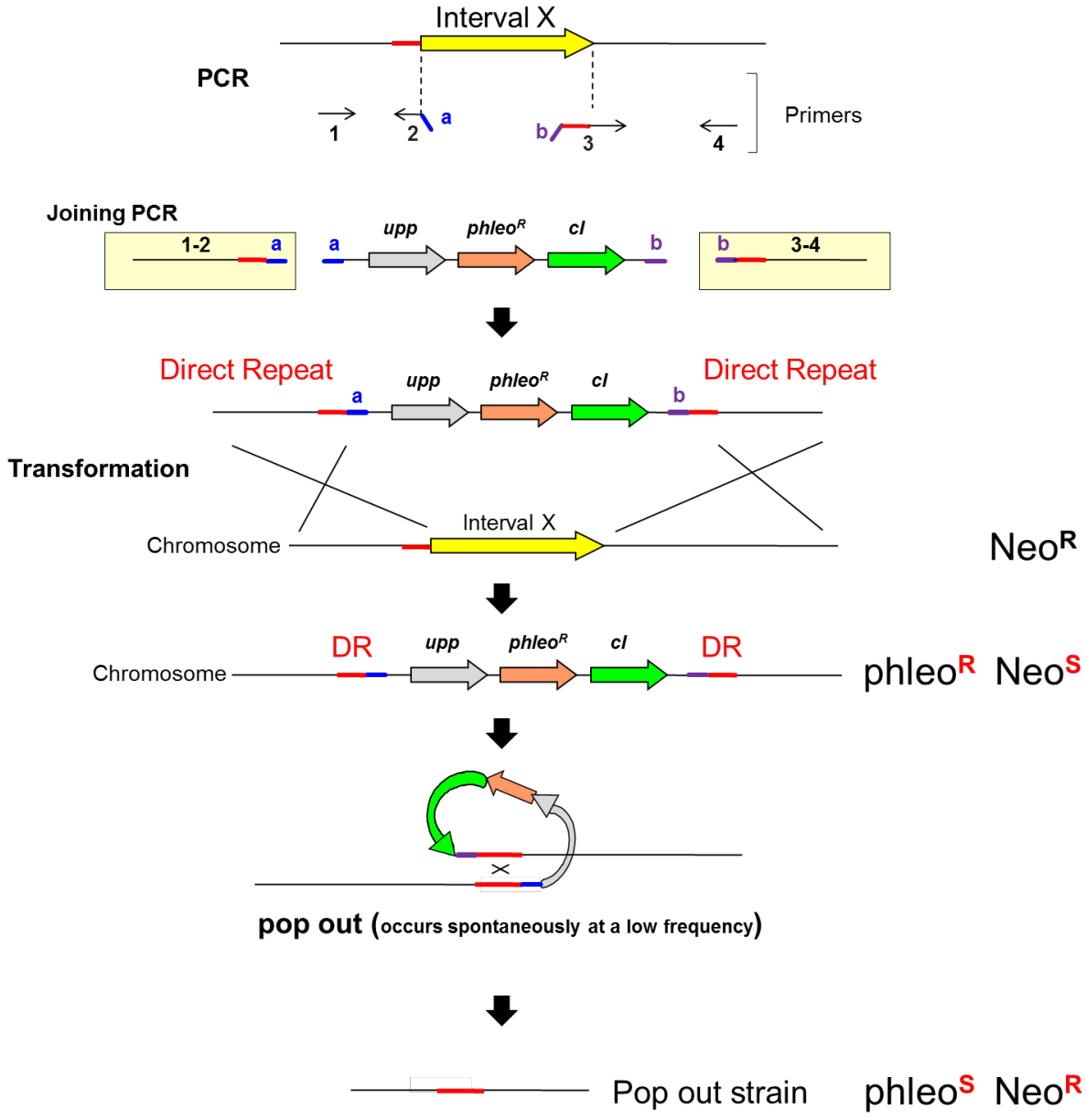


Figure 8.4: Design of the cassette *upp-phleo-cI*'. Primers used to amplify the 'cassette *upp-phleo-cI*' correspond to the a and b overlaps. Primer sets P1-P2 and P3-P4 are designed to amplify ≈1.5 kb upstream and downstream of the deletion target respectively. This is not an absolute size requirement of the product, smaller portions are possible (minimum of 500bp). The DNA strands amplified with these primers pairs are joined by the a and b overlaps region (P2 5' end = GACCTGCAGGCATGCAAGCT and P3 5' end = CGAGCTCGAATTCACTGGCCGTCG) using either joining PCR reaction or Gibson assembly method followed by a PCR on the obtained product with primers P1 and P4. The fragment thus obtained is chromosomally integrated into the selected *B. subtilis* strain by double cross-over leading to a Phleo^{R} Neo^{S} strain. The eviction of the marker is critical for obtaining a marker-less deletion. This occurs spontaneously at a low frequency when a DR is present on both sides of the marker cassette in the absence of antibiotic. This DR can be created by adding the last 30 bp of the P1-P2 product to the 5' end of the binding region of the P3 primer.

then added (per liter of final medium): 1 ml 0.1 M $CaCl_2 \bullet 2H_2O$, 1 ml 1 M $MgSO_4 \bullet 7H_2O$, 1 ml 50 mM $FeCl_3 \bullet 6H_2O$ and 10 ml trace salts solution. The trace salts solution contained (per liter): 170.0 mg $ZnCl_2$, 100 mg $MnCl_2 \bullet 4H_2O$, 60.0 mg $CoCl_2 \bullet 6H_2O$, 60.0 mg $Na_2MoO_4 \bullet 2H_2O$ and 43.0 mg $CuCl_2 \bullet 2H_2O$. The modified M9 was then supplemented with carbon sources at concentrations of 3 g/L (in M9G and M9IMLVMalGlc) or 5 g/L (in CH) for glucose, 4 g/L (in M9M) or 5g/L (in M9IMLVMalGlc) for malate, at 8 g/L for pyruvate, 5g/L for both glutamate and succinate; or with amino acids at 0.025 g/L for isoleucine, at 0.05 g/L for leucine, at 0.04 g/L for valine, and 0.02 g/L for methionine.

For *B. subtilis* transformation, the MGI medium is made of $(NH_4)_2SO_4$ (2 g/L), $C_6H_5Na_3O_7$ (1 g/L), $K_2HPO_4 \bullet 3H_2O$ (14 g/L), KH_2PO_4 (6 g/L), glucose (5 g/L), $MgSO_4 \bullet 7H_2O$ (40 mg/L), casein hydrolyzate (2.50 mg/L) and yeast extract (10 mg/L); the MGII medium is made of $(NH_4)_2SO_4$ (2 g/L), $C_6H_5Na_3O_7$ (1 g/L), $K_2HPO_4 \bullet 3H_2O$ (14 g/L), KH_2PO_4 (6 g/L), glucose (5 g/L), $MgSO_4 \bullet 7H_2O$ (100 mg/L), casein hydrolyzate (1.25 mg/L), yeast extract (2.50 mg/L) and $Ca(NO_3)_2$ (0.75 mg/L); and the MC completed medium is made of $K_2HPO_4 \bullet 3H_2O$ (14 g/L), KH_2PO_4 (5.2 g/L), glucose (20 g/L), $C_6H_5Na_3O_7$ (8.8 g/L), $C_6H_5O_7FeNH_4$ (2.2 g/L), casein hydrolyzate (1 g/L), $C_5H_8KNO_4 \bullet H_2O$ (2 g/L) and $MgSO_4$ (3.35 mM).

When required, media were supplemented with antibiotics at the indicated concentrations for transformants selection: for *E. coli*, ampicillin (100 μg/mL); for *B. subtilis*, spectinomycin (100 μg/mL), neomycin (15 μg/mL), phleomycin (8 μg/mL), tetracyclin (7.5 μg/mL), chloramphenicol (5 μg/mL). Otherwise, the linezolid, chloramphenicol and erythromycin antibiotics were injected to the growth medium at the final concentration indicated in part II.

8.5 Bacterial strains

The *B. subtilis* strains used were derived from the BSB168 strain which is a *trp+* derivative of *B. subtilis* 168 [Nicolas et al., 2012, Buescher et al., 2012].

8.5.1 RelA$^+$ and (p)ppGpp^0 strains construction

To obtain the RelA$^+$ and (p)ppGpp^0 strains, the deletions of the *relP*, *relQ* and *relA* genes in the BSB168 strains and its derivatives were carried out according the pop-in/pop-out system described in section 8.3.3 using the appropriate primers listed in table 8.9. The master strain CLB020 required to achieve this marker-less deletion technique was obtained by transforming the BSB168 strain with the genomic DNA extraction of the CB319 strain.

8.5.2 Inducible *guaB* expression

To control *guaB* expression, we first built a plasmid which contained the *guaB* gene under the control of the $P_{hyperspank}$ promoter: we amplified the *guaB* gene with appropriate primers (see table 8.9) and also amplified by reverse PCR the plasmid pOB1 [Borkowski et al., 2016] which contains the

P_{hs} promoter; and then, after a purification step, we assembled the two fragments using the Gibson assembly method[Gibson et al., 2009]. The resulting plasmid was then amplified by appropriate primers (see table 8.9) to obtain the DNA sequence which contained the *guaB* gene fused to the P_{hs} promoter. We also amplified the *lacI* gene from the plasmid pDR111 using appropriate primers (see table 8.9). The plasmid pAH328 [De San Eustaquio-Campillo et al., 2017] was digested by *EcoR*I/*BamH*I to obtain a vector which contained the *sacA* locus where the resulting construct will be inserted by double crossover. The vector and the two inserts were then digested by the *dpn*I enzyme and purified to be assembled together using the HiFi DNA assembly protocol based on the Gibson assembly method [Gibson et al., 2009]. The resulting product is the plasmid pCLB22 (Figure 8.5) which was then transformed into the derivative strains of *B. subtilis* as specified in table S1. This step was followed by the deletion of the *guaB* gene at its locus by sequence replacement of the *tetL* gene which confers tetracycline resistance to the cell transformed with this fragment. To obtain it, the *tetL* gene open reading frame (ORF) was amplified as well as the sequences (≈1500 bp) upstream and downstream of the *guaB* ORF using the appropriate primers (see table 8.9), and then assembled following the HiFi DNA assembly protocol. The resulting DNA fragment was then amplified by the appropriate primers (see table 8.9) and then selected strains were transformed with the resulting PCR product purified.

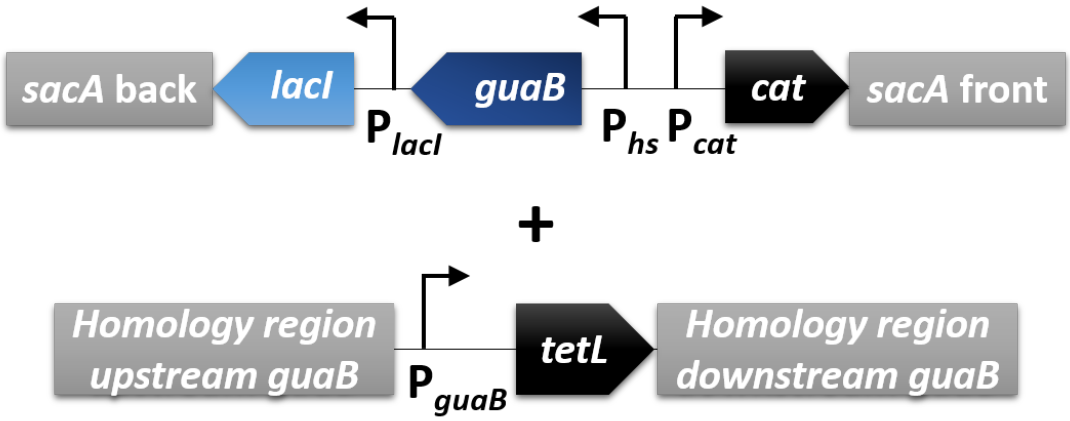


Figure 8.5: Genetic constructs to obtain the *guaB* inducible expression system. The first construct consists in puting the guaB ORF (in dark blue) under the control of the P_{hs} promoter followed by the *lacI* gene (in light blue). Upstream is the gene conferring the strain the resistance to chloramphenicol (in black) which has been inserted with an opposite direction for transcription in order to avoid interferences between the expression of the cat gene and the *guaB* gene. The overall genetic construct is inserted at the *sacA* locus (*sacA* gene in grey) by double cross-over. The fragment allowing the deletion of the *guaB* gene consists in flanking the ORF of the *tetL* gene (in black) conferring the strain the resistance to tetracycline by two homology regions (in grey); and each of them corresponds respectively to the 1500 base pairs upstream and downstream of the *guaB* ORF. The *guaB* promoter is thus included in the upstream homology region.

8.5.3 Promoter reporter fusions

8.5.3.1 The translational error reporter system

The translational error reporter system was obtained by modifying the superfolder GFP (sfGFP) [Overkamp et al., 2013] sequence which is under the constitutive promoter P_{veg} and contained in the plasmid pDG1730 (plasmid SG13, [Guiziou et al., 2016]) (Figure 8.6). To generate sfGFP with frameshift prone sequences, we amplified by inverse PCR the entire SG13 plasmid using primers introducing point mutations at the beginning of the original GFP ORF sequence. The resulting plasmids were extracted from *E. coli* DH5α strains and used to transform the selected strains by double crossover at the *amyE* locus.

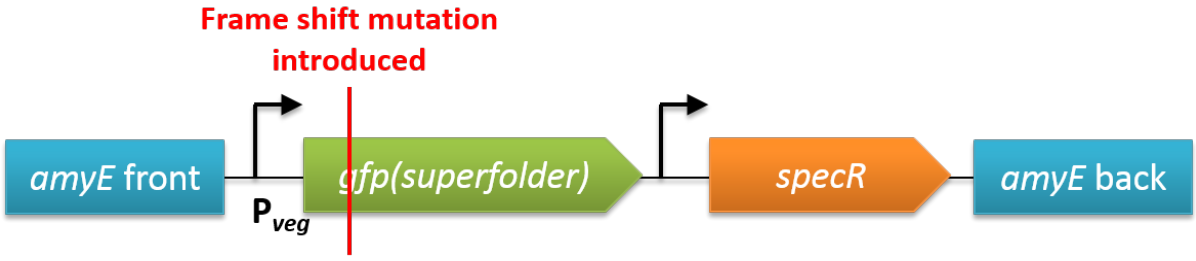


Figure 8.6: Genetic constructs to obtain the translational error reporter system. The translational error reporter system consists in having a GFP superfolder ORF (green) that have undergo modifications at the beginning (Frameshift mutation introduced in red) under the control of the constitutive promoter P_{veg}. This sequence is followed by the gene *specR* (orange) conferring resistance to spectinomycin to the strain. The overall genetic construct is inserted at the *amyE* locus (*amyE* gene in light blue) by double cross-over.

8.5.3.2 *rrns* promoter fusion

The strains carrying the *rrnJ* or *rrnO* promoter were constructed by assembling a reverse PCR product of the SG13 plasmid (where the P_{veg} sequence was removed) with the amplified sequence of either the *rrnJ* or *rrnO* promoter using the Gibson assembly method. The promoter region amplified starts right after the gene sequence upstream of the *rrn* sequence and ends ten base pairs after the TSS of the second promoter of the *rrn*, which thus comprise both promoters (Figure 8.7). The *rrn* promoter region does not end just before the *rrns*' sequence given that upstream of it there exists a site where the primary transcript is cleaved during the posttranscriptional processing events that generate mature 16S, 23S, and 5S rRNAs [Natori et al., 2009]. After purifying both PCR products, the vector and the insert (*rrnJ* or *rrnO* promoters) were assembled using the HiFI DNA assembly protocol and then transformed into *E. coli* to obtain the desired plasmids. The cloned and fused sequences were then chromosomally inserted into chosen *B. subtilis* strains as described in table 8.9.

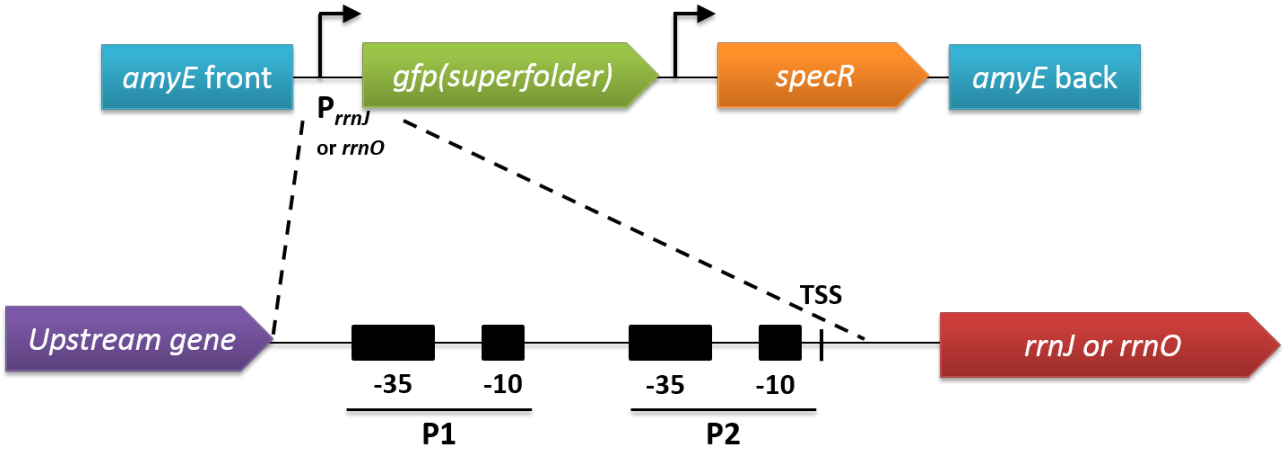


Figure 8.7: The *rrn* promoter fusion. The system reporting the transcription level of *rrnJ* and *rrnO* consists in fusing the gfp superfolder ORF (in green) to the promoter region of one of the *rrns*. The promoter region (in black) starts right after the gene sequence upstream of the *rrn* sequences (in purple) and ends ten base pairs after the TSS of the second promoter of the *rrn* (in red), which thus comprises both promoters. The fusion sequence is followed by the *specR* gene (orange) conferring resistance to spectinomycin to the strain. The overall genetic construct is inserted at the *amyE* locus (*amyE* gene in light blue) by double cross-over.

8.5.3.3 The TSS reporter system

To report the transcription level of genes according to their TSS, a plasmid carrying the *fbaA* promoter (constitutive) and its TIR fused to a *gfpmut3* sequence was built using the LIC method (described in section 8.1.5) and the original pBSBII plasmid. The region to be inserted was obtained by amplifying the upstream sequences of the *fbaA* gene ($\approx$ 1000 bp) using the appropriate primers (see table 8.9). The LIC product was purified and then mixed with TOP10 *E. coli* bacteria for transformation to obtain the newly formed plasmid pCLB49. The native TSS of *fbaA* is composed of two guanines (GG). To replace it by two adenines (AA), a reverse PCR was performed on the plasmid pCLB49 using primers which contained the exact same sequences of the TIR and its upstream region except for the two first nucleotides of the TIR which were replaced by adenines (figure 8.8). Then the PCR product was purified and transformed with TOP10 *E. coli* strains to form the pCLB51 plasmid. The pCLB49 and pCLB51 (Figure 8.9) were then transformed into chosen *B. subtilis* strains as described in table 8.9.

8.6 Live-Cell Array

The Live Cell Array is a technique which allows the high throughput study of gene expression. It measures both the optical density and the fluorescence emission of the reporter protein contained in the different strains of interest (figure 8.10).

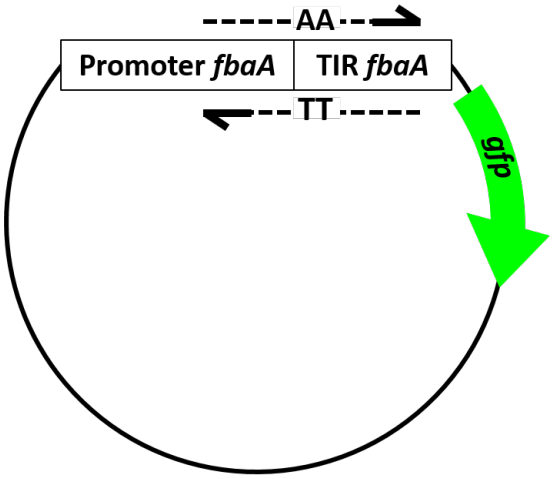


Figure 8.8: Replacement of the two first nucleotides of the TIR region. To obtain plasmid pCLB51, we used a forward primer with the exact same sequences of the *fbaA* TIR and its upstream region except for the two first nucleotides of the TIR which were replaced by adenines (AA) while two thymines (TT) replaced the two cytosines for the reverse primer.

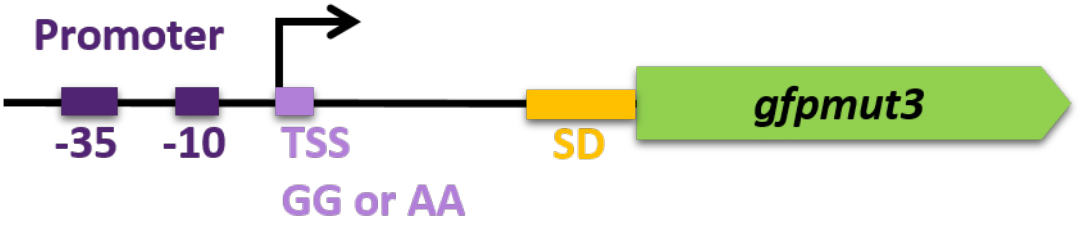


Figure 8.9: The TSS reporter system. The TSS reporter system consists in fusing the *gfpmut3* ORF (in green) to the *fbaA* promoter region (dark purple) followed by its native TIR or modified native TIR where two adenines replace the two guanines as TSS (in light purple) corresponding to the sequence starting at the TSS and ending just before the *gfpmut3* ORF. The TIR includes the Shine-Dalgarno (SD) sequence in yellow.

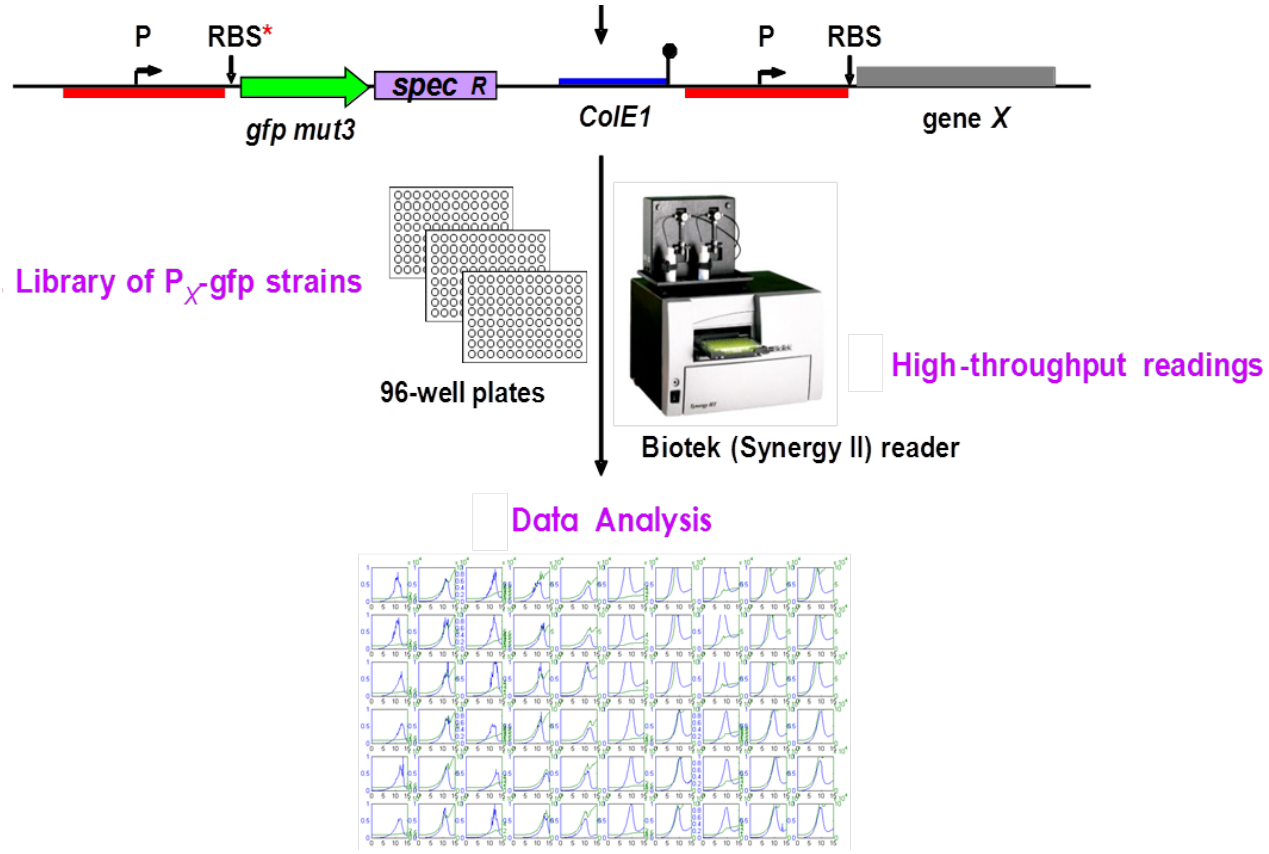


Figure 8.10: Live Cell Array. *gfp*-reporter strains are grown in 96-well microtiterplates in the Biotek (Syngery II) which measures the OD as well as the fluorescence levels at different time intervals generating important amounts of data.

8.6.1 *B. subtilis* culture and data acquisition

The *B. subtilis* strains were grown in 96-well microtiterplate (CELLSTAR®, Greiner bio-one) on LB medium overnight and then diluted 20-fold in fresh LB medium. The LB pre-cultures were grown until an OD_{600} of 0.3-0.4 was reached and then diluted 20-fold into the medium of interest. The cells grown in this pre-culture were diluted 20-fold into the final culture with the medium of interest. For each step, the cultures were inoculated into microtiterplates and incubated under constant shaking at 37°C in a Synergy^{TM}2 multimode microtiterplate reader (BioTek®). OD_{600} and fluorescence (excitation 485/20 nm, emission 528/20 nm) were measured at time intervals of 7 minutes. Only the 60 wells located in the center of the microtiter plates were inoculated since the wells around the edges exhibited evaporation higher than 5 % over 20 hours and were only filled in with sterile medium (figure 8.11).

	Fluorescein 10nM	Fluorescein 1nM									
Blank	BSB168	strain b	strain d	strain f	strain h	strain j	strain l	strain n	strain p	strain r	
Blank	BSB168	strain b	strain d	strain f	strain h	strain j	strain l	strain n	strain p	strain r	
Blank	BSB168	strain b	strain d	strain f	strain h	strain j	strain l	strain n	strain p	strain r	
Blank	strain a	strain c	strain e	strain g	stran i	strain k	strain m	strain o	strain q	BSB168	
Blank	strain a	strain c	strain e	strain g	stran i	strain k	strain m	strain o	strain q	BSB168	
Blank	strain a	strain c	strain e	strain g	stran i	strain k	strain m	strain o	strain q	BSB168	
									Fluorescein 1nM	Fluorescein 10nM	

Figure 8.11: Exemple of an LCA 96-well microtiterplate design. The different strains were cultivated on a 96-well microtiterplates with at least three replicates each time and were grown in a defined sterile medium. The BSB168 are the wild type strains used to correct the intrinsic fluorescence level of *B. subtilis* while the Blank is used to correct the OD and fluorescence measurements.

8.6.2 Data treatment

The measured fluorescence per well at any time ($RawFluo$) corresponds to the sum of the fluorescence signals from the OD_{600}-dependent GFP produced per well (GFP^{OD}) but also reflects the OD_{600}-dependent *B. subtilis* and medium auto-fluorescences ($AutoFluo^{OD}$), the time-dependent intra-day variability ($Drift^t$) and the day-to-day variability ($D2DFluo^d$) as follows:

$$RawFluo = (GFP^{OD} + AutoFluo^{OD}) \times Drift^t \times D2DFluo^d$$

In order to extract GFP^{OD} from the LCA data, we set up a specific design for the microtiterplate structure that sequentially allowed removing unwanted fluorescence.

8.6.3 Blank wells for the intra-day variability, $Drift^t$

A drift of the measured fluorescence intensity was observed with time. A set of 6 external wells, only filled in by sterile medium, were used to estimate and compensate for the observed very slow drift of the excitation and receptor devices (figure 8.11).

8.6.4 Fluorescein for the plate-to-plate normalization, $D2DFluo^d$

In order to correct possible day-to-day variability between LCA experiments, we systematically included in each microtiterplates 4 wells filled in with different concentrations of fluorescein (twice 1nM and 10nM).

8.6.5 Summary of a typical microtiterplate design

Following all these previous constraints, the 96-wells microtiterplate contained:

- 4 wells around the edges that were filled in by fluorescein (at the top left and bottom right of the microtiterplate) (figure 8.11)
- 6 wells on the right and left edges that were used as blank (figure 8.11)
- Each culture was performed in at least six technical replicates by two biological replicates (more than twelve values) with always BSB168 strain or derivatives which do not possess any fluorescent reporter protein in order to measure the autofluorescence level of the studied strain.

8.7 Estimates of growth rate and fluorescence levels from LCA data

8.7.1 Growth rate calculation

In order to estimate both the growth rate and the GFP abundance (GFP per well divided by OD at each time point), we subtracted in each well the first measured value of OD_{600} (time 0) to all OD_{600} measurements. The corrected OD_{600} time-series in logarithmic scale were fitted by a linear model, which directly provided the corresponding growth rates (its slopes). The bootstrap procedure [Efron, 1987] was used to obtain the 95% confidence intervals for the growth rate values.

8.7.2 Estimate of fluorescence levels in GFP reporter strains used to detect translational errors

8.7.2.1 Intra-day compensation, $Drift$

We estimated the drift using a loess regression as function of time, and we obtained the desired compensation after a suitable normalization, α_0 which is such that:

$$\alpha_0 \int_0^{tf} Drift(t)\,\mathrm{d}t = 1$$

Eventually, RawFluo data were corrected at each time point as follows:

$$Fluorescence\_corrected(t) = \frac{Fluorescence(t)}{\alpha_0 Drift(t)}$$

Nota Bene: in the sequel, we will use the term "$Fluorescence$" in place of "$Fluorescence\_corrected$".

8.7.2.2 Normalization by fluorescein, $D2D_Fluo^d$

Fluorescein is a light-sensitive molecule that is subject to photobleaching. Therefore, we normalized fluorescence levels of the entire dataset using the mean value of the fluorescein-related fluorescence obtained during 5 measurements (from 35 minutes to 70 min).

8.7.2.3 Fitting of the OD and fluorescence level as function of time

In order to realign the growth and fluorescence curves between the different wells of the microtiterplate, we fitted the OD_{600} and fluorescence as function of time by a linear approximation between each time point (measurements every 7 minutes):

$$OD(t) = a_n t + b_n \qquad t \in [n \times 7min, (n+1) \times 7min[$$

$$Fluorescence(t) = c_n t + d_n \qquad t \in [n \times 7min, (n+1) \times 7min[$$

where a_n, b_n, c_n and d_n are variables that change at each time interval ($[n \times 7min, (n+1) \times 7min[$). After fitting the OD_{600} and fluorescence data, we resampled them every 10 seconds in order to superpose the growth curves such that they reach a certain OD_{600} value at a fixed time point. The fluorescence level corresponding to this specific OD_{600} value was thus also found at this time point.

8.7.2.4 Autofluorescence correction

In order to estimate GFP fluorescence, we subtracted an auto-fluorescence function from the fluorescence datasets. We defined the auto-fluorescence function $AutoFluo(t)$ as the fluorescence fitting function defined in the previous part and applied to the fluorescence data of the wells where the strains without GFP reporters were grown (either BSB168, CLB038, CLB028, CLB240 or CLB247). Thus, the $Fluorescence$ function of the strains carrying GFP reporter systems was defined as follows:

$$GFP(t) = Fluorescence(t) - AutoFluo(t)$$

8.7.2.5 Bootstrap

The bootstrap procedure [Efron, 1987] was used to obtain the 95% confidence intervals for the fluorescence level and thus the translational frameshift error (TFE) rate. Once the fluorescence has been treated as previously described, wells corresponding to the replicates for each strain genotype were randomly chosen to calculate the mean fluorescence level and in fine the TFE rate. 100 rounds of random sampling with replacement were performed and provided the distribution of fluorescence and

TFE rate values as if 100 independent experiments have been done. The figures representing the evolution of the TFE rate as function of time are made with boxplots based on the obtained distribution (figure 8.12).

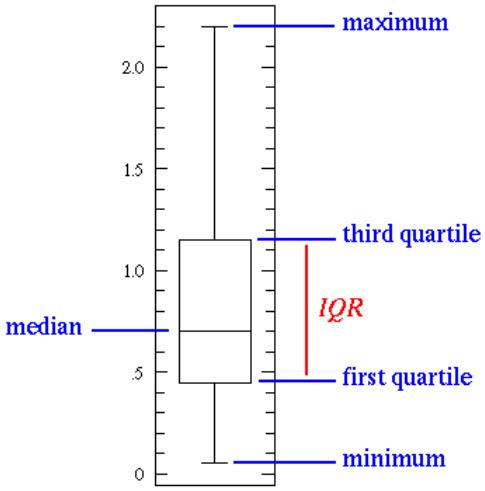


Figure 8.12: Schematic representation of a boxplot. The central rectangle spans the first quartile to the third quartile; a segment inside the rectangle shows the median; and "whiskers" above and below the box show the locations of the minimum and maximum values obtained.

8.7.3 Estimate of fluorescence levels in GFP reporter strains used to report transcription levels

8.7.3.1 Dedicated wells for auto-fluorescence correction, $AutoFluo^{OD}$

As already mentioned in the literature [Aïchaoui et al., 2012, Botella et al., 2010], the raw fluorescence values are the sum of the GFP fluorescence level and the *B. subtilis* and medium-related auto-fluorescence ($AutoFluo^{OD}$). We subtracted autofluorescences from the raw fluorescence values as previously described [Botella et al., 2010, Buescher et al., 2012].

8.7.3.2 Definition of the exponential steady-state regimen in LCA

After systematic removal of non-growing cultures, we determined the "OD_{600} range" in which all clones were considered to be in exponential growth and therefore in steady-state regimen. The range was manually defined based on the best linear fit of the OD_{600} (in logarithmic scale) with respect to the time. This step allowed further removal of data from specific wells in which cells did not grow in steady-state regimen.

8.7.3.3 Autofluorescence correction

In order to estimate GFP fluorescence, we subtracted an auto-fluorescence function from the raw fluorescence datasets. We defined the auto-fluorescence function using a 3^{rd}-degree polynomial applied

to the OD_{600} and fluorescence measurements of the strains reporting the autofluorescence (either BSB168, CLB038, or CLB028):

$$AutoFluo(OD) = b_0 + b_1 OD + b_2 OD^2 + b_3 OD3$$

where OD is used in place of OD_{600}.
The GFP abundance during the exponential phase is calculated for each time point (t_i) as follows:

$$[GFP(t_i)] = \frac{RawFluo(t_i) - AutoFluo(OD(t_i))}{OD(t_i)}$$

The final GFP abundance is given by the median of $[GFP(t_i)]$ with time.

8.8 Definition and calculation of the translational error rate

8.8.1 Expression of protein production

In the sequel $X(t)$ corresponds to the bacterial population density ($\sim$ OD(t)) and $GFP(t)$ is the total amount of fluorescent proteins ($\sim$ fluorescence). By definition, the accumulation of a specific protein is given by:

$$\frac{dGFP}{dt}(t) = X(t) * Protein\ production(t)$$

and thus:

$$Protein\ production(t) = \frac{1}{X(t)} * \frac{dGFP}{dt}(t) \tag{8.1}$$

If we now consider the GFP concentration, we then have by definition:

$$Protein\ concentration : [GFP](t) = \frac{GFP(t)}{X(t)}$$

from which we can deduce that the protein concentration as function of time:

$$\frac{d[GFP](t)}{dt} = \frac{\frac{dGFP}{dt}(t)X(t) - GFP(t) * \frac{dX}{dt}(t)}{X(t)^2}$$

and thus

$$\frac{d[GFP]}{dt}(t) = \frac{\frac{dGFP}{dt}(t)}{X(t)} - \frac{GFP(t) * \frac{dX}{dt}(t)}{X(t)^2}$$

By introducing the growth rate given by $\mu(t) = \frac{\frac{dX}{dt}(t)}{X(t)}$, we thus deduce:

$$\frac{d[GFP]}{dt}(t) = Protein\ production(t) - \mu(t) * \frac{GFP(t)}{X(t)}$$

and thus :

$$\frac{d[GFP](t)}{dt} = Protein\ production(t) - \mu(t) * [GFP](t)$$

$$\boxed{Protein\ production(t) = \frac{d[GFP]}{dt}(t) + \mu(t) * [GFP](t)}$$

During an exponential phase, the concentration and the growth rate are constant and the protein production is consequently given by :

$$\boxed{Protein\ production = \mu * [GFP]} \tag{8.2}$$

8.8.2 Expression of the translational error rate

The instantaneous translational error rate is defined as:

$$\tau(t) = \frac{GFP_{frameshift}\ production(t)}{GFP_{native}\ production(t)}.$$

and relation (8.1) allows us to deduce that :

$$\tau(t) = \frac{X_{native}(t)}{X_{frameshift}(t)} \frac{\frac{dGFP_{frameshift}}{dt}(t)}{\frac{dGFP_{native}}{dt}(t)}$$

By definition, the instantaneous error is well-defined if $X_{native}(t)$ and $X_{frameshift}(t)$ coincide and we then deduce that :

$$\boxed{\tau(t) = \frac{\frac{dGFP_{frameshift}}{dt}(t)}{\frac{dGFP_{native}}{dt}(t)}}.$$

We finally note that during an exponential phase, relation (8.2) allows to deduce that the error is constant and given by:

$$\boxed{\tau = \frac{[GFP]_{frameshift}}{[GFP]_{native}}}$$

8.8.3 Calculation of the translational error rate with experimental data

8.8.3.1 Method 1

As seen in section 8.8.2, during the exponential (or steady-state) phase, the translational ER is defined as:

$$\tau = \frac{[GFP]_{frameshift}}{[GFP]_{native}} = \frac{\frac{Fluo_{GFP_{frameshift}}}{OD}}{\frac{Fluo_{GFP_{native}}}{OD}}$$

$$\boxed{\tau = \frac{Fluo_{GFP_{frameshift}}}{Fluo_{GFP_{native}}}}$$

where $Fluo_{GFP_{native}}$ corresponds to the fluorescence level of the strain carrying a native GFP gene; and $Fluo_{GFP_{frame-shift}}$ corresponds to the fluorescence level of the strain carrying a mutated GFP gene with a nucleotide added (+1 frameshift) or deleted (-1 frameshift) which leads to the production of a complete GFP only if a frameshift event occurs. This definition is commonly used in the literature to report the level of translational errors [Hagervall et al., 1993, Li et al., 1997a, Meyerovich et al., 2010].

8.8.3.2 Method 2

The derivation of the GFP produced can be approximated as the GFP accumulated during a small period of time Δt such that:

$$\frac{dGFP}{dt}(t) = \frac{GFP_{t+1} - GFP_t}{\Delta t}$$

Consequently, the instantaneous translational error rate is calculated as follow:

$$\tau(t) = \frac{\frac{dGFP_{frame-shift}}{dt}(t)}{\frac{dGFP_{native}}{dt}(t)} \approx \frac{\frac{GFP_{frameshift}(t+1) - GFP_{frameshift}(t)}{\Delta t}}{\frac{GFP_{native}(t+1) - GFP_{native}(t)}{\Delta t}}$$

$$\tau(t) \approx \frac{\frac{Fluo_{GFP_{frameshift}}(t+1) - Fluo_{GFP_{frameshift}}(t)}{\Delta t}}{\frac{Fluo_{GFP_{native}}(t+1) - Fluo_{GFP_{native}}(t)}{\Delta t}}$$

$$\tau(t) \approx \frac{Fluo_{GFP_{frameshift}}(t+1) - Fluo_{GFP_{frameshift}}(t)}{Fluo_{GFP_{native}}(t+1) - Fluo_{GFP_{native}}(t)}$$

$$\boxed{\tau(t) \approx \frac{\Delta Fluo_{GFP_{frameshift}}(t)}{\Delta Fluo_{GFP_{native}}(t)}}$$

where $\Delta Fluo_{GFP_{native}}(t)$ is the amount of GFP_{native} accumulated during Δt and $\Delta Fluo_{GFP_{frameshift}}(t)$ is the amount of $GFP_{frameshift}$ accumulated during Δt.

8.8.4 Comparison of method 1 and method 2 during exponential and non-steady state growth

We experimentally measured and treated the fluorescence levels (see section 8.7.2) of the RelA$^+$ GFP^{Tyr}_{ctrl} and RelA$^+$ GFP^{Tyr}_{fs+1} strains grown in M9G; as well as the (p)ppGpp^0 GFP^{Tyr}_{ctrl} and (p)ppGpp^0 GFP^{Tyr}_{fs+1} strains grown in CH. We computed the data to calculate the TFE rate with method 1 and 2 as described in section 8.8.3. The bootstrap procedure [Efron, 1987] was used to obtain the 95% confidence intervals for the TFE rate calculated with either method 1 or method 2. We found that the resulting TFE rates were constant during the steady-state growth of the RelA$^+$ derivative strains grown in M9G (figure 8.13 A). The TFE rate was of ≈0.25% for both methods of calculation. These results validated the theoretical prediction made in part 2.3.2 that the two methods (1 and 2) are equivalent for calculating the TFE rate during steady-state growth. Thus, the redefined computation method (2) of the TFE rate is equivalent to the former computation method (1) when applied to exponentially growing cells but it can also be applied to non-steady state growth. Indeed, figure 8.13 B shows that during the transition to the stationary phase, method 1 did not provide a complete information on the evolution of the (p)ppGpp^0 TFE rate: the TFE rate calculated with method 1 peaked during the transition phase while the TFE rate calculated with method 2 only showed a slight increase.

8.9 List of the strains and primers used in this work

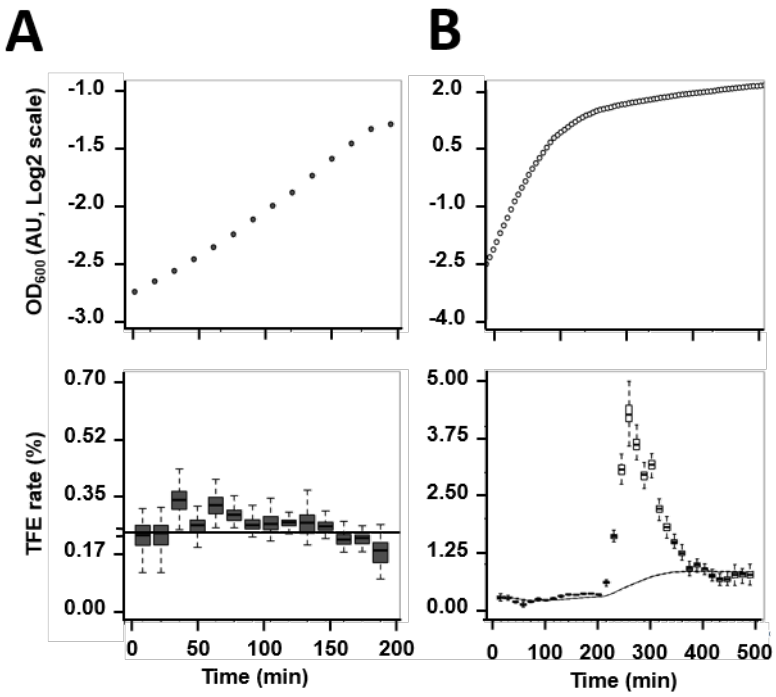


Figure 8.13: Growth and TFE rate calculated with the two computation methods during growth in M9G and CH. A and B. The graph at the top represents the growth curve. The bottom graph represents the TFE rate of the GFP^{Tyr}_{fs+1} reporter system as function of time calculated with method 1 (black line) and with method 2 (boxplots). A. The RelA$^+$ strain was grown in M9G. B. The (p)ppGpp^0 strain was grown in CH. The dashed lines up and down of the black line correspond to the bootstrap standard deviation of the TFE rate calculated with method 1.

Table 8.1: Table of the strains used in this work

Strain	Designation	Relevant genotype	Strains & construction[1]
BSB168	WT	Wild type (prototroph)	[Buescher et al., 2012, Nicolas et al., 2012]
CB319	–	BSB168 Δ*upp*::λ*Pr-neo* Δ*relP*::*upp-phleo-cI*	[Benoist et al., 2015]
CLB020	–	Δ*upp*::λ*Pr-neo*	CB319 → BSB168
CLB034	–	Δ*upp*::λ*Pr-neo* Δ*relP*	CB319 → CLB020 ↑ p/o
CLB038	RelA$^+$	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ*	↓ p/i *relQ* → CLB034 ↑ p/o
CLB028	(p)ppGpp^0	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ* Δ*relA*	↓ p/i *relA* → CLB038 ↑ p/o
CLB080	WT GFP^{Tyr}_{ctrl}	Δ*amyE*::P$_{veg}$*sfGFP*$^{Tyr}_{ctrl}$/*spec*	pSG13 → BSB168
CLB082	WT GFP^{Tyr}_{fs+1}	Δ*amyE*::P$_{veg}$*sfGFP*$^{Tyr}_{fs+1}$/*spec*	CLB09 → BSB168
CLB090	RelA$^+$ GFP^{Tyr}_{ctrl}	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ* Δ*amyE*::P$_{veg}$*sfGFP*$^{Tyr}_{ctrl}$/*spec*	pSG13 → CLB038
CLB092	RelA$^+$ GFP^{Tyr}_{fs+1}	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ* Δ*amyE*::P$_{veg}$*sfGFP*$^{Tyr}_{fs+1}$/*spec*	pCLB09 → CLB038
CLB130	(p)ppGpp^0 GFP^{Tyr}_{ctrl}	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ* Δ*relA* Δ*amyE*::P$_{veg}$*sfGFP*$^{Tyr}_{ctrl}$/*spec*	↓ p/i *relA* → CLB090 ↑ p/o
CLB132	(p)ppGpp^0 GFP^{Tyr}_{fs+1}	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ* Δ*relA* Δ*amyE*::P$_{veg}$*sfGFP*$^{Tyr}_{fs+1}$/*spec*	↓ p/i *relA* → CLB092 ↑ p/o
CLB155	–	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ* Δ*sacA*::P$_{hs}$*guaB/cm* Δ*amyE*::P$_{veg}$*sfGFP*$^{Tyr}_{ctrl}$/*spec*	pCLB22 → CLB090
CLB158	–	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ* Δ*sacA*::P$_{hs}$*guaB/cm* Δ*amyE*::P$_{veg}$*sfGFP*$^{Tyr}_{fs+1}$/*spec*	pCLB22 → CLB092

[1]Horizontal arrows (→) indicate construction by transformation; vertical arrows (↓) indicate that deletions were introduced by homologous replacement of the targeted gene by a linear DNA fragment containing a marker or the evictable cassette *upp-phleo-cI* (p/i for pop in); "↑ p/o" indicates the eviction (pop out) of the cassette *upp-phleo-cI* (see section 8.3.3)

CLB236	–	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta sacA$::$\mathrm{P}_{hs} guaB/cm$	pCLB22 → CLB038
CLB240	RelA$^+$ $guaB^{inducible}$	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta sacA$::$\mathrm{P}_{hs} guaB/cm$ $\Delta guaB$::$tetL$	↓ $guaB$ → CLB236
CLB242	RelA$^+$ GFP^{Tyr}_{ctrl} $guaB^{inducible}$	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta sacA$::$\mathrm{P}_{hs} guaB/cm$ $\Delta guaB$::$tetL$ $\Delta amyE$::$\mathrm{P}_{veg} sfGFP^{Tyr}_{ctrl}/spec$	↓ $guaB$ → CLB155
CLB244	RelA$^+$ GFP^{Tyr}_{fs+1} $guaB^{inducible}$	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta sacA$::$\mathrm{P}_{hs} guaB/cm$ $\Delta guaB$::$tetL$ $\Delta amyE$::$\mathrm{P}_{veg} sfGFP^{Tyr}_{fs+1}/spec$	↓ $guaB$ → CLB158
CLB247	(p)ppGpp^0 $guaB^{inducible}$	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta relA$::upp-$phleo$-cI $\Delta sacA$::$\mathrm{P}_{hs} guaB/cm$ $\Delta guaB$::$tetL$	↓ p/i $relA$ → CLB240
CLB249	(p)ppGpp^0 GFP^{Tyr}_{ctrl} $guaB^{inducible}$	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta relA$::upp-$phleo$-cI $\Delta sacA$::$\mathrm{P}_{hs} guaB/cm$ $\Delta guaB$::$tetL$ $\Delta amyE$::$\mathrm{P}_{veg} sfGFP^{Tyr}_{ctrl}/spec$	↓ p/i $relA$ → CLB242
CLB250	(p)ppGpp^0 GFP^{Tyr}_{fs+1} $guaB^{inducible}$	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta relA$::upp-$phleo$-cI $\Delta sacA$::$\mathrm{P}_{hs} guaB/cm$ $\Delta guaB$::$tetL$ $\Delta amyE$::$\mathrm{P}_{veg} sfGFP^{Tyr}_{fs+1}/spec$	↓ p/i $relA$ → CLB244
CLB254	–	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta amyE$::$\mathrm{P}_{veg} sfGFP^{codon10}_{ctrl}/spec$	pCLB33 → CLB038
CLB256	–	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta amyE$::$\mathrm{P}_{veg} sfGFP^{codon13}_{ctrl}/spec$	pCLB35 → CLB038
CLB258	–	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta amyE$::$\mathrm{P}_{veg} sfGFP^{codon14}_{ctrl}/spec$	pCLB37 → CLB038
CLB260	–	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta amyE$::$\mathrm{P}_{veg} sfGFP^{codon16}_{ctrl}/spec$	pCLB39 → CLB038
CLB264	RelA$^+$ GFP^{Ile}_{fs-1}	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta amyE$::$\mathrm{P}_{veg} sfGFP^{Ile}_{fs-1}/spec$	pCLB43 → CLB038
CLB266	RelA$^+$ GFP^{Leu}_{fs-1}	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta amyE$::$\mathrm{P}_{veg} sfGFP^{Leu}_{fs-1}/spec$	pCLB48 → CLB038
CLB273	(p)ppGpp^0 GFP^{Leu}_{fs-1}	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta relA$::upp-$phleo$-cI $\Delta amyE$::$\mathrm{P}_{veg} sfGFP^{Leu}_{fs-1}/spec$	↓ p/i $relA$ → CLB266
CLB301	(p)ppGpp^0 GFP^{Ile}_{fs-1}	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta relA$::upp-$phleo$-cI $\Delta amyE$::$\mathrm{P}_{veg} sfGFP^{Ile}_{fs-1}/spec$	↓ p/i $relA$ → CLB264
CLB337	RelA$^+$ GFP^{Leu}_{ctrl}	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta amyE$::$\mathrm{P}_{veg} sfGFP^{Leu}_{ctrl}/spec$	pCLB69 → CLB038

CLB339	RelA$^+$ GFP^{Ile}_{ctrl}	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta amyE$::$\mathrm{P}_{veg}sfGFP^{Ile}_{ctrl}/spec$	pCLB71 → CLB038
CLB341	(p)ppGpp^0 GFP^{Leu}_{ctrl}	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta relA$::upp-$phleo$-cI $\Delta amyE$::$\mathrm{P}_{veg}sfGFP^{Leu}_{ctrl}/spec$	↓ p/i $relA$ → CLB337
CLB343	(p)ppGpp^0 GFP^{Ile}_{ctrl}	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta relA$::upp-$phleo$-cI $\Delta amyE$::$\mathrm{P}_{veg}sfGFP^{Ile}_{ctrl}/spec$	↓ p/i $relA$ → CLB339
CLB365	WT GFP^{Ile}_{fs-1}	$\Delta amyE$::$\mathrm{P}_{veg}sfGFP^{Ile}_{fs-1}/spec$	CLB264 → BSB168
CLB366	WT GFP^{Leu}_{fs-1}	$\Delta amyE$::$\mathrm{P}_{veg}sfGFP^{Leu}_{fs-1}/spec$	CLB266 → BSB168
CLB367	WT GFP^{Leu}_{ctrl}	$\Delta amyE$::$\mathrm{P}_{veg}sfGFP^{Leu}_{ctrl}/spec$	CLB337 → BSB168
CLB368	WT GFP^{Ile}_{ctrl}	$\Delta amyE$::$\mathrm{P}_{veg}sfGFP^{Ile}_{ctrl}/spec$	CLB339 → BSB168
CLB004	WT $\mathrm{P}^{+1GG}_{fbaA}gfp$	$fbaA$::$\mathrm{P}^{+1GG}_{fbaA}gfpmut3/spec$	pCLB49 → BSB168
CLB001	WT $\mathrm{P}^{+1AA}_{fbaA}gfp$	$fbaA$::$\mathrm{P}^{+1AA}_{fbaA}gfpmut3/spec$	pCLB51 → BSB168
CLB305	RelA$^+$ $\mathrm{P}^{+1GG}_{fbaA}gfp$	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $fbaA$::$\mathrm{P}^{+1GG}_{fbaA}gfpmut3/spec$	pCLB49 → CLB038
CLB307	RelA$^+$ $\mathrm{P}^{+1AA}_{fbaA}gfp$	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $fbaA$::$\mathrm{P}^{+1AA}_{fbaA}gfpmut3/spec$	pCLB51 → CLB038
CLB309	RelA$^+$ $\mathrm{P}^{+1GG}_{fbaA}gfp$ $guaB^{inducible}$	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta sacA$::$\mathrm{P}_{hs}guaB/cm$ $\Delta guaB$::$tetL$ $fbaA$::$\mathrm{P}^{+1GG}_{fbaA}gfpmut3/spec$	pCLB49 → CLB240
CLB311	RelA$^+$ $\mathrm{P}^{+1AA}_{fbaA}gfp$ $guaB^{inducible}$	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta sacA$::$\mathrm{P}_{hs}guaB/cm$ $\Delta guaB$::$tetL$ $fbaA$::$\mathrm{P}^{+1AA}_{fbaA}gfpmut3/spec$	pCLB51 → CLB240
CLB313	(p)ppGpp^0 $\mathrm{P}^{+1GG}_{fbaA}gfp$	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta relA$::upp-$phleo$-cI $fbaA$::$\mathrm{P}^{+1GG}_{fbaA}gfpmut3/spec$	↓ p/i $relA$ → CLB305
CLB315	(p)ppGpp^0 $\mathrm{P}^{+1AA}_{fbaA}gfp$	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta relA$::upp-$phleo$-cI $fbaA$::$\mathrm{P}^{+1AA}_{fbaA}gfpmut3/spec$	↓ p/i $relA$ → CLB307
CLB317	(p)ppGpp^0 $\mathrm{P}^{+1GG}_{fbaA}gfp$ $guaB^{inducible}$	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta relA$::upp-$phleo$-cI $\Delta sacA$::$\mathrm{P}_{hs}guaB/cm$ $\Delta guaB$::$tetL$ $fbaA$::$\mathrm{P}^{+1GG}_{fbaA}gfpmut3/spec$	↓ p/i $relA$ → CLB313
CLB319	(p)ppGpp^0 $\mathrm{P}^{+1AA}_{fbaA}gfp$ $guaB^{inducible}$	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta relA$::upp-$phleo$-cI $\Delta sacA$::$\mathrm{P}_{hs}guaB/cm$ $\Delta guaB$::$tetL$ $fbaA$::$\mathrm{P}^{+1AA}_{fbaA}gfpmut3/spec$	↓ p/i $relA$ → CLB315
CLB322	RelA$^+$ $\mathrm{P}_{rrnJ}sfGFP$	Δupp::λPr-neo $\Delta relP$ $\Delta relQ$ $\Delta amyE$::$\mathrm{P}_{rrnJ}sfGFP/spec$	pCLB61 → CLB038

CLB322	RelA$^+$ P$_{rrnJ}$*sfGFP*	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ* Δ*amyE*::P$_{rrnJ}$*sfGFP/spec*	pCLB61 → CLB038
CLB327	RelA$^+$ P$_{rrnO}$*sfGFP*	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ* Δ*amyE*::P$_{rrnO}$*sfGFP/spec*	pCLB66 → CLB038
CLB329	RelA$^+$ P$_{rrnJ}$*sfGFP guaB^{inducible}*	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ* Δ*sacA*::P$_{hs}$*guaB/cm* Δ*guaB*::*tetL* Δ*amyE*::P$_{rrnJ}$*sfGFP/spec*	pCLB61 → CLB240
CLB335	RelA$^+$ P$_{rrnO}$*sfGFP guaB^{inducible}*	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ* Δ*sacA*::P$_{hs}$*guaB/cm* Δ*guaB*::*tetL* Δ*amyE*::P$_{rrnO}$*sfGFP/spec*	pCLB66 → CLB240
CLB349	(p)ppGpp^0 P$_{rrnJ}$*sfGFP*	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ* Δ*relA*::*upp-phleo-cI* Δ*amyE*::P$_{rrnJ}$*sfGFP/spec*	↓ p/i *relA* → CLB322
CLB353	(p)ppGpp^0 P$_{rrnO}$*sfGFP*	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ* Δ*relA*::*upp-phleo-cI* Δ*amyE*::P$_{rrnO}$*sfGFP/spec*	↓ p/i *relA* → CLB327
CLB355	(p)ppGpp^0 P$_{rrnJ}$*sfGFP guaB^{inducible}*	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ* Δ*relA*::*upp-phleo-cI* Δ*sacA*::P$_{hs}$*guaB/cm* Δ*guaB*::*tetL* Δ*amyE*::P$_{rrnJ}$*sfGFP/spec*	↓ p/i *relA* → CLB329
CLB359	(p)ppGpp^0 P$_{rrnO}$*sfGFP guaB^{inducible}*	Δ*upp*::λ*Pr-neo* Δ*relP* Δ*relQ* Δ*relA*::*upp-phleo-cI* Δ*sacA*::P$_{hs}$*guaB/cm* Δ*guaB*::*tetL* Δ*amyE*::P$_{rrnO}$*sfGFP/spec*	↓ p/i *relA* → CLB335

Table 8.2: Table of the primers used in this work

plasmid & construction[2]	**Sequence amplified**	**Primer forward**	**Primer reverse**
↓ p/i *relQ*	Sequence upstream of the *relQ* gene	P1_relQ TGTTTGATGACGGCGCGGAT	P2_relQ CGACCTGCAGGCATGCAACGTCATACATCCCCCAATTCCGAACC
	Sequence downstream of the *relQ* gene	P3_relQ AGCTCGAATTCACTGGCCGTCGAGAAACTGGTTCGGAATTGGGGGATGTATGGGTAAAGGGGAAGAAGAGCATGA	P4_relQ CTGGCGTTACCGTCCGTTCA
↓ p/i *relA*	Sequence upstream of the *relA* gene	P1_relA AGAGTGACCAAACCGAGGAT	P2_relA CGACCTGCAGGCATGCAACGTAGGGGTTAGAAAAGAGATTAGTTGC
	Sequence downstream of the *relA* gene	P3_relA CGAGCTCGAATTCACTGGCCGTCGGCTGAACAACTAATCTCTTTTCTAACCCCTCGTTCGCCATGGAATCACCT	P4_relA CGTCCCAGCTTGATTGCGTC
pCLB09	pSG13 reverse PCR	P_fw_GFP_fs CB223 ATGTCAAAAGGAGAAAGAACTTTTTACAGGTGTAGTACCTATCTTGGT	P_rev_GFP_fs CB227 TGTAAAAAGTTCTTTCTCCTTTTGACATGTTTGTCCTCCTTATTAGTT
pCLB22	pOB1 reverse PCR containing the $P_{hyperspank}$ *guaB* gene ORF	CB087_P_fw_PCRinv_sansGFP GCTTAATTAGCTGAGCTTGGACTCCT CB083_P_fw_guaB GATAACAATTGGTGGGAAGGAGGAACTACTATGTGGGAAAGTAAATTTTCAAAAGA	CB088_P_rev_PCRinv AGTAGTTCCTCCTTCCCACCAATTGTTAT CB085_P_rev_guaB_sansGFP CAACAGGAGTCCAAGCTCAGCTAATTAAGCTTATGAAATTGTATAGTTAGGTGATTCT

[2]vertical arrows (↓) indicate that deletions were introduced by homologous replacement of the targeted gene by a linear DNA fragment containing a marker or the evictable cassette *upp-phleo-cI* (p/i for pop in, see section 8.3.3)

	Sequence $P_{hyperspank}$-*guaB*	P_fw_Phs_plasmide CB242 GATAAGCTGTCAAACATGAGATTTTGCAAAAGTTGTTGAC	P_rev_terminator_plasmide CB253 TGTCGACTAATTCTCACCAATAAAAAACGC
	lacI gene	P_fw_lacI_plasmide CB241 TTGGTGAGAATTAGTCGACAGCTAGCCGCA	P_rev_lacI_plasmide CB248 TGGTAGCGACCGGCGCTCAGTCACTGCCCGCTTTCCAGTC
↓ *guaB*	Sequence upstream of the *guaB* gene	P_fw_zone_amont_guaB 268 AGGAAATTTGAGAGGAGCT	P_rev_zone_amont_guaB 261 TGTGAATAGGATGTATTCACTAGTAAATCCCCCTCTTTTC
	Sequence downstream of the *guaB* gene	P_fw_zone_aval_guaB CB265 TCTCAAAGGGATTTCTAATAAATTGTTACAAATTAAAAAC	P_rev_zone_aval_guaB CB258 TTTCAATTGATTAAAACCA
	tetL gene ORF	P_fw_tetL CB270 TCTCAAAGGGATTTCTAATAAATTGTTACAAATTAAAAAC	P_rev_tetL CB263 TTTCAATTGATTAAAACCA
pCLB33	pSG13 reverse PCR	P_fw_fs_codon_10 CB300 AAGAACTTTTTACACGGTGTAGTACCTATCTTGGTTGAAT	P_rev_fs_codon_10 CB309 ACACCGTGTAAAAAGTTCTTCTCCTTTTGACATGTTTGTC
pCLB35	pSG13 reverse PCR	P_fw_fs_codon_13 CB301 GGTGTAGTACACTATCTTGGTTGAATTGGATGGTGATGTT	P_rev_fs_codon_13 CB310 CCAAGATAGTGTACTACACCTGTAAAAAGTTCTTCTCCTT
pCLB37	pSG13 reverse PCR	P_fw_fs_codon_14 CB302 TGTAGTACCTCGTCTTGGTTGAATTGGATGGTGATGTTAA	P_rev_fs_codon_14 CB311 AACCAAGACGAGGTACTACACCTGTAAAAAGTTCTTCTCC
pCLB39	pSG13 reverse PCR	P_fw_fs_codon_16 CB303 CTTGGGTTGCATTGGATGGTGATGTTAACGGTCACAAATT	P_rev_fs_codon_16 CB312 ACCATCCAATGCAACCCAAGATAGGTACTACACCTGTAAA

pCLB43	pSG13 reverse PCR	P_fw_fs_isoleucine_AUC CB305 AGAACTTTTCATCGTGTAGTACC-TATCTTGGTTGAATTGG	P_rev_fs_isoleucine_AUC CB314 ACTA-CACGATGAAAAGTTCTTCTCCTTTTGA-CATGTTTGT
pCLB48	pSG13 reverse PCR	P_fw_fs_leucine_pos2 CB307 GT-CAAATTTCTTAAACTTTTTACAGGTG-TAGTACCTATCT	P_rev_fs_leucine_pos2 CB316 AAAAGTT-TAAGAAATTTGACATGTTTGTCCTCCT-TATTAG
pCLB69	pSG13 reverse PCR	P_fw_isoleucine_AUC_GFPnat CB339 AGAACTTTTCATCGTGGTAGTACC-TATCTTGGTTGAATTGG	P_rev_isoleucine_AUC_GFPnat CB341 ACTACACCGATGAAAAGTTCTTCTC-CTTTTGACATGTTTGT
pCLB71	pSG13 reverse PCR	P_fw_leucine_pos2_GFPnat CB340 GT-CAAATTTCTTAGAACTTTTTACAGGTG-TAGTACCTATCT	P_rev_leucine_pos2_GFPnat CB342 AAAAGTTCTAAGAAATTTGACAT-GTTTGTCCTCCTTATTAG
pCLB49	sequence upstream of the *fbaA* ORF	SENSE_PRM CCGCGGGCTTTC-CCAGCTTCGGCTGTGCCCGTCCCT-GAGCAAAACAT	ANTISENSE_PRM GTTCCTCCTTCCCAC-CTGTAGCCTGATTGTCTTCTAGCG
pCLB51	pCLB49 reverse PCR	primer_fw_AA_+1 CB6 TGGTATCCTAGT-TATAAAGAAAAAGCGATCTGAGTATTTA-CATATGACAGCA	CB7_rev_AA_+1 CAGATCGCTTTTTCTT-TATAACTAGGATACCAACTTTTCGA-CAAATCGGCAA
pCLB61	pSG13 reverse PCR	P_fw_SG13_rrnJ_1 CB323 AGTCGCTTGAACGAAGCTAGCGAT-TAACTAATAAG	P_rev_SG13_rrnJ TACCGCTCTTTTTT-TATAAGAATAGTAATACAGGAT
	P*rrnJ* promoter sequence	P_fw_rrnJ CB321 CGGATCCTGTATTAC-TATTCTTATAAAAAAGAGCGGTATCCTC	P_rev_rrnJ_1 CB327 TAGTTAATCGC-TAGCTTCGTTCAAGCGACTTTATT
pCLB66	pSG13 reverse PCR	P_fw_SG13_rrnO_2 CB326 GTAGTCGCTTTGAGAGCTAGCGAT-TAACTAATAAG	P_rev_SG13_rrnO CB332 TAGTTAATCGC-TAGCTCTCAAAGCGACTACTTAATAG

	P*rrnO* promoter sequence	P_fw_rrnO CB322 CGGATCCTGTATTACTATTCTTAAAAAAAGCGCAGCTGAAATA	P_rev_rrnO_2 CB330 CAGCTGCGCTTTTTTTAAGAATAGTAATACAGGAT

Bibliography

Agris, P. F., Eruysal, E. R., Narendran, A., Väre, V. Y., Vangaveti, S., and Ranganathan, S. V. Celebrating wobble decoding: Half a century and still much is new. *RNA biology*, pages 1–17, 2017.

Aïchaoui, L., Jules, M., Le Chat, L., Aymerich, S., Fromion, V., and Goelzer. Basylica: a tool for automatic processing of a bacterial live cell array. *Bioinformatics*, 28(20):2705–2706, 2012.

Anagnostopoulos, C. and Spizizen, J. Requirements for transformation in *Bacillus subtilis*. *Journal of bacteriology*, 81(5):741, 1961.

Andersen, G. R., Nissen, P., and Nyborg, J. Elongation factors in protein biosynthesis. *Trends in biochemical sciences*, 28(8):434–441, 2003.

Arenz, S., Abdelshahid, M., Sohmen, D., Payoe, R., Starosta, A. L., Berninghausen, O., Hauryliuk, V., Beckmann, R., and Wilson, D. N. The stringent factor rela adopts an open conformation on the ribosome to stimulate ppgpp synthesis. *Nucleic acids research*, 44(13):6471–6481, 2016.

Atkins, J., Elseviers, D., and Gorini, L. Low activity of β-galactosidase in frameshift mutants of *escherichia coli*. *Proceedings of the National Academy of Sciences*, 69(5):1192–1195, 1972.

Bais, H. P., Weir, T. L., Perry, L. G., Gilroy, S., and Vivanco, J. M. The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu. Rev. Plant Biol.*, 57:233–266, 2006.

Barak, Z., Lindsley, D., and Gallant, J. On the mechanism of leftward frameshifting at several hungry codons. *Journal of molecular biology*, 256(4):676–684, 1996.

Barbe, V., Cruveiller, S., Kunst, F., Lenoble, P., Meurice, G., Sekowska, A., Vallenet, D., Wang, T., Moszer, I., Médigue, C., et al. From a consortium sequence to a unified sequence: the *Bacillus subtilis* 168 reference genome a decade later. *Microbiology*, 155(6):1758–1775, 2009.

Belitsky, B. R. and Sonenshein, A. L. Genome-wide identification of *Bacillus subtilis* cody-binding sites at single-nucleotide resolution. *Proceedings of the National Academy of Sciences*, 110(17): 7026–7031, 2013.

Benoist, C., Guérin, C., Noirot, P., and Dervyn, E. Constitutive stringent response restores viability of *Bacillus subtilis* lacking structural maintenance of chromosome protein. *PloS one*, 10(11):e0142308, 2015.

Berg, J., Tymoczko, J., and Stryer, L. Dna replications of both strands proceeds rapidly from specific start sites. WH Freeman, 2002.

Bertani, G. Lysogeny at mid-twentieth century: P1, p2, and other experimental systems. *Journal of bacteriology*, 186(3):595–600, 2004.

Bidnenko, V., Nicolas, P., Grylak-Mielnicka, A., Delumeau, O., Auger, S., Aucouturier, A., Guerin, C., Francis, R., Bardowski, J., Aymerich, S., et al. Termination factor rho: From the control of pervasive transcription to cell fate determination in *Bacillus subtilis*. *PLoS genetics*, 13(7):e1006909, 2017.

Bimboim, H. and Doly, J. A rapid alkaline extraction procedure for screening recombinant plasmid dna. *Nucleic acids research*, 7(6):1513–1523, 1979.

Bittner, A. N., Kriel, A., and Wang, J. D. Lowering gtp level increases survival of amino acid starvation but slows growth rate for *Bacillus subtilis* cells lacking (p)ppgpp. *Journal of bacteriology*, 196(11): 2067–2076, 2014.

Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., et al. The complete genome sequence of *escherichia coli* k-12. *science*, 277(5331):1453–1462, 1997.

Blinkowa, A. and Walker, J. Programmed ribosomal frameshifting generates the *escherichia coli* dna polymerase iii γ subunit from within the τ subunit reading frame. *Nucleic acids research*, 18(7): 1725–1729, 1990.

Boehm, A., Steiner, S., Zaehringer, F., Casanova, A., Hamburger, F., Ritz, D., Keck, W., Ackermann, M., Schirmer, T., and Jenal, U. Second messenger signalling governs *escherichia coli* biofilm induction upon ribosomal stress. *Molecular microbiology*, 72(6):1500–1516, 2009.

Boor, K. J., Duncan, M. L., and Price, C. W. Genetic and transcriptional organization of the region encoding the β subunit of *Bacillus subtilis* rna polymerase. *Journal of Biological Chemistry*, 270 (35):20329–20336, 1995.

Borkowski, O., Goelzer, A., Schaffer, M., Calabre, M., Mäder, U., Aymerich, S., Jules, M., and Fromion, V. Translation elicits a growth rate-dependent, genome-wide, differential protein production in *Bacillus subtilis*. *Molecular systems biology*, 12(5):870, 2016.

Boros, I., Kiss, A., and Venetianer, P. Physical map of the seven ribosomal rna genes of *escherichia coli*. *Nucleic acids research*, 6(5):1817–1830, 1979.

Bosdriesz, E., Molenaar, D., Teusink, B., and Bruggeman, F. How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization. *The FEBS journal*, 282 (10):2029–2044, 2015.

Botella, E., Fogg, M., Jules, M., Piersma, S., Doherty, G., Hansen, E., Aand Denham, Le Chat, L., Veiga, P., Bailey, K., et al. pbasysbioii: an integrative plasmid generating *gfp* transcriptional fusions for high-throughput analysis of gene expression in *Bacillus subtilis*. *Microbiology*, 156(6):1600–1608, 2010.

Bremer, H. and Dennis, P. Modulation of chemical composition and other parameters of the cell at different exponential growth rates. *EcoSal-Escherichia coli and Salmonella*, 2008.

Brinsmade, S. R. and Sonenshein, A. L. Dissecting complex metabolic integration provides direct genetic evidence for cody activation by guanine nucleotides. *Journal of bacteriology*, 193(20):5637–5648, 2011.

Brown, A., Fernández, I. S., Gordiyenko, Y., and Ramakrishnan, V. Ribosome-dependent activation of stringent control. *Nature*, 534(7606):277, 2016.

Browning, D. F. and Busby, S. J. The regulation of bacterial transcription initiation. *Nature Reviews Microbiology*, 2(1):57, 2004.

Browning, D. F. and Busby, S. J. Local and global regulation of transcription initiation in bacteria. *Nature Reviews Microbiology*, 14(10):638, 2016.

Buckingham, R. H., Grentzmann, G., and Kisselev, L. Polypeptide chain release factors. *Molecular microbiology*, 24(3):449–456, 1997.

Buescher, J., Liebermeister, W., Jules, M., Uhr, M., Muntel, E., Jand Botella, Hessling, B., Kleijn, R., Le Chat, L., Lecointe, F., et al. Global network reorganization during dynamic adaptations of *Bacillus subtilis* metabolism. *Science*, 335(6072):1099–1103, 2012.

Burgess, A. B. and Mach, B. Formation of an initiation complex with purified mammalian ribosomal subunits. *Nature New Biology*, 233(41):209, 1971.

Burley, S. K., Berman, H. M., Kleywegt, G. J., Markley, J. L., Nakamura, H., and Velankar, S. Protein data bank (pdb): the single global macromolecular structure archive. *Protein Crystallography: Methods and Protocols*, pages 627–641, 2017.

Caliskan, N., Katunin, V., Belardinelli, R., Peske, F., and Rodnina, M. Programmed -1 frameshifting by kinetic partitioning during impeded translocation. *Cell*, 157(7):1619–1631, 2014.

Caliskan, N., Wohlgemuth, I., Korniy, N., Pearson, M., Peske, F., and Rodnina, M. Conditional switch between frameshifting regimes upon translation of *dnaX* mrna. *Molecular cell*, 66(4):558–567, 2017.

Carter, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., Hartsch, T., Wimberly, B. T., and Ramakrishnan, V. Crystal structure of an initiation factor bound to the 30s ribosomal subunit. *Science*, 291(5503):498–501, 2001.

Cashel, M. The control of ribonucleic acid synthesis in *escherichia coli* iv. relevance of unusual phosphorylated compounds from amino acid-starved stringent strains. *Journal of Biological Chemistry*, 244(12):3133–3141, 1969.

Cashel, M. and Gallant, J. Two compounds implicated in the function of the rc gene of *escherichia coli*. *Nature*, 221(5183):838, 1969.

Chatterji, D., Fujita, N., and Ishihama, A. The mediator for stringent control, ppgpp, binds to the β-subunit of *escherichia coli* rna polymerase. *Genes to Cells*, 3(5):279–287, 1998.

Chiaramello, A. E. and Zyskind, J. W. Coupling of dna replication to growth rate in *escherichia coli*: a possible role for guanosine tetraphosphate. *Journal of bacteriology*, 172(4):2013–2019, 1990.

Cho, B.-K., Kim, D., Knight, E. M., Zengler, K., and Palsson, B. O. Genome-scale reconstruction of the sigma factor network in *escherichia coli*: topology and functional states. *BMC biology*, 12(1):4, 2014.

Commichau, F. M., Pietack, N., and Stülke, J. Essential genes in *Bacillus subtilis*: a re-evaluation after ten years. *Molecular BioSystems*, 9(6):1068–1075, 2013.

Consortium, U. Uniprot: the universal protein knowledgebase. *Nucleic acids research*, 45(D1):D158–D169, 2016.

Corrigan, R. M., Bellows, L. E., Wood, A., and Gründling, A. ppgpp negatively impacts ribosome assembly affecting growth and antimicrobial tolerance in gram-positive bacteria. *Proceedings of the National Academy of Sciences*, 113(12):E1710–E1719, 2016.

Craigen, W. and Caskey, C. Expression of peptide chain release factor 2 requires high-efficiency frameshift. *Nature*, 322(6076):273, 1986.

Crick, F. Central dogma of molecular biology. *Nature*, 227(5258):561, 1970.

Crick, F. H. CodonŮanticodon pairing: the wobble hypothesis. *Journal of molecular biology*, 19(2): 548–555, 1966.

Curran, J. Analysis of effects of trna: message stability on frameshift frequency at the *escherichia coli* rf2 programmed frameshift site. *Nucleic acids research*, 21(8):1837–1843, 1993.

Curran, J. and Yarus, M. Use of trna suppressors to probe regulation of *escherichia coli* release factor 2. *Journal of molecular biology*, 203(1):75–83, 1988.

Dalebroux, Z. D. and Swanson, M. S. ppgpp: magic beyond rna polymerase. *Nature Reviews Microbiology*, 10(3):203–212, 2012.

De Hoon, M. J., Makita, Y., Nakai, K., and Miyano, S. Prediction of transcriptional terminators in *Bacillus subtilis* and related species. *PLoS computational biology*, 1(3):e25, 2005.

De Hoon, M. J., Eichenberger, P., and Vitkup, D. Hierarchical evolution of the bacterial sporulation network. *Current biology*, 20(17):R735–R745, 2010.

De San Eustaquio-Campillo, A., Cornilleau, C., Guérin, C., Carballido-López, R., and Chastanet, A. Pamr, a new marr-like regulator affecting prophages and metabolic genes expression in *Bacillus subtilis*. *PloS one*, 12(12):e0189694, 2017.

DeNapoli, J., Tehranchi, A. K., and Wang, J. D. Dose-dependent reduction of replication elongation rate by (p)ppgpp in *escherichia coli* and *Bacillus subtilis*. *Molecular microbiology*, 88(1):93–104, 2013.

Dervyn, E., Suski, C., Daniel, R., Bruand, C., Chapuis, J., Errington, J., Jannière, L., and Ehrlich, S. D. Two essential dna polymerases at the bacterial replication fork. *Science*, 294(5547):1716–1719, 2001.

Dirheimer, G., Keith, G., Dumas, P., and Westhof, E. Rna : structure, biosynthesis and function. 1995.

Dix, D. B. and Thompson, R. C. Elongation factor tu. guanosine 3'-diphosphate 5'-diphosphate complex increases the fidelity of proofreading in protein biosynthesis: mechanism for reducing translational errors introduced by amino acid starvation. *Proceedings of the National Academy of Sciences*, 83(7):2027–2031, 1986.

Doerfel, L., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H., and Rodnina, M. Ef-p is essential for rapid synthesis of proteins containing consecutive proline residues. *Science*, page 1229017, 2012.

Doherty, G. P., Fogg, M. J., Wilkinson, A. J., and Lewis, P. J. Small subunits of rna polymerase: localization, levels and implications for core enzyme composition. *Microbiology*, 156(12):3532–3543, 2010.

Donly, B., Edgar, C., Adamski, F., and Tate, W. Frameshift autoregulation in the gene for *escherichia coli* release factor 2: partly functional mutants result in frameshift enhancement. *Nucleic acids research*, 18(22):6517–6522, 1990.

Earl, A. M., Losick, R., and Kolter, R. Ecology and genomics of *Bacillus subtilis*. *Trends in microbiology*, 16(6):269–275, 2008.

Ebright, R. H. and Busby, S. The *Escherichia coli* rna polymerase α subunit: structure and function. *Current opinion in genetics & development*, 5(2):197–203, 1995.

Efron, B. Better bootstrap confidence intervals. *Journal of the American statistical Association*, 82 (397):171–185, 1987.

Epshtein, V. and Nudler, E. Cooperation between rna polymerase molecules in transcription elongation. *Science*, 300(5620):801–805, 2003.

Erickson, D., Schink, S., Patsalo, V., Williamson, J., and Gerland, T., Uand Hwa. A global resource allocation strategy governs growth transition kinetics of *escherichia coli*. *Nature*, 551(7678):119, 2017.

Fabbretti, A., Brandi, L., Milon, P., Spurio, R., Pon, C. L., and Gualerzi, C. O. Translation initiation without if2-dependent gtp hydrolysis. *Nucleic acids research*, 40(16):7946–7955, 2012.

Fabret, C., ES, D., and Noirot, P. A new mutation delivery system for genome-scale approaches in *Bacillus subtilis*. *Molecular microbiology*, 46(1):25–36, 2002.

Farabaugh, P. Programmed translational frameshifting. *Annual review of genetics*, 30(1):507–528, 1996.

Farabaugh, P. and Björk, G. How translational accuracy influences reading frame maintenance. *The EMBO journal*, 18(6):1427–1434, 1999.

Feng, B., Mandava, C. S., Guo, Q., Wang, J., Cao, W., Li, N., Zhang, Y., Zhang, Y., Wang, Z., Wu, J., et al. Structural and functional insights into the mode of action of a universally conserved obg gtpase. *PLoS biology*, 12(5):e1001866, 2014.

Feng, D.-F., Cho, G., and Doolittle, R. F. Determining divergence times with a protein clock: update and reevaluation. *Proceedings of the National Academy of Sciences*, 94(24):13028–13033, 1997.

Flower, A. M. and McHenry, C. S. The gamma subunit of dna polymerase iii holoenzyme of *escherichia coli* is produced by ribosomal frameshifting. *Proceedings of the National Academy of Sciences*, 87 (10):3713–3717, 1990.

Fukuoka, T., Moriya, S., Yoshikawa, H., and Ogasawara, N. Purification and characterization of an initiation protein for chromosomal replication, dnaa, in *Bacillus subtilis*. *The Journal of Biochemistry*, 107(5):732–739, 1990.

Gaca, A. O., Colomer-Winter, C., and Lemos, J. A. Many means to a common end: the intricacies of (p)ppgpp metabolism and its control of bacterial homeostasis. *Journal of bacteriology*, 197(7): 1146–1156, 2015.

Gallant, J., Irr, J., and Cashel, M. The mechanism of amino acid control of guanylate and adenylate biosynthesis. *Journal of Biological Chemistry*, 246(18):5812–5816, 1971.

Gamper, H., Masuda, I., Frenkel-Morgenstern, M., and Hou, Y. Maintenance of protein synthesis reading frame by ef-p and m1 g37-trna. *Nature communications*, 6:7226, 2015.

Geiger, T., Kästle, B., Gratani, F. L., Goerke, C., and Wolz, C. Two small (p)ppgpp synthases in *Staphylococcus aureus* mediate tolerance against cell envelope stress conditions. *Journal of bacteriology*, 196(4):894–902, 2014.

Gerhart, E., Wagner, H., Ehrenberg, M., and Kurland, C. Kinetic suppression of translational errors by (p)ppgpp. *Molecular and General Genetics MGG*, 185(2):269–274, 1982.

Gerosa, L., Kochanowski, K., Heinemann, M., and Sauer, U. Dissecting specific and global transcriptional regulation of bacterial gene expression. *Molecular systems biology*, 9(1):658, 2013.

Gibson, D., Young, L., Chuang, R., Venter, J., Hutchison III, C., and Smith, H. Enzymatic assembly of dna molecules up to several hundred kilobases. *Nature methods*, 6(5):343, 2009.

Giedroc, D., Theimer, C., and Nixon, P. Structure, stability and function of rna pseudoknots involved in stimulating ribosomal frameshifting. *Journal of molecular biology*, 298(2):167–185, 2000.

Giordano, N., Mairet, F., Gouzé, J., Geiselmann, J., and De Jong, H. Dynamical allocation of cellular resources as an optimal control problem: novel insights into microbial growth strategies. *PLoS computational biology*, 12(3):e1004802, 2016.

Glass, R. E., Jones, S. T., Nene, V., Nomura, T., Fujita, N., and Ishihama, A. Genetic studies on the β subunit of *Escherichia coli* rna polymerase. *Molecular and General Genetics MGG*, 203(3): 487–491, 1986.

Goelzer, A., Fromion, V., and Scorletti, G. Cell design in bacteria as a convex optimization problem. In *Proceedings of the 48th IEEE Conference on Decision and Control*, page 4517, 2009.

Goelzer, A. and Fromion, V. Bacterial growth rate reflects a bottleneck in resource allocation. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1810(10):978–988, 2011.

Goelzer, A., Fromion, V., and Scorletti, G. Cell design in bacteria as a convex optimization problem. *Automatica*, 47(6):1210–1218, 2011.

Goelzer, A., Muntel, J., Chubukov, V., Jules, M., Prestel, E., Nölker, R., Mariadassou, M., Aymerich, S., Hecker, M., Noirot, P., et al. Quantitative prediction of genome-wide resource allocation in bacteria. *Metabolic engineering*, 32:232–243, 2015.

González-Pastor, J., Hobbs, E., and Losick, R. Cannibalism by sporulating bacteria. *Science*, 301 (5632):510–513, 2003.

Guiziou, S., Sauveplane, V., Chang, H.-J., Clerté, C., Declerck, N., Jules, M., and Bonnet, J. A part toolbox to tune genetic expression in *Bacillus subtilis*. *Nucleic acids research*, 44(15):7495–7508, 2016.

Gurvich, O., Baranov, P. V., Zhou, J., Hammer, A., Gesteland, R., and Atkins, J. Sequences that direct significant levels of frameshifting are frequent in coding regions of *escherichia coli*. *The EMBO journal*, 22(21):5941–5950, 2003.

Gusarov, I. and Nudler, E. The mechanism of intrinsic transcription termination. *Molecular cell*, 3 (4):495–504, 1999.

Hagervall, T., Tuohy, T., Atkins, J. F., and Björk, G. Deficiency of 1-methylguanosine in trna from *Salmonella typhimurium* induces frameshifting by quadruplet translocation. *Journal of molecular biology*, 232(3):756–765, 1993.

Haldenwang, W. G. The sigma factors of *Bacillus subtilis*. *Microbiological reviews*, 59(1):1–30, 1995.

Hall, B. and Gallant, J. Defective translation in rc- cells. *Nature New Biology*, 237(74):131, 1972.

Hamel, E. and Cashel, M. Role of guanine nucleotides in protein synthesis. elongation factor g and guanosine 5'-triphosphate, 3'-diphosphate. *Proceedings of the National Academy of Sciences*, 70 (11):3250–3254, 1973.

Hamoen, L. W., Venema, G., and Kuipers, O. P. Controlling competence in *Bacillus subtilis*: shared use of regulators. *Microbiology*, 149(1):9–17, 2003.

Handke, L. D., Shivers, R. P., and Sonenshein, A. L. Interaction of *Bacillus subtilis* cody with gtp. *Journal of bacteriology*, 190(3):798–806, 2008.

Harger, J., Meskauskas, A., and Dinman, J. An 'integrated model' of programmed ribosomal frameshifting. *Trends in biochemical sciences*, 27(9):448–454, 2002.

Harwood, C. R. and Cutting, S. M. *Molecular biological methods for Bacillus*. Wiley, 1990.

Haugen, S. P., Ross, W., and Gourse, R. L. Advances in bacterial promoter recognition and its control by factors that do not bind dna. *Nature Reviews Microbiology*, 6(7):507, 2008.

Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T., and Gerdes, K. Recent functional insights into the role of (p)ppgpp in bacterial physiology. *Nature Reviews Microbiology*, 13(5):298, 2015.

Henkin, T. M. The role of the ccpa transcriptional regulator in carbon metabolism in *Bacillus subtilis*. *FEMS microbiology letters*, 135(1):9–15, 1996.

Henkin, T. M. Ribosomes, protein synthesis factors, and trna synthetases. In *Bacillus subtilis and Its Closest Relatives*, pages 313–322. American Society of Microbiology, 2002.

Henkin, T. M. and Yanofsky, C. Regulation by transcription attenuation in bacteria: how rna provides instructions for transcription termination/antitermination decisions. *Bioessays*, 24(8):700–707, 2002.

Hirokawa, G., Kiel, M. C., Muto, A., Selmer, M., Raj, V. S., Liljas, A., Igarashi, K., Kaji, H., and Kaji, A. Post-termination complex disassembly by ribosome recycling factor, a functional trna mimic. *The EMBO journal*, 21(9):2272–2281, 2002.

Hubert, M., Nyengaard, N. R., Shazand, K., Mortensen, K. K., Lassen, S. F., Grunberg-Manago, M., and Sperling-Petersen, H. U. Tandem translation of *Bacillus subtilis* initiation factor if2 in *E. coli* over-expression of $infB_{B.su}$ in *E. coli* and purification of α-and β-forms of $if2_{B.su}$. *FEBS letters*, 312 (2-3):132–138, 1992.

Ieong, K.-W., Uzun, Ü., Selmer, M., and Ehrenberg, M. Two proofreading steps amplify the accuracy of genetic code translation. *Proceedings of the National Academy of Sciences*, 113(48):13744–13749, 2016.

Ingham, C. J., Dennis, J., and Furneaux, P. A. Autogenous regulation of transcription termination factor rho and the requirement for nus factors in *Bacillus subtilis*. *Molecular microbiology*, 31(2): 651–663, 1999.

Jäger, G., Nilsson, K., and Björk, G. The phenotype of many independently isolated +1 frameshift suppressor mutants supports a pivotal role of the p-site in reading frame maintenance. *PLoS One*, 8(4):e60246, 2013.

Jameson, K. H. and Wilkinson, A. J. Control of initiation of dna replication in *Bacillus subtilis* and *escherichia coli*. *Genes*, 8(1):22, 2017.

Javid, B., Sorrentino, F., Toosky, M., Zheng, W., Pinkham, J. T., Jain, N., Pan, M., Deighan, P., and Rubin, E. J. Mycobacterial mistranslation is necessary and sufficient for rifampicin phenotypic resistance. *Proceedings of the National Academy of Sciences*, 111(3):1132–1137, 2014.

Jin, D. J., Burgess, R. R., Richardson, J. P., and Gross, C. A. Termination efficiency at rho-dependent terminators depends on kinetic coupling between rna polymerase and rho. *Proceedings of the National Academy of Sciences*, 89(4):1453–1457, 1992.

Jin, D. J., Cagliero, C., and Zhou, Y. N. Growth rate regulation in *Escherichia coli*. *FEMS microbiology reviews*, 36(2):269–287, 2012.

Jishage, M., Kvint, K., Shingler, V., and Nyström, T. Regulation of ς factor competition by the alarmone ppgpp. *Genes & development*, 16(10):1260–1270, 2002.

Johnsen, K., Molin, S., Karlstrom, O., and Maaløe, O. Control of protein synthesis in *escherichia coli*: analysis of an energy source shift-down. *Journal of bacteriology*, 131(1):18–29, 1977.

Juang, Y.-L. and Helmann, J. D. The δ subunit of *Bacillus subtilis* rna polymerase: an allosteric effector of the initiation and core-recycling phases of transcription. *Journal of molecular biology*, 239(1):1–14, 1994.

Jun, S., Si, F., Pugatch, R., and Scott, M. Fundamental principles in bacterial physiology-history, recent progress, and the future with focus on cell size control: A review. *Reports on Progress in Physics*, 2018.

Kaczanowska, M. and Rydén-Aulin, M. Ribosome biogenesis and the translation process in *escherichia coli*. *Microbiology and Molecular Biology Reviews*, 71(3):477–494, 2007.

Kanaya, S., Yamada, Y., Kudo, Y., and Ikemura, T. Studies of codon usage and trna genes of 18 unicellular organisms and quantification of *Bacillus subtilis* trnas: gene expression level and species-specific diversity of codon usage based on multivariate analysis. *Gene*, 238(1):143–155, 1999.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. Kegg : new perspectives on genomes, pathways, diseases and drugs. *Nucleic acids research*, 45(D1):D353–D361, 2016.

Kanjee, U., Gutsche, I., Alexopoulos, E., Zhao, B., El Bakkouri, M., Thibault, G., Liu, K., Ramachandran, S., Snider, J., Pai, E., et al. Linkage between the bacterial acid stress and stringent responses: the structure of the inducible lysine decarboxylase. *The EMBO Journal*, 30(5):931–944, 2011.

Kanjee, U., Ogata, K., and Houry, W. A. Direct binding targets of the stringent response alarmone (p)ppgpp. *Molecular microbiology*, 85(6):1029–1043, 2012.

Keiler, K. and Feaga, H. Resolving nonstop translation complexes is a matter of life or death. *Journal of bacteriology*, 196(12):2123–2130, 2014.

Kihira, K., Shimizu, Y., Shomura, Y., Shibata, N., Kitamura, M., Nakagawa, A., Ueda, T., Ochi, K., and Higuchi, Y. Crystal structure analysis of the translation factor rf3 (release factor 3). *FEBS letters*, 586(20):3705–3709, 2012.

Kleijn, R. J., Buescher, J. M., Le Chat, L., Jules, M., Aymerich, S., and Sauer, U. Metabolic fluxes during strong carbon catabolite repression by malate in *Bacillus subtilis*. *Journal of Biological Chemistry*, 285(3):1587–1596, 2010.

Klumpp, S. and Hwa, T. Growth-rate-dependent partitioning of rna polymerases in bacteria. *Proceedings of the National Academy of Sciences*, 105(51):20245–20250, 2008.

Klumpp, S., Zhang, Z., and Hwa, T. Growth rate-dependent global effects on gene expression in bacteria. *Cell*, 139(7):1366–1375, 2009.

Klumpp, S., Scott, M., Pedersen, S., and Hwa, T. Molecular crowding limits translation and cell growth. *Proceedings of the National Academy of Sciences*, 110(42):16754–16759, 2013.

Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., et al. Essential *Bacillus subtilis* genes. *Proceedings of the National Academy of Sciences*, 100(8):4678–4683, 2003.

Kolor, K., Lindsley, D., and Gallant, J. On the role of the p-site in leftward ribosome frameshifting at a hungry codon, 1993.

Koo, B.-M., Kritikos, G., Farelli, J. D., Todor, H., Tong, K., Kimsey, H., Wapinski, I., Galardini, M., Cabal, A., Peters, J. M., et al. Construction and analysis of two genome-scale deletion libraries for *Bacillus subtilis*. *Cell systems*, 4(3):291–305, 2017.

Kornberg, A. and Baker, T. Biosynthesis of dna precursors. *DNA replication*, 2:68–69, 1992.

Kramer, E. and Farabaugh, P. The frequency of translational misreading errors in *e. coli* is largely determined by trna competition. *Rna*, 13(1):87–96, 2007.

Krásnỳ, L. and Gourse, R. L. An alternative strategy for bacterial ribosome synthesis: *Bacillus subtilis rRNA* transcription regulation. *The EMBO journal*, 23(22):4473–4483, 2004.

Krásnỳ, L., Tišerová, H., Jonák, J., Rejman, D., and Šanderová, H. The identity of the transcription +1 position is crucial for changes in gene expression in response to amino acid starvation in *Bacillus subtilis*. *Molecular microbiology*, 69(1):42–54, 2008.

Kriel, A., Bittner, A. N., Kim, S. H., Liu, K., Tehranchi, A. K., Zou, W. Y., Rendon, S., Chen, R., Tu, B. P., and Wang, J. D. Direct regulation of gtp homeostasis by (p)ppgpp: a critical component of viability and stress resistance. *Molecular cell*, 48(2):231–241, 2012.

Kriel, A., Brinsmade, S. R., Jessica, L. T., Tehranchi, A. K., Bittner, A. N., Sonenshein, A. L., and Wang, J. D. Gtp dysregulation in *Bacillus subtilis* cells lacking (p)ppgpp results in phenotypic amino acid auxotrophy and failure to adapt to nutrient downshift and regulate biosynthesis genes. *Journal of bacteriology*, 196(1):189–201, 2014.

Kudrin, P., Dzhygyr, I., Ishiguro, K., Beljantseva, J., Maksimova, E., Oliveira, S. R. A., Varik, V., Payoe, R., Konevega, A. L., Tenson, T., et al. The ribosomal a-site finger is crucial for binding and activation of the stringent factor rela. *Nucleic acids research*, 2018.

Kumar, V., Chen, Y., Ero, R., Ahmed, T., Tan, J., Li, Z., Wong, A. S. W., Bhushan, S., and Gao, Y.-G. Structure of bipa in gtp form bound to the ratcheted ribosome. *Proceedings of the National Academy of Sciences*, 112(35):10944–10949, 2015.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A., Alloni, G., Azevedo, V., Bertero, M., Bessieres, P., Bolotin, A., Borchert, S., et al. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature*, 390(6657):249, 1997.

Kuroda, A., Murphy, H., Cashel, M., and Kornberg, A. Guanosine tetra-and pentaphosphate promote accumulation of inorganic polyphosphate in *escherichia coli*. *Journal of Biological Chemistry*, 272 (34):21240–21243, 1997.

Lainé, S., Thouard, A., Komar, A., and Rossignol, J. Ribosome can resume the translation in both+ 1 or- 1 frames after encountering an aga cluster in *Escherichia coli*. *Gene*, 412(1):95–101, 2008.

Larsen, B., Wills, N., Gesteland, R., and Atkins, J. rrna-mrna base pairing stimulates a programmed -1 ribosomal frameshift. *Journal of bacteriology*, 176(22):6842–6851, 1994.

Larsen, B., Gesteland, R., and Atkins, J. Structural probing and mutagenic analysis of the stem-loop required for *Escherichia coli dnaX* ribosomal frameshifting: programmed efficiency of 50%. *Journal of molecular biology*, 271(1):47–60, 1997.

Laursen, B. S., Sørensen, H. P., Mortensen, K. K., and Sperling-Petersen, H. U. Initiation of protein synthesis in bacteria. *Microbiology and molecular biology reviews*, 69(1):101–123, 2005.

Lawrence, J. G. and Ochman, H. Molecular archaeology of the *escherichia coli* genome. *Proceedings of the National Academy of Sciences*, 95(16):9413–9417, 1998.

Lechat, P., Hummel, L., Rousseau, S., and Moszer, I. Genolist : an integrated environment for comparative analysis of microbial genomes. *Nucleic acids research*, 36(suppl_1):D469–D474, 2007.

Li, J., Esberg, B., Curran, J., and BjoÈrk, G. Three modified nucleosides present in the anticodon stem and loop influence the in vivo aa-trna selection in a trna-dependent manner. *Journal of molecular biology*, 271(2):209–221, 1997a.

Li, W., Bouveret, E., Zhang, Y., Liu, K., Wang, J. D., and Weisshaar, J. C. Effects of amino acid starvation on rela diffusive behavior in live *escherichia coli*. *Molecular microbiology*, 99(3):571–585, 2016.

Li, X., Lindahl, L., Sha, Y., and Zengel, J. M. Analysis of the *Bacillus subtilis* s10 ribosomal protein gene cluster identifies two promoters that may be responsible for transcription of the entire 15-kilobase s10-spc-alpha cluster. *Journal of bacteriology*, 179(22):7046–7054, 1997b.

Liang, S.-T., Xu, Y.-C., Dennis, P., and Bremer, H. mrna composition and control of bacterial gene expression. *Journal of Bacteriology*, 182(11):3037–3044, 2000.

Liao, P., Gupta, P., Petrov, A., Dinman, J., and Lee, K. A new kinetic model reveals the synergistic effect of e-, p-and a-sites on + 1 ribosomal frameshifting. *Nucleic acids research*, 36(8):2619–2629, 2008.

Liao, P., Choi, Y., and Lee, K. Fsscan: a mechanism-based program to identify +1 ribosomal frameshift hotspots. *Nucleic acids research*, 37(21):7302–7311, 2009.

Lindsley, D. and Gallant, J. On the directional specificity of ribosome frameshifting at a "hungry" codon. *Proceedings of the National Academy of Sciences*, 90(12):5469–5473, 1993.

Liu, K., Bittner, A. N., and Wang, J. D. Diversity in (p)ppgpp metabolism and effectors. *Current opinion in microbiology*, 24:72–79, 2015a.

Liu, K., Myers, A. R., Pisithkul, T., Claas, K. R., Satyshur, K. A., Amador-Noguez, D., Keck, J. L., and Wang, J. D. Molecular mechanism and evolution of guanylate kinase regulation by (p)ppgpp. *Molecular cell*, 57(4):735–749, 2015b.

Lopez, J. M., Dromerick, A., and Freese, E. Response of guanosine 5'-triphosphate concentration to nutritional changes and its significance for *Bacillus subtilis* sporulation. *Journal of bacteriology*, 146 (2):605–613, 1981.

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

Loveland, A. B., Demo, G., Grigorieff, N., and Korostelev, A. A. Ensemble cryo-em elucidates the mechanism of translation fidelity. *Nature*, 546(7656):113, 2017.

Maaløe, O. and Kjeldgaard, N. O. Control of macromolecular synthesis: a study of dna, rna, and protein synthesis in bacteria. 1966.

Maciag, M., Kochanowska, M., Lyzen, R., Wegrzyn, G., and Szalewska-Palasz, A. ppgpp inhibits the activity of *escherichia coli* dnag primase. *Plasmid*, 63(1):61–67, 2010.

Maciag-Dorszynska, M., Szalewska-Palasz, A., and Wegrzyn, G. Different effects of ppgpp on *escherichia coli* dna replication in vivo and in vitro. *FEBS Open Bio*, 3(1):161–164, 2013.

Maehigashi, T., Dunkle, J., Miles, S., and Dunham, C. Structural insights into +1 frameshifting promoted by expanded or modification-deficient anticodon stem loops. *Proceedings of the National Academy of Sciences*, 111(35):12740–12745, 2014.

Makrides, S. C. Strategies for achieving high-level expression of genes in *escherichia coli*. *Microbiological reviews*, 60(3):512–538, 1996.

Malys, N. Shine-dalgarno sequence of bacteriophage t4: Gagg prevails in early genes. *Molecular biology reports*, 39(1):33–39, 2012.

Maracci, C. and Rodnina, M. V. translational gtpases. *Biopolymers*, 105(8):463–475, 2016.

Margeat, E., Kapanidis, A. N., Tinnefeld, P., Wang, Y., Mukhopadhyay, J., Ebright, R. H., and Weiss, S. Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes. *Biophysical journal*, 90(4):1419–1431, 2006.

Marr, A. Growth rate of *escherichia coli*. *Microbiological reviews*, 55(2):316–333, 1991.

Masucci, J., Gallant, J., Lindsley, D., and Atkinson, J. Influence of the rela gene on ribosome frameshifting. *Molecular Genetics and Genomics*, 268(1):81–86, 2002.

Mathew, R. and Chatterji, D. The evolving story of the omega subunit of bacterial rna polymerase. *Trends in microbiology*, 14(10):450–455, 2006.

Médigue, C., Rose, M., Viari, A., and Danchin, A. Detecting and analyzing dna sequencing errors: toward a higher quality of the *Bacillus subtilis* genome sequence. *Genome research*, 9(11):1116–1127, 1999.

Mejlhede, N., Atkins, J., and Neuhard, J. Ribosomal -1 frameshifting during decoding of*Bacillus subtilis cdd* occurs at the sequence cga aag. *Journal of bacteriology*, 181(9):2930–2937, 1999.

Meyer, F. M., Jules, M., Mehne, F. M., Le Coq, D., Landmann, J. J., Görke, B., Aymerich, S., and Stülke, J. Malate-mediated carbon catabolite repression in *Bacillus subtilis* involves the hprk/ccpa pathway. *Journal of bacteriology*, 193(24):6939–6949, 2011.

Meyerovich, M., Mamou, G., and Ben-Yehuda, S. Visualizing high error levels during gene expression in living bacterial cells. *Proceedings of the National Academy of Sciences*, 107(25):11543–11548, 2010.

Miller, D. L., Cashel, M., and Weissbach, H. The interaction of guanosine 5'-diphosphate, 2'(3')-diphosphate with the bacterial elongation factor tu. *Archives of biochemistry and biophysics*, 154 (2):675–682, 1973.

Milon, P., Tischenko, E., Tomšic, J., Caserta, E., Folkers, G., La Teana, A., Rodnina, M. V., Pon, C. L., Boelens, R., and Gualerzi, C. O. The nucleotide-binding site of bacterial translation initiation factor 2 (if2) as a metabolic sensor. *Proceedings of the National Academy of Sciences*, 103(38):13962–13967, 2006.

Milon, P., Konevega, A. L., Gualerzi, C. O., and Rodnina, M. V. Kinetic checkpoint at a late step in translation initiation. *Molecular cell*, 30(6):712–720, 2008.

Milón, P., Maracci, C., Filonava, L., Gualerzi, C. O., and Rodnina, M. V. Real-time assembly landscape of bacterial 30s translation initiation complex. *Nature Structural and Molecular Biology*, 19(6):609, 2012.

Mitkevich, V. A., Ermakov, A., Kulikova, A. A., Tankov, S., Shyp, V., Soosaar, A., Tenson, T., Makarov, A. A., Ehrenberg, M., and Hauryliuk, V. Thermodynamic characterization of ppgpp binding to ef-g or if2 and of initiator trna binding to free if2 in the presence of gdp, gtp, or ppgpp. *Journal of molecular biology*, 402(5):838–846, 2010.

Molenaar, D., van Berlo, R., de Ridder, D., and Teusink, B. Shifts in growth strategies reflect tradeoffs in cellular economics. *Molecular systems biology*, 5(1), 2009.

Molle, V., Nakaura, Y., Shivers, R. P., Yamaguchi, H., Losick, R., Fujita, Y., and Sonenshein, A. L. Additional targets of the *Bacillus subtilis* global regulator cody identified by chromatin immunoprecipitation and genome-wide transcript analysis. *Journal of bacteriology*, 185(6):1911–1922, 2003.

Monod, J. The growth of bacterial cultures. *Annual Reviews in Microbiology*, 3(1):371–394, 1949.

Moreno-Campuzano, S., Janga, S. C., and Pérez-Rueda, E. Identification and analysis of dna-binding transcription factors in *Bacillus subtilis* and other firmicutes-a genomic approach. *BMC genomics*, 7(1):147, 2006.

Mukherjee, K., Nagai, H., Shimamoto, N., and Chatterji, D. Groel is involved in activation of *Escherichia coli* rna polymerase devoid of the ω subunit in vivo. *The FEBS Journal*, 266(1):228–235, 1999.

Murakami, K. S. and Darst, S. A. Bacterial rna polymerases: the whole story. *Current opinion in structural biology*, 13(1):31–39, 2003.

Murakami, K. Structural biology of bacterial rna polymerase. *Biomolecules*, 5(2):848–864, 2015.

Nanamiya, H., Kasai, K., Nozawa, A., Yun, C.-S., Narisawa, T., Murakami, K., Natori, Y., Kawamura, F., and Tozawa, Y. Identification and functional analysis of novel (p)ppgpp synthetase genes in *Bacillus subtilis*. *Molecular microbiology*, 67(2):291–304, 2008.

Näsvall, S., Nilsson, K., and Björk, G. The ribosomal grip of the peptidyl-trna is critical for reading frame maintenance. *Journal of molecular biology*, 385(2):350–367, 2009.

Natori, Y., Tagami, K., Murakami, K., Yoshida, S., Tanigawa, O., Moh, Y., Masuda, K., Wada, T., Suzuki, S., Nanamiya, H., et al. Transcription activity of individual *rrn* operons in *Bacillus subtilis* mutants deficient in (p)ppgpp synthetase genes, *relA*, *yjbM*, and *ywaC*. *Journal of bacteriology*, 191 (14):4555–4561, 2009.

Neidhardt, F. Roles of amino acid activating enzymes in cellular physiology. *Bacteriological reviews*, 30(4):701, 1966.

Neidhardt, F. and Magasanik, B. Studies on the role of ribonucleic acid in the growth of bacteria. *Biochimica et biophysica acta*, 42:99–116, 1960.

Newton, A. Isolation and characterization of frameshift mutations in the *lac* operon. *Journal of molecular biology*, 49(3):589–601, 1970.

Nicolas, P., Mäder, U., Dervyn, E., Rochat, T., Leduc, A., Pigeonneau, N., Bidnenko, E., Marchadier, E., Hoebeke, M., Aymerich, S., et al. Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science*, 335(6072):1103–1106, 2012.

Nomura, T., Ishihama, A., Kajitani, M., Takahashi, T., Nakada, N., and Yoshinaga, K. Promoter selectivity of *Escherichia coli* rna polymerase. *Molecular and General Genetics MGG*, 193(1):8–16, 1984.

Nudler, E., Avetissova, E., Markovtsov, V., and Goldfarb, A. Transcription processivity: protein-dna interactions holding together the elongation complex. *Science*, 273(5272):211–217, 1996.

O'Brien, E. J., Utrilla, J., and Palsson, B. O. Quantification and classification of *E. coli* proteome utilization and unused protein costs across environments. *PLoS computational biology*, 12(6):e1004998, 2016.

O'Farrell, P. The suppression of defective translation by ppgpp and its role in the stringent response. *Cell*, 14(3):545–557, 1978.

Ogle, J. M., Brodersen, D. E., Clemons, W. M., Tarry, M. J., Carter, A. P., and Ramakrishnan, V. Recognition of cognate transfer rna by the 30s ribosomal subunit. *Science*, 292(5518):897–902, 2001.

Ogle, J. M., Murphy IV, F. V., Tarry, M. J., and Ramakrishnan, V. Selection of trna by the ribosome requires a transition from an open to a closed form. *Cell*, 111(5):721–732, 2002.

Olubajo, B. and Taylor, E. A -1 frameshift in the hiv-1 *env* gene is enhanced by arginine deficiency via a hungry codon mechanism. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 579(1):125–132, 2005.

Orth, J., Thiele, I., and Palsson, B. What is flux balance analysis? *Nature biotechnology*, 28(3):245, 2010.

Overkamp, W., Beilharz, K., Weme, R., Solopova, A., Karsens, H., Kovács, A., Kok, J., Kuipers, O., and Veening, J.-W. Benchmarking various green fluorescent protein variants in *Bacillus subtilis*, *Streptococcus pneumoniae*, and *Lactococcus lactis* for live cell imaging. *Applied and environmental microbiology*, 79(20):6481–6490, 2013.

Pande, S., Vimaladithan, A., Zhao, H., and Farabaugh, P. Pulling the ribosome out of frame by +1 at a programmed frameshift site by cognate binding of aminoacyl-trna. *Molecular and cellular biology*, 15(1):298–304, 1995.

Pao, C. C. and Dyes, B. T. Effect of unusual guanosine nucleotides on the activities of some *escherichia coli* cellular enzymes. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 677(3-4):358–362, 1981.

Park, M., Jand Dies, Lin, Y., Hormoz, S., Smith-Unna, S., Quinodoz, M., Sand Hernández-Jiménez, Garcia-Ojalvo, J., Locke, J., and Elowitz, M. Molecular time sharing through dynamic pulsing in single cells. *Cell systems*, 2018.

Parker, J. Errors and alternatives in reading the universal genetic code. *Microbiological reviews*, 53 (3):273, 1989.

Partridge, S. and Errington, J. The importance of morphological events and intercellular interactions in the regulation of prespore-specific gene expression during sporulation in *Bacillus subtilis*. *Molecular microbiology*, 8(5):945–955, 1993.

Paul, B. J., Barker, M. M., Ross, W., Schneider, D. A., Webb, C., Foster, J. W., and Gourse, R. L. Dksa: a critical component of the transcription initiation machinery that potentiates the regulation of rrna promoters by ppgpp and the initiating ntp. *Cell*, 118(3):311–322, 2004a.

Paul, B. J., Ross, W., Gaal, T., and Gourse, R. L. *rRNA* transcription in *escherichia coli*. *Annu. Rev. Genet.*, 38:749–770, 2004b.

Paul, B. J., Berkmen, M. B., and Gourse, R. L. Dksa potentiates direct activation of amino acid promoters by ppgpp. *Proceedings of the National Academy of Sciences*, 102(22):7823–7828, 2005.

Peters, J. M., Vangeloff, A. D., and Landick, R. Bacterial transcription terminators: the rna 3'-end chronicles. *Journal of molecular biology*, 412(5):793–813, 2011.

Pon, C. L. and Gualerzi, C. O. Mechanism of protein biosynthesis in prokaryotic cells: effect of initiation factor if1 on the initial rate of 30 s initiation complex formation. *FEBS letters*, 175(2): 203–207, 1984.

Potrykus, K. and Cashel, M. (p)ppgpp: still magical? *Annual review of microbiology*, 62, 2008.

Ramakrishnan, V. Ribosome structure and the mechanism of translation. *Cell*, 108(4):557–572, 2002.

Rao, F., See, R. Y., Zhang, D., Toh, D. C., Ji, Q., and Liang, Z.-X. Yybt is a signaling protein that contains a cyclic dinucleotide phosphodiesterase domain and a ggdef domain with atpase activity. *Journal of Biological Chemistry*, 285(1):473–482, 2010.

Ratkowsky, D., Lowry, R., McMeekin, T., Stokes, A., and Chandler, R. Model for bacterial culture growth rate throughout the entire biokinetic temperature range. *Journal of bacteriology*, 154(3): 1222–1226, 1983.

Ratnayake-Lecamwasam, M., Serror, P., Wong, K.-W., and Sonenshein, A. L. *Bacillus subtilis* cody represses early-stationary-phase genes by sensing gtp levels. *Genes & development*, 15(9):1093–1103, 2001.

Richardson, J. P. Loading rho to terminate transcription. *Cell*, 114(2):157–159, 2003.

Roberts, M. W. and Rabinowitz, J. The effect of *Escherichia coli* ribosomal protein s1 on the translational specificity of bacterial ribosomes. *Journal of Biological Chemistry*, 264(4):2228–2235, 1989.

Robinson, V. L. et al. Salmonella enterica serovar typhimurium bipa exhibits two distinct ribosome binding modes. *Journal of bacteriology*, 190(17):5944–5952, 2008.

Rocha, E. P., Danchin, A., and Viari, A. Translation in *Bacillus subtilis*: roles and trends of initiation and termination, insights from a genome analysis. *Nucleic acids research*, 27(17):3567–3576, 1999.

Roghanian, M., Zenkin, N., and Yuzenkova, Y. Bacterial global regulators dksa/ppgpp increase fidelity of transcription. *Nucleic acids research*, 43(3):1529–1536, 2015.

Rojas, A. and Ehrenberg, M. How does ppgpp affect translational accuracy in the stringent response? *Biochimie*, 73(5):599–605, 1991.

Rojas, A.-M., Ehrenberg, M., Andersson, S. G., and Kurland, C. ppgpp inhibition of elongation factors tu, g and ts during polypeptide synthesis. *Molecular and General Genetics MGG*, 197(1): 36–45, 1984.

Ross, W., Vrentas, C. E., Sanchez-Vazquez, P., Gaal, T., and Gourse, R. L. The magic spot: a ppgpp binding site on *e. coli* rna polymerase responsible for regulation of transcription initiation. *Molecular cell*, 50(3):420–429, 2013.

Ross, W., Sanchez-Vazquez, P., Chen, A. Y., Lee, J.-H., Burgos, H. L., and Gourse, R. L. ppgpp binding to a site at the rnap-dksa interface accounts for its dramatic effects on transcription initiation during the stringent response. *Molecular cell*, 62(6):811–823, 2016.

Rozov, A., Demeshkina, N., Westhof, E., Yusupov, M., and Yusupova, G. New structural insights into translational miscoding. *Trends in biochemical sciences*, 41(9):798–814, 2016.

Russell, A. Bacterial adaptation and resistance to antiseptics, disinfectants and preservatives is not a new phenomenon. *Journal of Hospital Infection*, 57(2):97–104, 2004.

Schaaper, R. M. Base selection, proofreading, and mismatch repair during dna replication in *Escherichia coli*. *Journal of Biological Chemistry*, 268(32):23762–23765, 1993.

Schaechter, M., Maaløe, O., and Kjeldgaard, N. O. Dependency on medium and temperature of cell size and chemical composition during balanced growth of *Salmonella typhimurium*. *Microbiology*, 19(3):592–606, 1958.

Schaechter, M., Ingraham, J. L., Neidhardt, F. C., et al. *Microbe*. ASM press, 2006.

Schallmey, M., Singh, A., and Ward, O. P. Developments in the use of *Bacillus* species for industrial production. *Canadian journal of microbiology*, 50(1):1–17, 2004.

Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. Interdependence of cell growth and gene expression: origins and consequences. *Science*, 330(6007):1099–1102, 2010.

Shivers, R. P. and Sonenshein, A. L. Activation of the *Bacillus subtilis* global regulator cody by direct interaction with branched-chain amino acids. *Molecular microbiology*, 53(2):599–611, 2004.

Shoemaker, C. and Green, R. Translation drives mrna quality control. *Nature Structural and Molecular Biology*, 19(6):594, 2012.

Slack, F., Mueller, J., and Sonenshein, A. Mutations that relieve nutritional repression of the *Bacillus subtilis* dipeptide permease operon. *Journal of bacteriology*, 175(15):4605–4614, 1993.

Slack, F. J., Serror, P., Joyce, E., and Sonenshein, A. L. A gene required for nutritional repression of the *Bacillus subtilis* dipeptide permease operon. *Molecular microbiology*, 15(4):689–702, 1995.

Sonenshein, A. L. Cody, a global regulator of stationary phase and virulence in gram-positive bacteria. *Current opinion in microbiology*, 8(2):203–207, 2005.

Sonenshein, A. L., Hoch, J. A., and Losick, R. *bacillus subtilis*: from cells to genes and from genes to cells. In *Bacillus subtilis and Its Closest Relatives*, pages 3–5. American Society of Microbiology, 2002.

Sonenshein, A. Control of key metabolic intersections in *Bacillus subtilis*. *nature reviews microbiology*, 5(12):917, 2007.

Sørensen, M. Charging levels of four trna species in *Escherichia coli* rel+ and rel- strains during amino acid starvation: a simple model for the effect of ppgpp on translational accuracy. *Journal of molecular biology*, 307(3):785–798, 2001.

Sørensen, M., Jensen, K., and Pedersen, S. High concentrations of ppgpp decrease the rna chain growth rate: implications for protein synthesis and translational fidelity during amino acid starvation in *Escherichia coli*. *Journal of molecular biology*, 236(2):441–454, 1994.

Spizizen, J. Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. *Proceedings of the National Academy of Sciences*, 44(10):1072–1078, 1958.

Steinchen, W. and Bange, G. The magic dance of the alarmones (p)ppgpp. *Molecular microbiology*, 101(4):531–544, 2016.

Steinchen, W., Schuhmacher, J. S., Altegoer, F., Fage, C. D., Srinivasan, V., Linne, U., Marahiel, M. A., and Bange, G. Catalytic mechanism and allosteric regulation of an oligomeric (p)ppgpp synthetase by an alarmone. *Proceedings of the National Academy of Sciences*, 112(43):13348–13353, 2015.

Steinchen, W., Vogt, M. S., Altegoer, F., Giammarinaro, P. I., Horvatek, P., Wolz, C., and Bange, G. Structural and mechanistic divergence of the small (p)ppgpp synthetases relp and relq. *Scientific reports*, 8(1):2195, 2018.

Stephens, C. Bacterial sporulation: a question of commitment? *Current biology*, 8(2):R45–R48, 1998.

Stewart, G. C., Wilson, F. E., and Bott, K. F. Detailed physical mapping of the ribosomal rna genes of *Bacillus subtilis*. *Gene*, 19(2):153–162, 1982.

Subramaniam, A., Zid, B., and O'Shea, E. An integrated approach reveals regulatory controls on bacterial translation elongation. *Cell*, 159(5):1200–1211, 2014.

Suh, J., Boylan, S. A., and Price, C. W. Gene for the alpha subunit of *Bacillus subtilis* rna polymerase maps in the ribosomal protein gene cluster. *Journal of bacteriology*, 168(1):65–71, 1986.

Sundararajan, A., Michaud, W., Qian, Q., Stahl, G., and Farabaugh, P. Near-cognate peptidyl-trnas promote +1 programmed translational frameshifting in yeast. *Molecular cell*, 4(6):1005–1015, 1999.

Svenningsen, S., Kongstad, M., Stenum, T. S., Muñoz-Gómez, A., and Sørensen, M. Transfer rna is highly unstable during early amino acid starvation in *escherichia coli*. *Nucleic acids research*, 45 (2):793–804, 2016.

Tagami, K., Nanamiya, H., Kazo, Y., Maehashi, M., Suzuki, S., Namba, E., Hoshiya, M., Hanai, R., Tozawa, Y., Morimoto, T., et al. Expression of a small (p)ppgpp synthetase, ywac, in the (p)ppgpp0 mutant of *Bacillus subtilis* triggers yvyd-dependent dimerization of ribosome. *MicrobiologyOpen*, 1 (2):115–134, 2012.

Taheri-Araghi, S., Bradde, S., Sauls, J. T., Hill, N. S., Levin, P. A., Paulsson, J., Vergassola, M., and Jun, S. Cell-size control and homeostasis in bacteria. *Current Biology*, 25(3):385–391, 2015.

Tanaka, K., Henry, C., Zinner, J., Jolivet, E., Cohoon, M., Xia, V., Fand Bidnenko, Ehrlich, S., Stevens, R., and Noirot, P. Building the repertoire of dispensable chromosome regions in *Bacillus subtilis* entails major refinement of cognate large-scale metabolic model. *Nucleic acids research*, 41 (1):687–699, 2012.

Thackray, P. D. and Moir, A. Sigm, an extracytoplasmic function sigma factor of *Bacillus subtilis*, is activated in response to cell wall antibiotics, ethanol, heat, acid, and superoxide stress. *Journal of bacteriology*, 185(12):3491–3498, 2003.

Timmusk, S. and Wagner, E. G. H. The plant-growth-promoting rhizobacterium paenibacillus polymyxa induces changes in arabidopsis thaliana gene expression: a possible connection between biotic and abiotic stress responses. *Molecular Plant-Microbe Interactions*, 12(11):951–959, 1999.

Tojo, S., Satomura, T., Kumamoto, K., Hirooka, K., and Fujita, Y. Molecular mechanisms underlying the positive stringent response of the *Bacillus subtilis ilv-leu* operon, involved in the biosynthesis of branched-chain amino acids. *Journal of bacteriology*, 190(18):6134–6147, 2008.

Tojo, S., Kumamoto, K., Hirooka, K., and Fujita, Y. Heavy involvement of stringent transcription control depending on the adenine or guanine species of the transcription initiation site in glucose and pyruvate metabolism in bacillus subtilis. *Journal of bacteriology*, 192(6):1573–1585, 2010.

Tortora, G., Funke, B., Case, C., and Johnson, T. *Microbiology: an introduction*, volume 9. Benjamin Cummings San Francisco, CA, 2004.

Toulokhonov, I. I., Shulgina, I., and Hernandez, V. J. Binding of the transcription effector ppgpp to *escherichia coli* rna polymerase is allosteric, modular, and occurs near the n terminus of the β'-subunit. *Journal of Biological Chemistry*, 276(2):1220–1225, 2001.

Traverse, C. and Ochman, H. Conserved rates and patterns of transcription errors across bacterial growth states and lifestyles. *Proceedings of the National Academy of Sciences*, 113(12):3311–3316, 2016.

Trueba, F. J. and Woldringh, C. L. Changes in cell diameter during the division cycle of *escherichia coli*. *Journal of bacteriology*, 142(3):869–878, 1980.

Tsai, T., Yang, H., Yin, H., Xu, S., and Wang, Y. High-efficiency "-1" and "-2" ribosomal frameshiftings revealed by force spectroscopy. *ACS chemical biology*, 12(6):1629–1635, 2017.

Tsuchihashi, Z. and Kornberg, A. Translational frameshifting generates the gamma subunit of dna polymerase iii holoenzyme. *Proceedings of the National Academy of Sciences*, 87(7):2516–2520, 1990.

Ude, S., Lassak, J., Starosta, A., Kraxenberger, T., Wilson, D., and Jung, K. Translation elongation factor ef-p alleviates ribosome stalling at polyproline stretches. *Science*, 339(6115):82–85, 2013.

Urbonavicius, J., StahL, G., Durand, J., Salem, S., Qian, Q., Farabaugh, P., and Björk, G. Transfer rna modifications that alter +1 frameshifting in general fail to affect -1 frameshifting.

Urbonavičius, J., Qian, Q., Durand, J., Hagervall, T., and Björk, G. Improvement of reading frame maintenance is a common function for several trna modifications. *The EMBO journal*, 20(17): 4863–4873, 2001.

van Dijl, J. and Hecker, M. *Bacillus subtilis*: from soil bacterium to super-secreting cell factory, 2013.

Vassylyev, D. G., Vassylyeva, M. N., Perederina, A., Tahirov, T. H., and Artsimovitch, I. Structural basis for transcription elongation by bacterial rna polymerase. *Nature*, 448(7150):157, 2007.

Vassylyeva, M. N., Lee, J., Sekine, S. i., Laptenko, O., Kuramitsu, S., Shibata, T., Inoue, Y., Borukhov, S., Vassylyev, D. G., and Yokoyama, S. Purification, crystallization and initial crystallographic analysis of rna polymerase holoenzyme from *Thermus thermophilus*. *Acta Crystallographica Section D: Biological Crystallography*, 58(9):1497–1500, 2002.

Wang, F. and Greene, E. C. Single-molecule studies of transcription: from one rna polymerase at a time to the gene expression profile of a cell. *Journal of molecular biology*, 412(5):814–831, 2011.

Wang, J. D., Sanders, G. M., and Grossman, A. D. Nutritional control of elongation of dna replication by (p)ppgpp. *Cell*, 128(5):865–875, 2007.

Weber, H. J. Stoichiometric measurements of 30s and 50s ribosomal proteins from *Escherichia coli*. *Molecular and General Genetics MGG*, 119(3):233–248, 1972.

Weiss, R., Dunn, D., Atkins, J., and Gesteland, R. Slippery runs, shifty stops, backward steps, and forward hops: -2, -1, +1, +2, +5, and +6 ribosomal frameshifting. In *Cold Spring Harbor Symposia on Quantitative Biology*, volume 52, pages 687–693. Cold Spring Harbor Laboratory Press, 1987.

Weiße, A. Y., Oyarzún, D. A., Danos, V., and Swain, P. S. Mechanistic links between cellular trade-offs, gene expression, and growth. *Proceedings of the National Academy of Sciences*, 112(9): E1038–E1047, 2015.

Wendrich, T. M. and Marahiel, M. A. Cloning and characterization of a rela/spot homologue from *Bacillus subtilis*. *Molecular microbiology*, 26(1):65–79, 1997.

Wendrich, T. M., Blaha, G., Wilson, D. N., Marahiel, M. A., and Nierhaus, K. H. Dissection of the mechanism for the stringent factor rela. *Molecular cell*, 10(4):779–788, 2002.

Westhof, E., Yusupov, M., and Yusupova, G. Recognition of watson-crick base pairs: constraints and limits due to geometric selection and tautomerism. *F1000prime reports*, 6, 2014.

Wilson, D. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nature Reviews Microbiology*, 12(1):35, 2014.

Winther, K., Roghanian, M., and Gerdes, K. Activation of the stringent response by loading of rela-trna complexes at the ribosomal a-site. *Molecular cell*, 70(1):95–105, 2018.

Wolf, M., Müller, T., Dandekar, T., and Pollack, J. D. Phylogeny of firmicutes with special reference to mycoplasma (mollicutes) as inferred from phosphoglycerate kinase amino acid sequence data. *International journal of systematic and evolutionary microbiology*, 54(3):871–875, 2004.

Yan, S., Wen, J., Bustamante, C., and Tinoco Jr, I. Ribosome excursions during mrna translocation mediate broad branching of frameshift pathways. *Cell*, 160(5):870–881, 2015.

Yang, L., Yurkovich, J. T., Lloyd, C. J., Ebrahim, A., Saunders, M. A., and Palsson, B. O. Principles of proteome allocation are revealed using proteomic data and genome-scale models. *Scientific reports*, 6:36734, 2016.

Yang, L., Yurkovich, J. T., King, Z. A., and Palsson, B. O. Modeling the multi-scale mechanisms of macromolecular resource allocation. *Current opinion in microbiology*, 45:8–15, 2018.

Zhang, Y., Mooney, R. A., Grass, J. A., Sivaramakrishnan, P., Herman, C., Landick, R., and Wang, J. D. Dksa guards elongating rna polymerase against ribosome-stalling-induced arrest. *Molecular cell*, 53(5):766–778, 2014.

Zhu, B. and Stülke, J. Subti wiki in 2018: from genes and proteins to functional network annotation of the model organism *Bacillus subtilis*. *Nucleic acids research*, 46(D1):D743–D748, 2017.

Zucker, F. H. and Hershey, J. W. Binding of *escherichia coli* protein synthesis initiation factor if1 to 30s ribosomal subunits measured by fluorescence polarization. *Biochemistry*, 25(12):3682–3690, 1986.

Zuo, Y., Wang, Y., and Steitz, T. A. The mechanism of *e. coli* rna polymerase regulation by ppgpp is suggested by the structure of their complex. *Molecular cell*, 50(3):430–436, 2013.

Zweers, J. C., Nicolas, P., Wiegert, T., van Dijl, J. M., and Denham, E. L. Definition of the σw regulon of *Bacillus subtilis* in the absence of stress. *PloS one*, 7(11):e48471, 2012.

université PARIS-SACLAY | ÉCOLE DOCTORALE Structure et dynamique des systèmes vivants (SDSV)

Titre : Comprendre comment les métabolites, GTP et (p)ppGpp, contrôlent simultanément l'apparition d'erreurs traductionnelles et l'allocation des ressources chez les bactéries.

Mots clés : Erreur traductionnelle, Réponse stringente, GTP et (p)ppGpp, Bactéries, Adaptation cellulaire, Allocation des ressources

Résumé : Bien que divers mécanismes coopèrent pour empêcher les erreurs lors de la synthèse des protéines chez les bactéries, des erreurs traductionnelles de type « frameshift » (ETFs) ou « faux-sens » peuvent avoir lieu. En particulier, les ETFs ont été détectées à de faibles niveaux lors de la phase de croissance exponentielle et à des niveaux plus élevés durant la phase de croissance stationnaire chez *Escherichia coli* et *Bacillus subtilis*. Ces observations ont conduit les chercheurs à revoir le rôle de la "réponse stringente" dans la survenue des ETFs, qui constitue l'un des mécanismes clé de l'adaptation bactérienne aux changements nutritionnels. Elle découle de l'interaction entre un ribosome en cours de traduction et la protéines RelA/SpoT ce qui permet de détecter les ARNs de transfert (ARNts) non chargés et résulte en la production d'une molécule appelée (p)ppGpp. Dans une souche mutante *relA* incapable de synthétiser le (p)ppGpp, les ETFs sont fortement augmentées.

Dans ce contexte, notre objectif principal a été de revisiter le rôle de la réponse stringente dans le contrôle des erreurs traductionnelles et de clarifier le rôle des deux métabolites antagonistes GTP et (p)ppGpp. Par exemple, le GTP stimule l'initiation de la traduction (en ciblant le facteur d'initiation IF2) alors que le (p)ppGpp inhibe l'initiation de la traduction (en rentrant en concurrence avec le GTP pour se fixer sur IF2).

A cette fin, nous avons utilisé le modèle des bactéries à Gram positif *B. subtilis*, conçu trois systèmes rapporteurs distincts pour détecter les ETFs et construit une souche incapable de synthétiser du (p)ppGpp (appelée "(p)ppGpp^0"). Nous avons observé qu'au cours de la croissance dans des milieux pauvres, les ETFs augmentent en l'absence de (p)ppGpp durant la phase exponentielle et que, contrairement à la souche sauvage, la souche (p)ppGpp^0 présente un pic d'ETFs en milieu riche pendant la transition à la phase stationnaire. En contrôlant les niveaux intracellulaires de GTP dans la souche (p)ppGpp^0, nous avons montré que l'abondance de GTP est le facteur qui déclenche l'apparition des ETFs. Néanmoins, après une "faible" induction de la biosynthèse du GTP conduisant à des taux de croissance sous-optimaux, le niveau d'ETFs forme toujours un pic lors de la transition vers la phase stationnaire, ce qui montre que le mode d'action du (p)ppGpp pour prévenir l'apparition des ETFs ne repose pas uniquement sur son action inhibitrice de la biosynthèse du GTP. Nous nous sommes alors concentrés sur l'effet inhibiteur du (p)ppGpp sur IF2 et avons mimé son action en injectant des drogues connues pour inhiber l'initiation de la traduction. Nous avons ainsi démontré qu'en réduisant l'initiation de la traduction lors de l'épuisement des aminoacyl-ARNts, la souche "(p)ppGpp^0" est capable de contrôler de façon optimale le taux d'ETFs lors de la transition vers la phase stationnaire.

Dans une deuxième partie, nous avons étudié comment la transcription et la traduction sont affectées par les variations du niveau de GTP et de (p)ppGpp. Nous avons observé que les gènes possédant un "+1" de transcription (TSS, « transcription start site ») composé de deux guanines (gènes artificiels et ARNs ribosomaux) ont vu leur taux de transcription positivement corrélés au taux de croissance à l'inverse des gènes possédant un TSS composé de deux adénines. Cette différence est encore plus prononcée pour la souche (p)ppGpp^0 cultivée en milieu riche lors de l'ajout de guanosine (ce qui conduit à un niveau élevé de GTP).

En conclusion, nous avons démontré que le (p)ppGpp contrôle le niveau d'erreurs traductionnelles lors de la croissance en régime permanent en abaissant les niveaux de GTP et lors d'un changement nutritionnel en inhibant spécifiquement l'initiation de la traduction, assurant une allocation parcimonieuse des ressources au sein de la bactérie.

université PARIS-SACLAY | ÉCOLE DOCTORALE Structure et dynamique des systèmes vivants (SDSV)

Title: How do the metabolites, GTP and (p)ppGpp, simultaneously control the occurrence of translational errors and resource allocation in bacteria?

Keywords: Translational error, Stringent response, GTP and (p)ppGpp, Bacteria, Cell adaptation, Resource allocation

Abstract: Even though diverse mechanisms cooperate to prevent protein synthesis errors in bacteria, missense and translational frameshift errors (TFEs) can occur. In particular, TFEs were detected at low levels in the exponential growth phase and at higher levels in the stationary phase in both *Escherichia coli* and *Bacillus subtilis*. This observation led researchers to revisit the role of the "stringent response" in the occurrence of TFEs since it is the key mechanism involved in the bacterial adaptation to nutritional downshifts. It relies on the interaction between the RelA/SpoT proteins and the translating ribosomes, which leads to the detection of uncharged tRNAs and to the production of an alarmone called (p)ppGpp. In a *relA* mutant strains unable to synthesize (p)ppGpp, translational errors are highly increased.

In this context, the main goal of our work was to revisit the role of the stringent response in the translational error control and to clarify the role of the two key, antagonistic metabolites GTP and (p)ppGpp. Indeed, while GTP enhances translation initiation (targeting the initiation factor IF2), (p)ppGpp inhibits GTP biosynthesis and translation initiation (competing with GTP on IF2).

For this purpose, we used the Gram positive model bacterium *B. subtilis*, designed three distinct reporter systems to detect TFEs and built a strain unable to synthesize (p)ppGpp (called "(p)ppGpp^{0}"). We observed that during growth in poor media TFEs were increased in the absence of (p)ppGpp in the exponential phase (i.e. steady-state growth) and that by contrast to the wild type, the (p)ppGpp^{0} strain exhibited a TFE burst during the transition in rich medium to the stationary phase. By controlling intracellular levels of GTP in the (p)ppGpp^{0} strain, we showed that GTP abundance is the trigger factor of TFEs occurrence. Nevertheless, upon a "weak" induction of GTP biosynthesis leading to sub-optimal growth rates, the TFEs rate still peaked during the transition to the stationary phase, which demonstrated that the mode of action of (p)ppGpp to prevent TFEs occurrence did not only rely on its inhibition of GTP biosynthesis. We then focused on the (p)ppGpp inhibitory effect on IF2 and mimicked its action by injecting drugs known to inhibit translation initiation. Hence, we demonstrated that by reducing translation initiation (injecting drugs) upon aminoacyl-tRNAs depletion (p)ppGpp^{0} strain is able to control the rate of TFEs in the transition to the stationary phase.

In a second part, we studied how transcription and translation are affected by variations in GTP and (p)ppGpp abundances. We observed that genes possessing a transcription start site (TSS) made of two guanines were more importantly transcribed at higher growth rates than genes possessing a TSS made of two adenines. This difference was even more pronounced for (p)ppGpp^{0} strains grown in rich medium upon guanosine addition (leading to a high level of GTP). Moreover, the ribosomal RNAs (rrns; for which the TSS is a guanine) synthesis level seemed to be positively correlated to GTP levels during exponential growth in poor and rich media as observed by the modulation of GTP biosynthesis.

In conclusion, we demonstrated that (p)ppGpp controls the occurrence of translational errors during steady-state growth by decreasing GTP levels and during a nutritional downshift by specifically inhibiting translation initiation ensuring a parsimonious resource allocation.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France