TUDelft

Delft University of Technology

Dynamical regulation in single cells

Wehrens, Martijn

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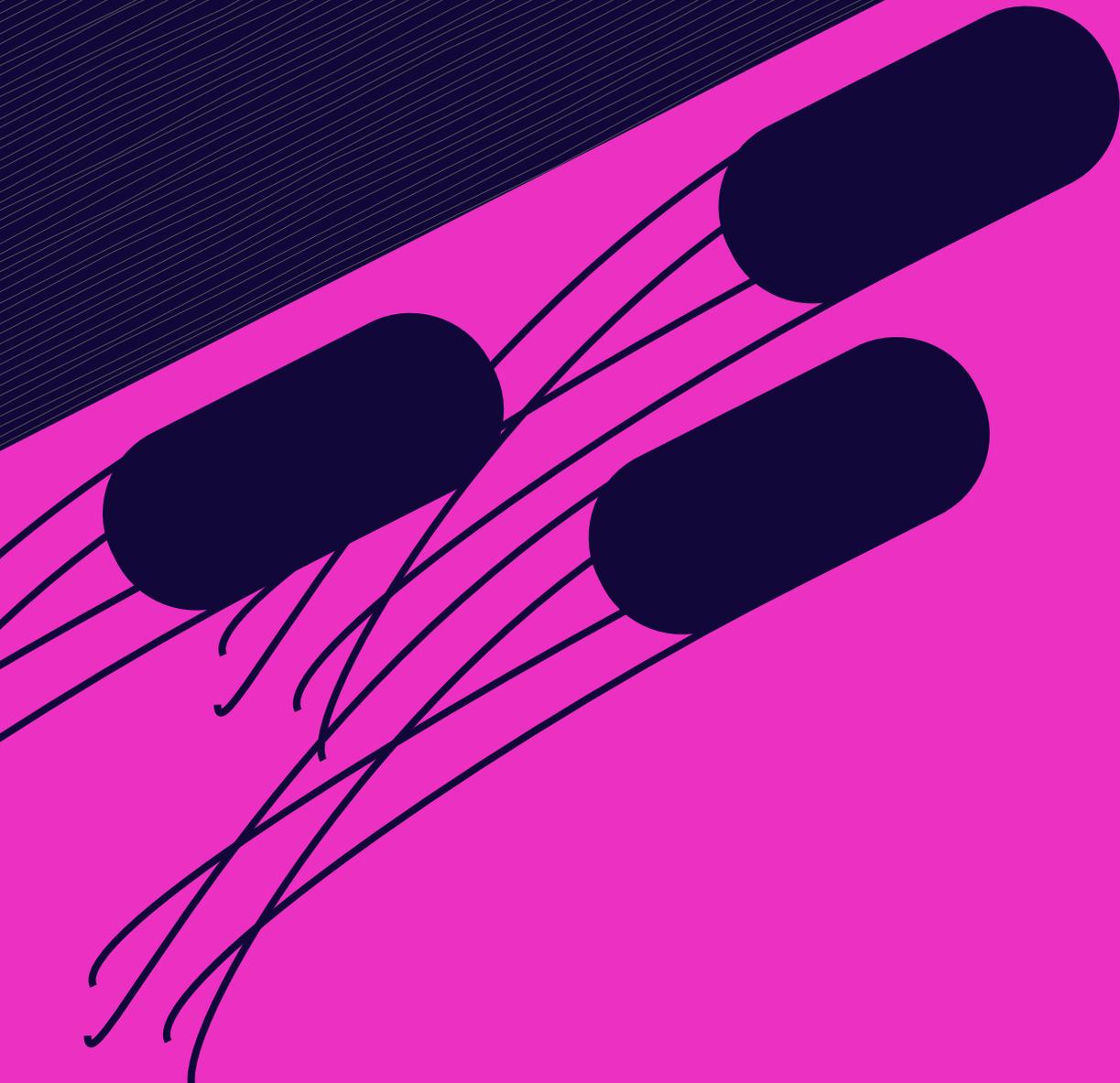Please check the document version above.

DYNAMICAL REGULATION IN SINGLE CELLS

A PHD THESIS BY MARTIJN WEHRENS

DYNAMICAL REGULATION IN SINGLE CELLS

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus Prof.dr.ir. T. H. J. J. van der Hagen;
Chair of the Board for Doctorates
to be defended publicly on
Monday 4 March 2019 at 12:30 o'clock

by

Martijn WEHRENS

Master of Science in Chemistry, University of Amsterdam, the Netherlands
born in Nijmegen, the Netherlands

This dissertation has been approved by the promotor.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. ir. S. J. Tans	Delft University of Technology, *promotor*

Independent members:

Prof.dr.ir. J.J. Heijnen	Delft University of Technology
Prof. P. Nghe	ESPCI Paris, France
Prof.dr. L.W. Hamoen	University of Amsterdam
Prof.dr. F.J. Bruggeman	Vrije Universiteit Amsterdam
Prof.dr. P.R. ten Wolde	Vrije Universiteit Amsterdam
Dr. R. Hermsen	Utrecht University
Prof.dr. N.H. Dekker	Delft University of Technology, substitute member

Contents

SUMMARY

In this thesis, we probe single bacterial cells to further understand both the regulation of cell divisions during adverse conditions and the phenomenon of cellular heterogeneity.

Firstly, in introductory **chapter 1** we provide an overview of the topics in this thesis and provide some context for the layman. Then, in **chapter 2** we describe how we expanded on current methods to investigate the behaviour of single *Escherichia coli* cells: we discuss the PDMS device that we used to subject microcolonies of cells to changes in growth medium and observe them, the software expansions that simplified and enabled analysis of our single cell time lapse data, and the application of cross-correlation analysis to branched lineage data that is acquired by observing growing microcolonies of cells (which is relevant for studying heterogeneity).

In **chapter 3** we discuss novel findings regarding the regulation of cell divisions during adverse conditions. Instead of growing *E. coli* in favorable conditions, as is often done in the lab, we subjected the bacteria to adverse real-world conditions like antibiotic exposure and high temperature, The cells responded to these adverse conditions by halting the cell division process, whilst they continued to grow. This process is called filamentation and it leads to characteristic elongated cell morphologies. When we switched growth conditions back to favorable conditions, the cells started dividing again to eventually recover their normal bacillary form. Remarkably, during this process, the cells continuously re-arranged potential division sites (Fts rings), to place them at specific locations along the cellular axis. Where Fts rings formed depended on the length of the individual bacteria, and we showed that this placement was regulated by Min proteins, which acted as a dynamical ruler. This regulation resulted in actual division locations appearing according to specific rules. We also showed that the timing of divisions was set by the so-called adder mechanism. This means bacteria on average divide each time they have grown by a specific volume. The observations on the Min regulation and adder mechanism were remarkable, since these systems were hitherto only known for regulation of division in bacillary shaped bacteria. Taken together, the results in chapter 3 indicate that E. coli cells continuously keep track of absolute length to control size, suggest a wider relevance for the adder principle and provide a new perspective on the function of the Fts and Min systems.

In chapters 4-6 we investigate the origins of cellular heterogeneity. As described in chapter 1, even individual bacteria in an isogenic population in a constant environment can show different behaviour, which ultimately stems from the stochastic nature of chemical reactions that go on inside the cells. Research into this heterogeneity often either focuses on processes where stochasticity in a process can be directly linked to a phenotypic effect or focuses on how signalling can be robust despite noise. How noise transmits through large biochemical networks, and what effects this has on the cellular state, is researched less often. In **chapter 4** we review literature that focuses on this question, with a specific focus on the metabolism, since this is an important large bio-

chemical network in the cell. In **chapter 5** we investigate the role of regulatory networks in heterogeneity. Specifically, we look at the cAMP receptor protein (CRP), which is a master regulator of metabolic enzyme expression. Metabolic enzymes convert carbon molecules taken up by the cell into smaller metabolites (which also generates energy to fuel the cell). Some of these small metabolites also inhibit CRP activity. This negative feedback was previously reported to be the regulatory interaction that is responsible for setting metabolic enzyme expression to the optimal level. Given the hypothesis described in chapter 4 that metabolite levels continuously fluctuate in a single cell, even in steady state, we speculated that the CRP regulatory system must also continuously receive different inputs, to which it might respond. To test this hypothesis we used a previously constructed *E. coli* strain that lacked the metabolic feedback. By subjecting this strain to alternating high and low input signals, we showed that regulatory systems can respond on the fast timescales that are associated with stochastic fluctuations. We then investigated whether the dynamic behaviour of the regulatory interaction was different between a wild type strain (that had feedback regulation) and the strain lacking the feedback. Using cross-correlation analysis of expression-growth dynamics and mathematical modelling of the dynamics, we revealed noteworthy differences between the case with feedback and the case without feedback. This suggested that regulatory interactions indeed respond to stochastic fluctuations that occur within the cell. These observations inspired us to speculate that even regulatory networks in cells in a constant environment continuously interact and adjust, resulting in a perpetually changing cellular state fuelled by random events.

Finally, in **chapter 6** we focus on the role ribosomal concentration fluctuations might have in cellular heterogeneity. It is often suggested that stochastic fluctuations in the ribosomal concentration might contribute to concerted fluctuations in gene expression. Given the pivotal role of protein production in cellular growth, we additionally hypothesized that fluctuations in ribosomal concentration might even transmit to single cell growth rates. We investigated these two hypotheses in this chapter. We studied ribosomal dynamics by using labeled ribosomal proteins and ribosomal RNA reporters. We additionally introduced a constitutively expressed fluorescent reporter, which allowed us to correlate ribosomal concentration with protein production. We measured the expression of these reporters and growth rates in single cells in different media, and also tested conditions where cells were exposed to sub-lethal concentrations of antibiotics that inhibited translation. Cross-correlation analysis of these experiments provided insufficient evidence to clearly support either transmission of ribosomal fluctuations to protein expression nor transmission to growth rates. We end the chapter by reflecting on the question whether a single labeled ribosomal protein is good proxy for the concentration of completely assembled active ribosomes; each of the 58 ribosomal proteins might have their own dynamics. We also suggest a future experiment to further investigate the role of ribosomal RNA in gene expression dynamics.

SAMENVATTING

In dit proefschrift beschrijven we onderzoek naar individuele bacteriecellen. We proberen hiermee meer inzicht te krijgen in het proces van celdeling tijdens ongunstige condities en het fenomeen van heterogeniteit in cellulaire populaties beter te begrijpen.

In **hoofdstuk 1** geven we een overzicht van de onderwerpen in dit proefschrift, alsook context voor de leek. In **hoofdstuk 2** beschrijven we vervolgens hoe we huidige methoden hebben uitgebreid om het gedrag van individuele *Escherichia coli* cellen te kunnen bestuderen. We bespreken een PDMS *microfluidics* opstelling die we gebruiken om microcolonies van bacteriecellen bloot te stellen aan veranderingen in groeimedium en te observeren. Daarnaast bespreken we software-uitbreidingen die onze analyse van individuele cellen faciliteren, en hoe we cross-correlaties kunnen toepassen op onze datastructuur (cross-correlaties zijn relevant in onze studie naar heterogeniteit).

In **hoofdstuk 3** beschrijven we nieuwe inzichten aangaande celdeling in ongunstige condities. In plaats van cellen te laten groeien in gunstige condities, wat meestal gedaan wordt in het lab, hebben we cellen laten groeien in ongunstige condities, namelijk blootstelling aan antibiotica en hoge temperaturen, Deze komen cellen in de echte wereld ook vaak tegen. De cellen reageren op deze ongunstige condities door te stoppen met delen terwijl ze wel blijven groeien. Dit proces wordt ook wel filamentatie genoemd, en leidt tot zeer lange cellen. Wanneer we overschakelden naar gunstige groeicondities, herstelden de cellen zich door weer te gaan delen. Tijdens dit proces verplaatsten de cellen hun potentiële delingsplekken (Fts-ringen) continu. De plaatsing van de Fts-ringen hing af van de lengte van de individuele bacterie en werd bepaald door regulatie door Min-eiwitten. De Min-eiwitten bleken te functioneren als waren zij een dynamische liniaal Dit resulteerde in zeer precieze regulatie van waar daadwerkelijke celdelingen plaatsvonden. We laten in dit hoofdstuk ook zien dat de timing van de delingen plaatsvond volgens het zogenaamde *adder*-principe (letterlijk vertaald: toevoeger-principe). Dit dicteert dat cellen gemiddeld delen wanneer zij met een specifieke volume gegroeid zijn. Deze observaties aangaande het Min systeem en het *adder*-principe zijn opzienbarend, aangezien deze systemen tot nu toe vooral beschouwd waren in de context van cellen met een normale morfologie. De resultaten in dit hoofdstuk laten zien dat *E. coli* cellen continu hun lengte meten om hun grootte te controleren, en duiden op een bredere relevantie van het *adder*-principe en werpen een nieuw licht op de functies van het Fts en Min systeem.

In de hoofdstukken 4-6 onderzoeken we hoe heterogeniteit in populaties ontstaat. Zoals beschreven in hoofdstuk 1, laten zelfs genetisch identieke individuele bacteriën in een constante omgeving verschillend gedrag zien. Uiteindelijk komt dat door het stochastische (willekeurige) karakter van de chemische reacties die plaatsvinden in de bacteriën. Onderzoek naar heterogeniteit focust vaak óf op processen waarin fenotypische effecten direct kunnen worden verbonden aan een bron van stochasticiteit, óf op hoe biochemische beslissingsnetwerken robuust kunnen zijn ondanks ruis die ontstaat door stochastische fluctuaties. Hoe ruis zich voortplant in grote biochemische netwerken, en

wat de effecten zijn op de staat van de cel, wordt minder vaak onderzocht. **Hoofdstuk 4** geeft een literatuuroverzicht over deze laatste vragen. We leggen hierin een focus op het metabolisme, aangezien dit een essentieel biochemisch netwerk in de cel is. In **hoofdstuk 5** onderzochten we de rol van regulatienetwerken in heterogeniteit. Om deze rol verder te begrijpen keken we naar het cAMP recepter-eiwit (CRP), een *master* regulatie-eiwit dat de expressie van metabole enzymen reguleert. Metabole enzymen zetten kool-waterstofmoleculen (zoals suikers) om in kleinere metabolieten. Dit dient ook om de cel van energie te voorzien. Sommige van deze metabolieten remmen CRP regulatie. Deze negatieve feedback is eerder onderzocht, er wordt gedacht dat deze interactie ertoe leidt dat het expressieniveau van metabole enzymen optimaal is. In hoofdstuk 4 stellen we echter dat de concentraties van metabolieten wellicht continu fluctueren door stochasti-sche ruis in de cel. Dat zou betekenen dat het CRP systeem constant verschillende inputs krijgt, zelfs in een constante omgeving, en hier wellicht op reageert. Om deze hypothese te testen gebruikten we een eerder gemaakte genetisch gemanipuleerde *E. coli* stam die deze metabole feedback niet heeft. Eerst gaven we op een artificiële manier alternerend hoge en lage signalen als input aan het CRP regulatiesysteem in deze stam. We alterneer-den dit signaal redelijk snel (elk uur), omdat stochastische fluctuaties waarschijnlijk ook op snelle tijdschalen plaatsvinden, en we wilden testen of bacteriën überhaupt op der-gelijke tijdschalen kunnen reageren. De snelle reactie van de bacteriën op deze signalen liet zien dat ze in principe in staat zijn te reageren op dergelijke snelle signaalwisselin-gen. Daarna onderzochten we het verschil in de dynamiek van de CRP regulatie in een wild type *E. coli* stam (waarin het feedbacksysteem intact is) en de *E. coli* stam zonder feedback. Met behulp van een cross-correlatie analyse van de expressie-groei dynamica en wiskundige modellen van de dynamica, konden we laten zien dat deze verschillen er zijn en hier een interpretatie aan geven. De analyse suggereerde dat het regulatienetwerk inderdaad reageert op stochastische fluctuaties in de cel. Daarnaast waren de resultaten consistent met een interpretatie van de dynamica waarin de regulatie er voor zorgde dat ruis zich minder goed kon voortplanten. In bredere zin zou dit erop kunnen duiden dat regulatienetwerken in cellen in een continu veranderende staat zijn doordat zij continu reageren op stochastische fluctuaties en daaropvolgende effecten.

Ten slotte onderzoeken we de rol van ribosomen in cellulaire heterogeniteit in **hoofd-stuk 6**. Een hypothese die vaak genoemd wordt in de literatuur is dat stochastische fluc-tuaties in de concentratie van ribosomen leiden tot simultane fluctuaties in genexpres-sie. Aangezien de productie van eiwitten een essentiële rol heeft in celgroei, hadden we tevens de hypothese dat op het niveau van één cel fluctuaties in ribosoomconcentraties wellicht ook een effect zouden hebben op de groeisnelheid van de cel. In dit hoofdstuk hebben we deze twee hypotheses onderzocht. We hebben de dynamica van ribosomen onderzocht met behulp van fluorescent gelabelde ribosomale eiwitten en fluorescente reporters voor de expressie van ribosomaal RNA. Daarnaast hebben we een fluorescent eiwit ingebracht dat constant tot expressie werd gebracht. Dit eiwit hebben we gebruikt om het verband te onderzoeken tussen de ribosoomconcentratie en eiwitproductie. We hebben de expressie van deze reporters en de groeisnelheden gemeten in individuele cellen in verschillende groeimedia en hebben daarnaast condities onderzocht waarin de cellen waren blootgesteld aan subletale concentraties translatie inhiberende antibiotica. Ook hier hebben we cross-correlatie analyses uitgevoerd, maar we hebben onvoldoende

bewijs gevonden om aan te tonen dat er transmissie plaatsvindt van ribosomale fluctu-aties naar eiwitexpressie of groeisnelheid. We eindigen dit hoofdstuk 6 met de vraag of een label aan een enkel ribosomaal eiwit een goede weergave is van de concentratie van compleet geassembleerde ribosomen; elk van de 58 ribosomale eiwitten heeft wellicht zijn eigen dynamica. Daarnaast stellen we voor hoe de rol van fluctuaties in ribosomale RNA concentraties verder onderzocht kan worden.

1

INTRODUCTION

1.1. THE POWER OF NUMBERS

Describing the world around us with numbers is not just an interesting hobby for number enthusiasts. Quantitative descriptions of phenomena can give insights that otherwise would have been out of reach. Some examples of such research can be found in the field of biology. Like the Lotka-Volterra model from the early 1900s, which explains fluctuations in predator and pray numbers using differential equations [1, 2]. Or the mathematical reaction-diffusion models created by Turing, which explain how large spatial inhomogeneities can spontaneously arise from molecules that interact by simple rules [3]. Such principles are for example hugely important during embryogenesis, when cells need to decide which part of the body they will develop into. However, perhaps partially because of reluctance by biology researchers [4] and partially because a lack of proper quantitative measuring tools [5, 6], such quantitative approaches were not commonplace in biology. But since the turn of the millennium, quantitative biology has taken a huge flight again. I hope the work in this thesis also offers exciting examples of new insights that are acquired with this new wave of quantitative measurements.

1.2. PRYING IN THE LIVES OF SINGLE CELLS

Before the onset of techniques that allowed high throughput quantitative measurements, bacterial researchers usually did not perform studies on individual specimens of their organism of interest. Experiments were conducted in test tubes on millions of individual cells at once. While this gives great information about the average behaviour of bacteria, which filled text-books with the detailed workings of bacteria, it does not allow one to learn everything about how a bacterium works. New techniques now allow us to better probe the life of single cells. Where a century ago people had to observe tiny bacterial colonies by eye through the microscope, painstakingly draw them, and quantify division times manually [7], we can now do the same using a computer and thoroughly track the lives of thousands of individual cells in a few days work, as described in chapter 2 of this
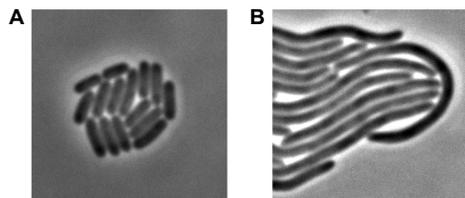
1

Figure 1.1: Bacterial colonies. These pictures were taken with a microscope in the Tans lab, and show a growing microcolony of bacteria in favorable conditions (A) and (part of) a colony of bacteria that are exposed to a sub-lethal dosis of the antibiotic tetracycline (B). This colony in panel B shows a so-called filamentous morphology, because under the influence of certain stresses bacteria can stop dividing but continue elongating.

thesis. Figure 1.2.A shows an example of a snapshot of a microcolony of the rod-shaped *Escherichia coli* bacterium, acquired by one of our microscopes. This computerized approach gives new insights, as illustrated by chapter 3. During single cell experiments in the Tans lab that probed the effect of antibiotics on *E. coli* bacteria we noticed a peculiarity. Bacteria stop dividing in this condition, but continue elongating, leading to so-called filamentous phenotypes with an elongated morphology, see also figure 1.2.B. When antibiotic stress was removed, and *E. coli* with elongated morphologies started dividing again, we noticed that they always divided at very defined relative positions. Before these observations, it was thought that filamentous bacteria behave as if they were a string of multiple bacteria merged together. But our measurement showed that instead of following rules according to that principle, bacteria followed another set of rules. This required an unexpected and not previously observed mobility of bacterial division site structures. It also revealed that one of the bacterial regulation systems for division site placement has a previously unrecognised functionality when cells are in their filamentous state. Aside from these two novel observations, the results implied that bacterial cells carefully regulate their size, even in this elongated morphology. This morphology is not often studied in the lab, but is very relevant in daily life where bacteria might for example survive antibiotic treatments or cold conditions (fridges) by adopting this filamentous form.

1.3. SINGLE CELLS ARE DIFFERENT

Single cell experiments also allow us to probe the differences between individual bacteria. In a pioneering study in 1976, Spudich and Koshland [8] found that genetically identical individual bacteria can still show different food-searching behaviour. In brief, some bacterial individuals favoured swimming straight for long stretches in search of food, whilst others preferred to re-orient their direction more often to create a more winding search pattern for food. This was attributed to 'chance occurrences [in their] internal processes'. It is now clear that in general, the biochemical reactions that finally lead to bacterial decision making do not always have the same outcome. It is thought that the origin of this stochasticity can be traced back to reactions that involve a small number of molecules in a relatively large volume. Since reactants need to find each other by diffusion, this introduces a component of chance. This stochasticity can also be observed in gene expression. Even with precisely the same amount of activation by regulatory molecules, gene expression will fluctuate over time and differ between individuals [9].

This implies that protein concentrations in bacterial cells fluctuate constantly, which in turn must affect all cellular processes. Indeed, it has been suggested that even fluctuations in single enzymes can result in fluctuations in single cell elongation rates (i.e. growth rates) [10]. We review literature about the consequences of stochasticity on the bacterial metabolism, and subsequent effects on single cell growth and population dynamics in chapter 4.

1.4. REGULATION AND STOCHASTICITY

One could ask to what extent these stochastic fluctuations disturb the cellular regulatory networks. In chapter 5 we show that an important metabolic regulatory protein not only responds to the cellular environment, but can also respond to stochastic fluctuations that occur inside the cell itself. This suggests that a stochastic protein concentration fluctuation occurring somewhere in the cell (e.g. in single a metabolic enzyme), might be followed by a fluctuation in the concentration of a specific metabolite, in turn leading to a regulatory response, leading for example to the production of additional proteins, which consequently triggers again different responses, etcetera. Fluctuations might thus have cell-wide consequences, for example on cellular growth rate, mediated by regulatory interactions. This implies that cells might not exist in a well-defined average state, but instead by nature have an ever-changing state. This chapter indicates that asking to what extent stochastic fluctuations disturb regulatory networks might in fact be the wrong question, and instead one should be asking to what extent fluctuations are an integral component of regulatory networks [6].

1.5. SOURCES OF CELLULAR INDIVIDUALITY

In the final chapter of this thesis, chapter 6, we try to further understand cellular individuality. We focus on the ribosome. As a general rule, with a few exceptions, all components of the cell either are proteins, or are produced by reactions that are catalyzed by proteins. Proteins themselves are also produced by complexes (i.e. superstructures) of many proteins, which are called ribosomes. Ribosomes are also an exception to aforementioned general rule, as they also contain RNA. Ribosomes could be major contributors to cellular individuality, as they are often cited as a cellular component that can result in cell-wide protein fluctuations [11–14]. The idea is that when the concentration of ribosomes fluctuates in a cell, the production rates of all proteins that are being produced in that cell also fluctuate simultaneously. These cell-wide fluctuations could potentially have implications on all cellular processes, including behaviour and growth. We investigated this hypothesis, but could neither validate nor disprove it. This might be because the ribosome is such a complex structure, of which the different components that we are able to track experimentally might show different dynamics. This final chapter therefore illustrates that there are still many open questions in our understanding of cellular individuality.

INTRODUCTIE (NL)

DE KRACHT VAN GETALLEN

De wereld om ons heen beschrijven met behulp van getallen is niet slechts een hobby voor getallenfetisjisten. Kwantitatieve beschrijvingen van fenomenen kunnen inzichten geven waar men anders niet toe gekomen was. Ook in de biologie zijn hier voorbeelden van. Zoals het Lotka-Volterra model uit de vroege 20e eeuw [1, 2], dat fluctuaties in hoeveelheden roofdieren en prooien verklaart met behulp van differentiaalvergelijkingen. Of het wiskundige reactie-diffusiemodel van Alan Turing, dat verklaart hoe ruimtelijke heterogeniteit spontaan kan ontstaan uit moleculen die interageren volgens simpele regels [3]. Dit principe is bijvoorbeeld extreem belangrijk tijdens de embryogenese, wanneer cellen moeten bepalen tot welk deel van het lichaam zij zich zullen ontwikkelen. Echter, wellicht deels vanwege een terughoudendheid bij biologen [4] en deels door een gebrek aan de juiste onderzoeksmiddelen [5, 6], zijn dit soort kwantitatieve methoden nooit breed toegepast in de biologie. Sinds het nieuwe millennium heeft kwantitatieve biologie echter een enorme vlucht genomen. Ik hoop dat het werk dat beschreven wordt in deze thesis ook spannende voorbeelden geeft van nieuwe inzichten die verkregen zijn met deze nieuwe golf van kwantitatieve onderzoeken.

HET LEVEN VAN INDIVIDUELE CELLEN BESPIEDEN

Voordat *high troughput* methoden werden uitgevonden, bestudeerden bacteriële onderzoekers doorgaans geen individuele exemplaren van het organisme waar ze in geïnteresseerd waren. Metingen werden gedaan aan reageerbuizen met miljoenen bacteriële cellen. Zulke metingen geven een goed beeld van het gemiddelde gedrag van een bacterie. Hele studieboeken zijn gevuld met gedetailleerde informatie die zo verkregen is. Deze manier van onderzoeken geeft echter geen volledig beeld van hoe een bacterie werkt. Nieuwe technieken stellen ons in staat het leven van de individuele bacterie beter in kaart te brengen. Een eeuw geleden moest men bacteriële kolonies nog met het blote oog door de microscoop observeren en handmatig natekenen om delingstijden (een maat voor de groeisnelheid) van individuele bacteriën te bepalen [7]. Tegenwoordig kunnen we hetzelfde doen met een computer, en het leven van duizenden bacterien in kaart brengen in slechts een paar dagen werk. Figuur 1.2.A laat een voorbeeld zien van een microscoopopname van een microkolonie van de staafvormige bacterie *Escherichia coli* (*E. coli*), verkregen met een van onze microscopen. Zie hoofdstuk 2 van deze thesis voor meer over hoe wij metingen verrichten. Een dergelijke geautomatiseerde aanpak om individuele cellen (*single cells*) te bestuderen geeft nieuwe inzichten, zoals bijvoorbeeld beschreven in hoofdstuk 3. Hierin beschrijven we dat toen we de effecten van antibiotica op individuele cellen wilden onderzoeken, ons iets eigenaardigs opviel. De bacteriën stopten met delen door de antibiotica, maar bleven wel groeien, waardoor

1
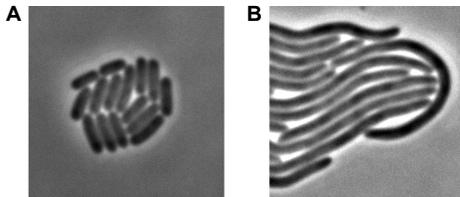
Figuur 1.2: Bacteriekolonies. Deze twee foto's zijn genomen met de microscopen van het Tans lab, en laten een groeiende microkolonie van bacteriën zien in gunstige condities (A), en (een deel van) een bacteriële microkolonie die is blootgesteld aan een niet-dodelijke concentratie van het antibioticum tetracycline (B). De bacteriën in de kolonie in figuur 1.2.B laten een gefilamenteerd uiterlijk zien doordat zij zijn gestopt met delen maar nog wel groeien. Dit is een reactie op stresscondities, waaronder blootstelling aan antibiotica.

zogeheten gefilamenteerde bacterien ontstonden. Deze bacteriën hebben een sterk verlengd uiterlijk, zie ook figuur 1.2.B. Wanneer de antibiotica verwijderd werden, en de *E. coli* met de verlengde morfologieën weer startten met delen, zagen we dat ze altijd deelden op zeer bepaalde relatieve posities. Voorheen dacht men dat gefilamenteerde bacteriën zich gedroegen als een ketting van samengevoegde bacteriën. Maar in plaats van te delen volgens regels zoals je die op basis van deze aanname zou verwachten, volgden de bacteriën een heel andere set regels. Dit vereist een onverwachte en niet eerder geziene mobiliteit van de bacteriële delingsstructuren. Dit onderzoek liet ook zien dat bacteriële regulatiesystemen die de delingsstructuren plaatsen een functionaliteit hebben in de filamenteuze bacteriën. Dit werd eerder niet onderkend. Daarnaast impliceren de resultaten dat bacteriën hun grootte nauwkeurig reguleren, zelfs als ze een verlengd uiterlijk hebben. Deze filamenteuze morfologie wordt niet vaak in het lab bestudeerd, maar is wel relevant in de praktijk, waar bacteriën bijvoorbeeld antibioticakuren of koude condities (zoals in koelkasten) overleven door een gefilamenteerde vorm aan te nemen.

INDIVIDUELE CELLEN ZIJN ANDERS

Single cell experimenten helpen ons ook verschillen tussen individuele bacteriën beter te begrijpen. In 1976 lieten de onderzoekers Spudich en Koshland in een baanbrekende studie zien dat genetisch identieke bacteriën toch ander gedrag laten zien wanneer zij naar voedsel zoeken [8]. Sommige bacteriën zwommen liever lange rechte stukken, terwijl anderen liever vaker van richting veranderden tijdens het zwemmen, waardoor ze een kronkeliger zoekpad aflegden. Dit verschil in gedrag werd toegeschreven aan kansgebeurtenissen in hun interne processen. Het is nu duidelijk dat in het algemeen geldt dat biochemische processen die leiden tot het maken van bacteriële beslissingen niet altijd dezelfde uitkomst hebben. Er wordt gedacht dat de bron van deze stochasticiteit ligt in reacties tussen moleculen die slechts in kleine aantallen aanwezig zijn, maar plaatsvinden in relatief grote volumes. Omdat de reactanten elkaar moeten vinden door diffusie (willekeurige bewegingen), introduceert dat een kanscomponent. Deze stochasticiteit beïnvloedt ook de expressie van genen. Zelfs wanneer een gen op een constant, gelijk niveau wordt geactiveerd door regulatiemoleculen, zal de expressie van dat gen fluctueren over de tijd en verschillend zijn tussen individuen [9]. Dit impliceert dat ei-

witconcentraties in cellen constant fluctueren, wat gevolgen heeft voor alle cellulaire processen. Inderdaad is aangetoond dat zelfs fluctuaties in de concentraties van één enzym kunnen leiden tot fluctuaties in de groeisnelheid van individuele cellen [10]. In hoofdstuk 4 diepen we dit onderwerp verder uit met een literatuurstudie naar de consequenties van stochasticiteit voor het bacteriele metabolisme, en daaruit volgende effecten op de groei van individuele cellen en populatiedynamica.

REGULATIE EN STOCHASTICITEIT

Een open vraag is in hoe verre stochastische fluctuaties cellulaire regulatienetwerken verstoren. In hoofdstuk 5 laten we zien dat een belangrijke metabool regulatie-eiwit niet alleen reageert op signalen uit cellulaire omgeving, maar ook reageert op stochastische fluctuaties die vanuit de cel zelf komen. Dit doet vermoeden dat een stochastische fluctuatie in een eiwitconcentratie die ergens in de cel ontstaat (bijvoorbeeld in de concentratie van een metabool enzym) gevolgd kan worden door een fluctuatie van een specifieke metaboliet, wat op zijn beurt weer leidt tot een regulatoire respons, wat leidt tot de productie van extra eiwitten, wat vervolgens weer leidt tot andere cellulaire responsen, et cetera. In andere woorden: fluctuaties kunnen via regulatoire interacties wellicht consequenties hebben voor alle processen in de cel. Dit zou ook een effect kunnen hebben op bijvoorbeeld de groeisnelheid van de cel. Op zijn beurt doet dat vermoeden dat cellen niet in een constante staat (chemische samenstelling) verkeren die lijkt op de gemiddelde cel, maar dat zij in plaats daarvan een constant veranderende staat hebben. Naar aanleiding van dit hoofdstuk kan worden getwijfeld aan de relevantie van de vraag in hoe verre stochastische fluctuaties regulatienetwerken verstoren. Wellicht is een relevantere vraag in hoe verre stochastische fluctuaties een integraal onderdeel zijn van die regulatie [6].

DE HERKOMST VAN CELLULAIRE INDIVIDUALITEIT

In het laatste hoofdstuk van deze thesis, hoofdstuk 6, proberen we cellulaire individualiteit beter te begrijpen. Onze focus ligt hierbij op het ribosoom. In het algemeen, op een paar uitzonderingen na, zijn alle onderdelen van de cel eiwitten, of geproduceerd door reacties die gekatalyseerd worden door eiwitten. Eiwitten zelf worden ook gemaakt door complexen (samengestelde structuren) van vele eiwitten. Deze worden ribosomen genoemd. Ribosomen zijn ook de uitzondering op eerder genoemde regel, aangezien zij naast eiwitten ook bestaan uit RNA. Ribosomen worden vaak genoemd als bron van cellulaire heterogeniteit [11–14]. Het idee hierachter is dat als de concentratie ribosomen fluctueert, de productiesnelheden van alle eiwitten die worden geproduceerd in de cel mee fluctueren. Deze celbrede fluctuaties kunnen in potentie consequenties hebben voor alle cellulaire processen, inclusief groei en gedrag. We hebben deze hypothese onderzocht, maar konden deze noch valideren noch weerleggen. Dit zou kunnen komen doordat de ribosomen zo een ingewikkeld complex van eiwitten en RNA zijn, dat elk onderdeel ervan dat wij kunnen labelen zijn eigen dynamiek kent en tot andere resultaten leidt. Dit laatste hoofdstuk laat dus zien dat het onderwerp cellulaire individualiteit nog vele open vragen kent.

2

METHODS

2.1. INTRODUCTION

A substantial part of the methods used for experiments performed in this thesis were already discussed in detail, either in publications from other labs [15, 16] or in previous work from the Tans lab at AMOLF [17–19]. An additional independent method section can be found in chapter 3 of this thesis as that chapter has been published in a scientific journal.

In this chapter, I will focus on how I further developed these existing experimental protocols and analyses. Importantly, we have recently employed an adapted "mother machine", a microfluidic device that allows growing and observing bacterial microcolonies for very long time as superfluous cells are washed away, and also allows for quick exchange of growth media.

2.2. SINGLE CELL EXPERIMENTS

Observing growing single cells under the microscope provides a novel perspective compared to measuring bacterial behavior in bulk. It can quantify single cell deviations from the mean behavior (see e.g. [9, 10] and chapters 5, 6), and it provides insights in single cell morphologic changes over time (of which the filamentation and division processes are good examples, see chapter 3). One straightforward way to visualize growing microcolonies of cells under the microscope is by using gel pads [9, 10, 15–17, 20] (Figure 2.1.A). Pads can be produced by adding agarose to the desired growth medium, solidifying the medium into a gel. Since it was found cells can also consume agarose, we employ polyacrylamide gel pads in our experiments [17]. The gel pads are soaked in the desired medium, and then transfered to a glass slide. An airtight chamber is created by using a second glass slide with a hole in the middle, a cover slip, silicon grease and a metal scaffold (Figure 2.1.A); for a detailed protocol see [18] and [16].

The disadvantage of growing cells on gel pads is that only a limited number of generations of cells can be recorded. At some point either (a) cells grow to such high densities

that they form multilayers and single cells cannot be distinguished or (b) nutrients provided by the pad run out and cell growth stops.

2.2.1. Microfluidic devices

This advantage can be overcome by using a microfluidic device. In a microfluidic device, fresh medium is typically pumped through the sample, such that bacteria have continuous access to fresh medium. Additionally, most designs allow for superfluous cells to be washed out of the sample, such that microcolonies can be observed for very long times (hundreds of generations, see e.g. [21]). Many microfluidic devices have been developed and employed recently [21–31].

Some experiments described in this thesis (see e.g. chapter 3) were performed with a microfluidic device as depicted in Figure 2.1.B. This device has a flow channel on its bottom. Bacteria are pipetted on a microscope cover slide, and covered by a very thin membrane. The microfluidic device is placed on top, such that fresh medium flowing through the channel can reach the bacteria under the thin membrane. The advantage of this device is that all bacteria have equal and close access to the fresh medium. Disadvantages of this device are that continuous growth eventually leads to multilayered colonies that cannot be analyzed (i.e. superfluous cells are not removed). Additionally, in practice, this device is prone to leakage. A more detailed description of this device can be found in [32].

2.2.2. The microfluidic device used in this work

Most work with microfluidics in this thesis was however done with a more recent microfluidic design. This device was developed by Daan J. Kiviet (unpublished), it is similar to the "mother machine" described in ref. [33] but it has wider microcolony wells. As shown in Figure 2.1.C, this design contains 4 replicate flow channels. Specifically, each flow channel contains a 200 µm wide main flow channel, splitting into two 100 µm flow channels. Perpendicular on these are chambers with the following widths: 1x 80 µm, 1x 60 µm, 2x 40 µm, 3x 20 µm, 3x 10 µm and 3x 5 µm; and with depths of 60 µm, 30 µm, 50 µm and 40 µm (Daan Kiviet, personal communications). These blocks forms chambers that are .75 µm high, also referred to as "wells", in which bacterial microcolonies can grow. The main channel is 23.5 µm high (heights reported here refer to the original wafer).

Epoxy mold fabrication

The PDMS devices were made by casting them into an epoxy mold, which was a gift from Daan J. Kiviet and the Ackermann lab. The production of the epoxy mold was performed following an online protocol [34]. Briefly, this protocol involved creating a copy of the original mold using R123 epoxy resin and R614 hardener. The epoxy mixture was prepared as described by the manufacturer, and bubbles were removed by vacuum pumping (ultrasonification is an alternative technique). The original PDMS cast from the wafer mold was placed in a container, and the epoxy mixture was poured on top (approximately 5 mm thick). The cast was left overnight at room temperature, the PDMS was then removed (using scalpel and tweezers), after which the mould was baked at 70

C for one day, and it was then left for another day at room temperature to harden. It was then cleaned with plasma and silanized using a trichloromethylsilane-saturated atmosphere for 5 min.

PDMS DEVICE FABRICATION

To produce the polydimethylsiloxane (PDMS) device, polymer and curing agent (Sylgard 184 elastomer, Dow Corning Corp.) were mixed in a 50 mL Falcon tube using a polystyrene dinner fork (product nr. 888223, Bright Packaging) and a vortex mixer, using 1 mL of curing agent for each 7.7 g of polymer (i.e. not the recommended 1:10 ratio). Then, this mixture was cast into an epoxy mold (made by Daan J. Kiviet) that is a positive copy of the original wafer mold (sometimes dust was removed from the mold by pressured Nitrogen gas). Air bubbles were removed from the mixture either by putting the mold and casting in a dessicator for 30 minutes, or by leaving the mixture for several hours before casting. The mold and casting were then put in a 80 C oven for 1-12 hours. Subsequently, the casting was removed from the device, and holes were punched for the liquid in- and outlets. The device was cut into a smaller size using a scalpel to remove rough, raised or uneven edges. Then the PDMS device was covalently bound to a clean glass cover slip by treating the PDMS and glass surface with a portable corona device [35] (5-10 sweeps of approx. 5 seconds for each surface from approx. 5-10 mm distance). The device was gently tapped using a gloved finger to improve contact between the PDMS and glass surfaces. (During this procedure the PDMS was only handled with clean metal pincers, a scalpel, or gloved hands.) Consecutively, the device was baked for another 1-12 hrs. We noticed bonding continued to improve during storage at RT for 1-2 weeks after completion of this protocol .

INOCULATION OF BACTERIA INTO THE DEVICE

To inoculate bacteria into the device, 2 mL culture of *Escherichia coli* is grown to high OD (>1) in a 10 mL Falcon culture tube on a rotator, either at 37 C or 30 C (O/N). The concentration of bacteria is further increased by spinning down 1 mL of the sample in an eppendorf tube at 2300-16100 RCF, and removing supernatant such that the concentration is increased by a factor of approximately 30. The sample is then resuspended. First, 1 μL of sterile 0.01% Tween (H_2O) solution is slowly pipetted (the plunger is pressed down in approximately 5-10 seconds) into the PDMS device trough one of the holes. Similarly, 1 μL of the concentrated culture is then introduced in the device. Alternatively, a syringe attached to a piece of polyethylene tubing (Fine Bore Polyethyline Tubing, 0.55 mm inner diameter, 0.96 mm outer diameter, Smiths Medical), attached to a small metal tube (outer diameter approx. 0.65 mm), which is then inserted into the device at one of the holes can be used to very slowly introduce first Tween solution and then the condensed bacterial culture into the device.

SETTING UP THE EXPERIMENT

Once the device is inoculated with bacteria, it can be placed under the microscope (usually placed in the same metal scaffold as used for the gel pad experiments, Figure 2.1.A). Two small metal tubes (outer diameter approx. 0.65 mm) are then inserted in the holes of the PDMS device, and used as connectors to connect polyethylene tubing (Fine Bore

2

Polyethyline Tubing, 0.55 mm inner diameter, 0.96 mm outer diameter, Smiths medical International Ltd.). One of the tubes is in turn connected to a syringe (either 10 mL or 50 mL) placed in a microfluidic pumps (ProSense NE-1000 and NE-300) containing the desired growth medium. The other end of the other tube is placed in a waste collection erlenmeyer flask. When desired, instead of one pump with a medium syringe, two pumps each containing a syringe with different medium and an automated valve (Modular valve positioner, RS232, Hamilton) can be used to quickly switch between two media. Typical flow rates are between 0.5 and 1.0 mL/hr.

2.3. DATA COLLECTION

As also described in chapter 3 and previously [10, 32], cells were imaged with an inverted microscope (Nikon, TE2000), equipped with 100X oil objective (Nikon, Plan Fluor NA 1.3), cooled CMOS camera (Hamamatsu, Orca Flash4.0), xenon lamp with liquid light guide (Sutter, Lambda LS), GFP, mCherry, CFP and YFP filter set (Chroma, 41017, 49008, 49001 and 49003), computer controlled shutters (Sutter, Lambda 10-3 with SmartShutter), automated stage (Marzhauzer, SCAN IM 120 x 100) and an incubation chamber (Solent) allowing precise 37 C temperature control. An additional 1.5X lens was used, resulting in images with pixel size of 0.0438 μm. The microscope was controlled by Meta-Morph software (Molecular Devices).

2.4. UPDATES TO THE ANALYSIS

To extract quantitative data from the time lapse movies, further analyses are performed. Briefly: Cells are first segmented to identify areas that constitute cells in the microscopy images. Then, after individual cells have been identified, the lineages of the cells are tracked over the frames of the time lapse. In subsequent analyses cellular parameters are characterized. Examples of important parameters are length, growth rate and the concentration of fluorescent reporter.

Many algorithms have been developed that can segment microcolony data, track individual cells from one frame to the next, and quantify cellular parameters to a greater or lesser extent. Examples include Schnitzcells [10, 16], Supersegger [36], Oufti [37], MoMA [38], Sachs et al. [39], Nobs et al. [40], MicrobeTracker [29, 41], CBA [42] and MAMLE [43]. The employed segmentation procedures range from simple image thresholding to more advanced combinations of image transformations. A recent development is the usage of machine learning to effectively segment microcolony data [44].

Throughout this work, we use a custom set of scripts based on the schnitzcells framework, which was developed in the Elowitz lab [16]. These custom scripts are a mostly written by Daan J. Kiviet [17], Philippe Nghe and Noreen Walker [18]. Updates and novel functions are also introduced in this work. These improvements are described in the sections below. Additionally, we introduced a script called

```
Schnitzcells_masterscript.m,
```

from which the whole analysis can be run. This script interacts with a custom made Microsoft Excel configuration file that holds important analysis parameters (directories, segmentation parameters, fluorescence parameters, etc). This script also provides
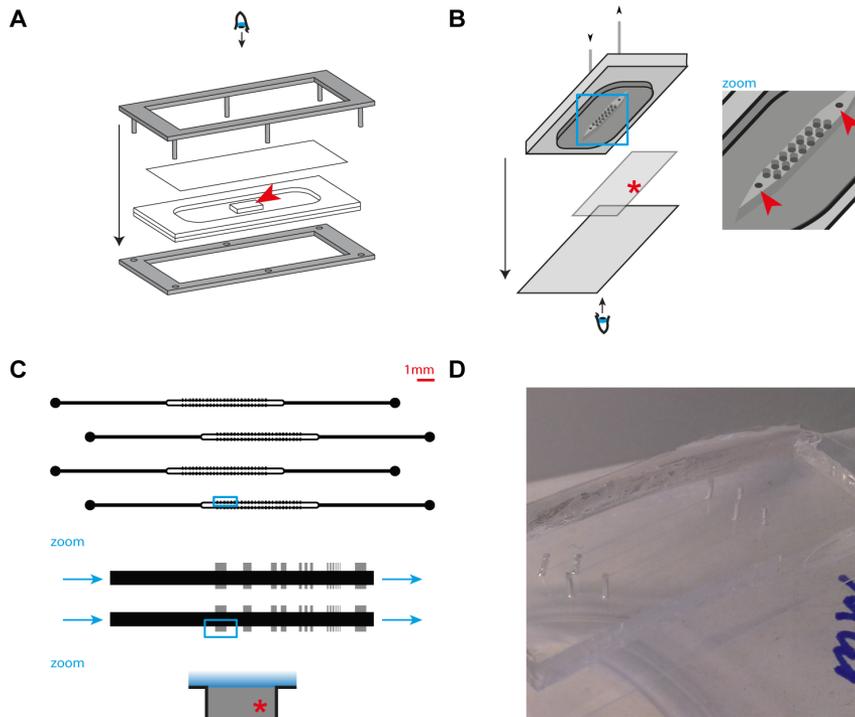
Figure 2.1: (A) Schematic view of a sample that employs a gel pad to grow cells. Two metal parts (grey) hold the glass slides together using screws. Bacteria are pipetted onto a gel pad (red arrow) which is placed in a chamber build from a glass microscope slide, a glass slide with a hole, and a cover slip. The image is placed upside down in the microscope, the eye indicate the direction from which the microscope will image the sample. Image by Noreen Walker. (B) Cartoon (not on scale) of a microfluidic device used in some experiments in this thesis. The asterisk indicates bacteria are grown in between a cover slip (shown at the bottom) and a thin polyacrylamide membrane (shown in the middle). On top a PDMS slab is placed, which has a flow chamber on its bottom (blue square and zoom), through which medium is flown using two holes (red arrows) that are connected to tubing. This fresh medium can access the bacteria by diffusion through the thin membrane. The 30 by 3 mm chamber contains pillars placed on a 0.6 mm grid (here only a few are shown for illustration purposes) to keep the thin polyacrylamide membrane in place. (C) A more advanced microfluidic device employed in this work was designed by Daan J. Kiviet. The mask shown here at the top is used to produce PDMS slabs with four identical growth medium channels, which each also fork into two main channels. These channels (which are 23.5 μm high) have protrusions on the side that are less high (0.75 μm), which we refer to as wells. The PDMS slab is covalently bound to a glass cover slip, and after the inoculation procedure (see main text) single layered microcolonies can grow in the wells. Medium is flown through the main channels such that cells receive fresh medium. When the colonies grows and divides, superfluous cells are washed away at the side of the well that faces the main channel. (D) A picture of a microfluidic PDMS device as described in panel C.
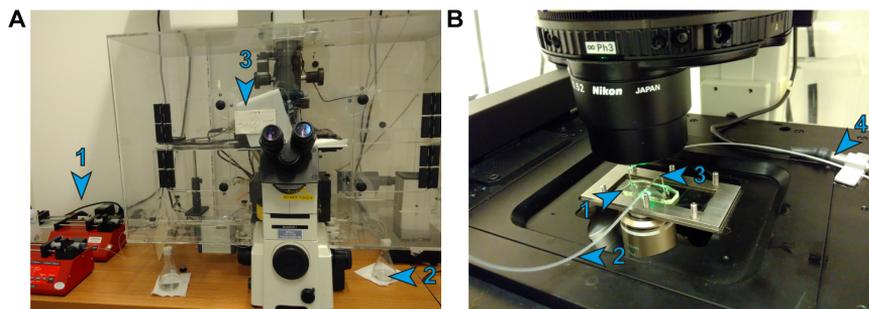
2

Figure 2.2: (A) Microscopy setup. One syringe pump (see label 1) pumps culture medium into the sample on the microscope stage in the temperature chamber and then to a waste collection Erlenmeyer flask (2), the other pump pumps medium directly to a waste Erlenmeyer. The two pumps in combination with an automated valve (3) that controls which medium goes to the sample and which directly goes to the waste flask allow for quick switching between the two media. (B) The microfluidic device under the microscope. Medium is supplied to the sample (1) through a polyethylene tube (2) and a metal connecting tube (3). Waste medium is disposed through another polyethylene tube (4).

a graphical user interface to increase the efficiency of the analysis (Figure 2.3). More detailed information about how this "master script" can be run can be found in the comments of the script. The scripts are available in a git repository at

```
https://bitbucket.org/microscopeguerrillas/schnitzcells_tans.git,
```

which holds the main files, and a git repository at

```
https://bitbucket.org/microscopeguerrillas/
schnitzcells_tans_extensions.git
```

that holds additional files required to run the full analysis.

2.4.1. SEGMENTATION

Throughout this work, we use a segmentation algorithm developed by Philippe Nghe [18]. Additionally, all frames are manually checked for segmentation mistakes. However, the Nghe script was developed for data from gel pad time lapse experiments. Thus, some minor modifications were introduced to handle data from the microfluidic device were cells also disappear from the experimental observations. Importantly, cells are flushed away once they reach the main channel, and thus cells that touch the edge of the image (see Figure 2.4 for an example of the segmentation output) need to be ignored from the analysis. The original algorithm could not deal with cells touching the edge of the image. As a simple fix, we introduced a gradient at the edge of the image (Figure 2.4.A) that goes from transparent to white, see the script
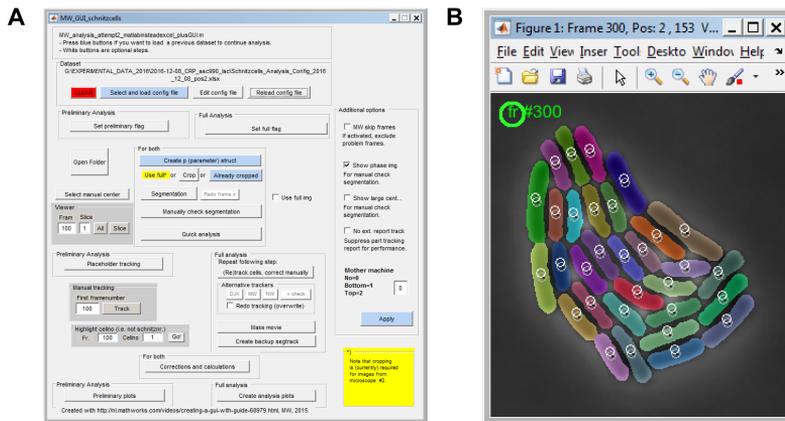
```
MW_preprocessimagefadeedge,
```

Figure 2.3: (A) Screenshot of the GUI introduced to increase the efficiency of the analysis. (B) Screenshot of the Matlab figure with which the user can interact to manually correct the analysis. The segmentation is shown by the colors, and the circles show (corrected and uncorrected) centers of cells from the previous frame.

such that cells touching that edge are considered to have their cell edge at the edge of the image with some margin by the algorithm (Figure 2.4.C-D). This procedure allows processing steps of the algorithm to function as if the data came from a gel pad experiment. Consequently, cells to which this was applied are removed from the analysis.

More adjustments were made during the tracking of cells, see next section.

2.4.2. TRACKING

Calculating the lineage of cells using the segmented images is an essential part of the desired analysis. Two tracking functions already existed, and we have introduced a third simple tracker matlab script to these. This was done because both the other trackers relied to methods that represented the location of cells as a point in one way or another (Figure 2.5.A-B). In one method, cells from both frames to be connected are represented by three points along their skeleton. In the second method, the centers of cells from one frame are compared with the cell areas from the other frame [18]. In specific cases this leads to issues. For example, if cells have atypical shapes (e.g. due to filamentation), the centers characterize their placement less well. Therefor the introduced method, see

```
MW_tracker.m, MW_linkframes,
```

tracks cells by looking at overlap between cell areas. The cell from frame n+1 is considered to be lineage-connected to that cell from the previous frame n that has most overlap with it (Figure 2.5.C).

Additionally, during image acquisition the sample can shift tiny amounts (usually less than a micrometer), which is corrected by the Nghe script by aligning the centers of the microcolonies for consecutive frames. This procedure leads to issues when cells

2
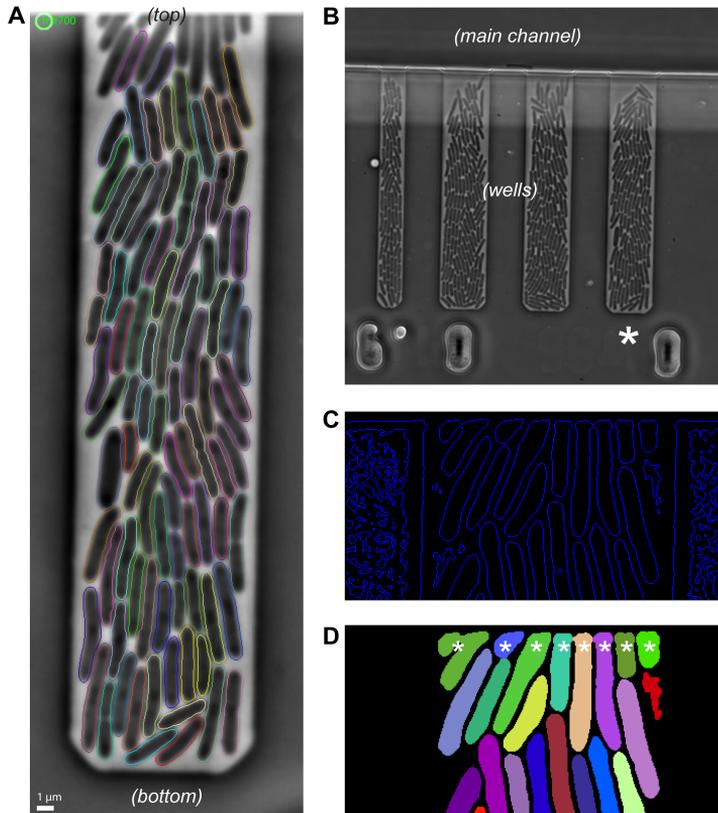
Figure 2.4: (A) The result of the segmentation algorithm, outlines of determined cell areas are shown in color. (A frame the nucleoid labeled cells time lapse data from chapter 3 is shown.) Images are cropped to increase computational efficiency, and also to remove the main channel structure. Note that cells are generally flushed out when they reach the main channel. Growth results in cells moving towards the well exit at the main channel, or towards the edge of the picture labeled top here. Cells that are partially outside the cropped image are ignored in the analysis. (B) The image in panel A is cropped, this panel shows the full image taken by the microscope. The asterisk indicates the well which is also displayed in other panels. (C-D) Image at the top of the cropped image are artificially considered to have a cell edge that is aligned with the image edge. These cells are later removed from the analysis (here indicated with asterisks). Shown are processing steps early (C) and late (D) in the algorithm. (Note that the artifact shown in red in panel D will also be removed later in the analysis, either automatically or manually.)

disappear from the analysis, as is the case with the microfluidic device. (When cells disappear, the centroids do not represent the same set of cells any more, and the procedure can misalign the microcolonies.) Thus, an additional adjustment was introduced in the tracking script, which aims to shift the joined areas that are recognized as cells from the two frames on top of each other, and maximizes the overlap between these two areas (Figure 2.5.D). Specifically, a binary representation of the image (cells detected or no cells detected for that pixel) is integrated along one axis, such that both the n-th frame and the n+1 frame are represented by a line along the other axis. Then, these lines are shifted until the difference between the lines is minimized along that axis. This is done both for the x and y directions (i.e. both axes).

2.4.3. SKELETON LENGTH AND STRAIGHTENING OF SKELETON

Previously [17], the length of the cells during the analysis was determined using a third order polynomial fit through the cell area. In some cases, e.g. in involving filamentation of cells, a 7-th order polynomial was used. An alternative way of determining length is using the skeleton of the cell area. The cellular skeletons are determined in the scripts

`NDL_addToSchnitzes_skeletonLengthMW`, `NDL_lengthforfillamentedcellsMW`,

which were developed by Nick de Lange. The skeleton is determined using Matlab's

`bwmorph`

function, and after removal of branches, it is extrapolated to the edge of the cell's areas using 0.95 µm long windows at both ends of the skeleton (Figure 2.6.A-B). Lengths of the skeleton and the extrapolated parts are determined using Matlab's

`bwdistgeodesic`

function (using the quasi-euclidean method). Additionally, an algorithm was developed to produce straightened representations of the cells using the cellular skeletons. This is useful when for example the localization or intensity of the fluorescence signal along the cellular axis needs to be quantified (this is for example done in chapter 3 for fluorescently labeled division rings and nucleoids). This is done in the script

`MW_straightenbacteria`

by placing lines of equal length tangential to the skeleton, and subsequently placing the image pixels corresponding to each line into the columns of a rectangular matrix (Figure 2.6).

2.4.4. CORRELATION FUNCTIONS, SCATTER PLOTS, WEIGHING AND CONTROLS

THE CROSS-CORRELATION FUNCTION

An important tool to gain insights into the dynamics and interactions between different cellular parameters (such as expression of different proteins and growth rate) is the correlation function. As this function is also central in this thesis, We'll discuss some basic

2
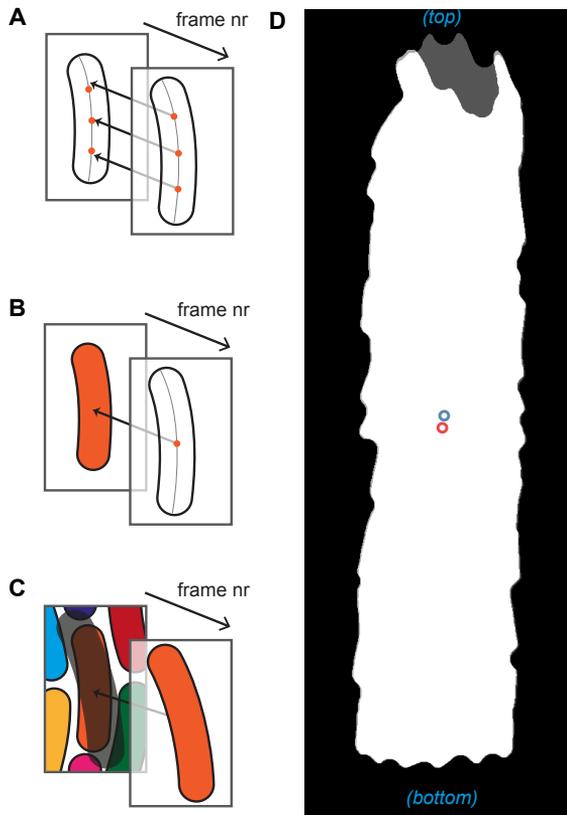
Figure 2.5: (A) Default Schnitzcells tracking algorithm, which characterizes cells by three points along its skeleton. Cells are connected between frames by minimizing the distance between these characteristic points. See [18] for more information. (B) The algorithm introduced by Noreen Walker, which links cells between frames by minimizing the distances of centroids in the n+1 frame and areas of cells in the n-th frame. See [18] for more information. (C) Algorithm introduced in this work, which links every cell in the n+1 frame to that cell in the n-th frame that has most overlap with the projection of the n+1 frame cell onto the n-th frame. In this example, the area of the orange cell in the n+1 frame overlaps with the blue, dark blue, orange, green and purple cells in the n-th frame, but shows most overlap with the orange cell. Thus, the two orange cells in these two consecutive frames are connected. (D) An example of the overlapping colony areas from two consecutive frames from a microfluidic device dataset. Overlap between the colonies is indicated in white, non-overlap in gray. Alignment using centroids would fail in this case, since the microcolony centers (shown as circles) do not represent the same subset of cells. For this reason the overlap of these two areas is maximized to align subsequent frames in data from microfluidic devices. (The microcolony shown is the same as shown in Figure 2.4.) Images (A) and (B) are made by Noreen Walker.
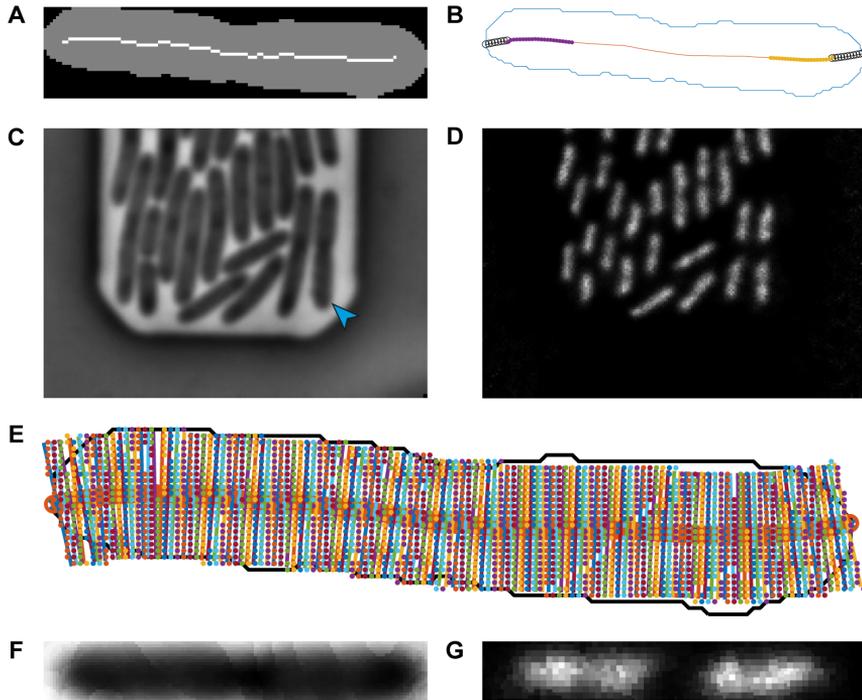
Figure 2.6: (A) Skeleton of an example cell area (gray) calculated by Matlab's `bwmorph` function, after removal of side branches. (B) The skeleton is extrapolated towards cell poles (black circles) using the ends of the skeleton (purple and yellow dots). The total cell length is calculated by summation of the lengths of the extrapolated and original skeleton parts. (C-G) Example of cellular straightening. Based on segmentation (see main text and Figure 2.4) both for the phase image (C) and the fluorescence image (D) cells can be straightened to produce a fluorescence intensity profile along the cellular axis. This is done by using a series of lines that are placed tangential to the skeleton (E), as shown here for the cell indicated by the blue arrow in panel (C). The pixels closest to this line (represented by colored dots here) are used to generate the straightened bacteria, both for the phase image (F) and the fluorescence image (G). The images shown here are part of the dataset with nucleoid labeled cells described in chapter 3.

definitions regarding this quantity here. Simply put, the aim here is to gain insight in to what extent the value of one signal expected to affect the value of another signal, considering there might be a delay in the effect. The cross-correlation function $R(\tau)$ quantifies to what extent the deviation from the mean value in one parameter $f(t)$ at time t is correlated with the deviation from the mean value in another parameter $g(t+\tau)$, at a delay τ later (or earlier). Mathematically, this is expressed as [45, see lemma "Cross-Correlation"]

$$S_{f,g}(\tau) = f \star g = \int_{\tau=\infty}^{\infty} \bar{f}(t)\bar{g}(t+\tau)\delta\tau \qquad (2.1)$$

or for a discrete signal the cross-correlation can be defined as [15]:

$$S_{f,g}(\tau) = \frac{1}{N-|\tau|} \sum_{n=0}^{N-|\tau|-1} \tilde{f}(n)\tilde{g}(n+\tau), \qquad (2.2)$$

where f and g are either continuous and mean-subtracted (indicated by the bar) or discrete and mean-subtracted (indicated by the tilde). N is the number of time points in the data series. When this is normalized by

$$\sqrt{S_{f,f}(0)S_{g,g}(0)}$$

the result is often also referred to as the cross-correlation [46]:

$$\begin{aligned} R_{f,g}(\tau) &= \frac{1}{\sqrt{S_{f,f}(0)S_{g,g}(0)}} \frac{1}{N-|\tau|} \sum_{n=0}^{N-|\tau|-1} \tilde{f}(n)\tilde{g}(n+\tau) \\ &= \frac{S_{f,g}(\tau)}{\sqrt{S_{f,f}(0)S_{g,g}(0)}} = \frac{S_{f,g}(\tau)}{\sqrt{\sigma_f^2 \cdot \sigma_f^2}} \end{aligned} \qquad (2.3)$$

$R_{f,g}(\tau)$ defined in equation 2.3 is also the function that we call the cross-correlation and use throughout chapters 5 and 6 to quantify relationships between different biological quantities. Note that $S_{f,f}(0)$ also equals the variance in f, σ_f^2. When $\tau = 0$, the cross-correlation in Equation 2.3 simply becomes the correlation coefficient $\rho_{f,g}$ (also known as the Pearson's correlation) between two the two parameters f and g:

$$\rho_{f,g} = \frac{1}{\sigma_f \sigma_g \cdot N} \sum_{n=0}^{N-1} \tilde{f}(n)\tilde{g}(n) = \frac{\sigma_{f,g}^2}{\sigma_f \sigma_g} = \frac{cov(f,g)}{\sqrt{var(f)var(g)}}, \qquad (2.4)$$

where σ_f^2 again indicates the variance of f, which can be equivalently written as $var(f)$. (σ_f simply indicates the standard deviation, the square root of the variance.) The covariance is indicated by $\sigma_{f,g}$ or $cov(f,g)$. These quantities are tightly related to the least squares fitting of b in $g = a + bf$, which is given by [45, see lemma "Correlation Coefficient"]:

$$b = \frac{cov(f,g)}{var(f)}; \qquad (2.5)$$

this also shows that $\rho_{f,g} = \sqrt{b \cdot b'}$ (defining $f = a' + b'g$ additionally to $g = a + bf$).

   Note that there is some ambiguity between the terms cross-covariance and cross-correlations when talking about time series data. While Munsky et al. [46] also define the cross-correlation $R(f,g)$ as we do in equation 2.3, Dunlop et al. [15] call the quantity $R(f,g)$ defined in equation 2.3 the cross-covariance, and call $S(f,g)$ in equations 2.1 and 2.2 the cross-correlation (as also defined in [45, see lemma "Cross-Correlation"]). To be consistent with the definitions of the Pearson's correlation coefficient and the covariance, we follow the definition as given by Munsky et al. [46], as also employed by Kiviet et al. [10, 17].

DELAYED SCATTER PLOTS

Cross-correlations are particularly good at detecting linear relationships. As an addition to the analysis using cross-correlations, it can therefor be instructive to analyze time traces of measured cellular parameters by using scatter plots (Figure 2.7), which might show non-linear relationships between the parameters of interest. However, cellular parameters are shown to sometimes correlate with delays; it has for example been shown that the production rates of some enzymes have an effect on growth rate at some later point in time [10]. Thus, simply plotting $f(n)$ and $g(n)$ (which could for example be enzyme production and growth rate, respectively) against each other for all n values might not give the maximum insight in the relationships between the two quantities x and y. Therefor, in the scripts

```
MW_delayedScatter, MW_getdelayedscatter
```

$f(n)$ is plotted against $g(n+\tau)$ in multiple plots, each plot corresponding to one of $(N-1)$ values of τ. This procedure is illustrated in Figure 2.8, which shows that the delayed relationship between two parameters becomes clearer when plotting a scatter plot for that specific delay. Note that we can also determine the correlation coefficient $R(\tau)$ for each of the scatter plots corresponding to a particular τ. Thus, the scatter plots also provide a way to generate a cross-correlation function.

MUTUAL INFORMATION

Though not applied in this work, it might be interesting to provide an outlook on how one could quantify the extent to which parameters are related (in non-linear ways) beyond the visual inspection of scatter plots, or the determination of the correlation coefficient $R(\tau)$. One way to investigate this is to look at the independence of two parameters. If the biological quantity x does not affect another quantity y, the probability to find a certain value of y should not depend on the value of x, and vice versa. In other words, when the two parameters are independent of each other $p(x,y) = p(x)p(y)$. To what extent values of x are constrained by values of y (either by a relationship between the two, or because of an indirect link) and vice versa, is quantified by the mutual information [47–49]. For two random variables X and Y, this defines a distance measure between the product of the two marginal probability densities $p(x)$ and $p(y)$ and the joint probability $p(x,y)$.

$$I(X;Y) = \sum_y \sum_x p(x,y) \log_2 \left(\frac{p(x,y)}{p(x)p(y)} \right), \tag{2.6}$$

2
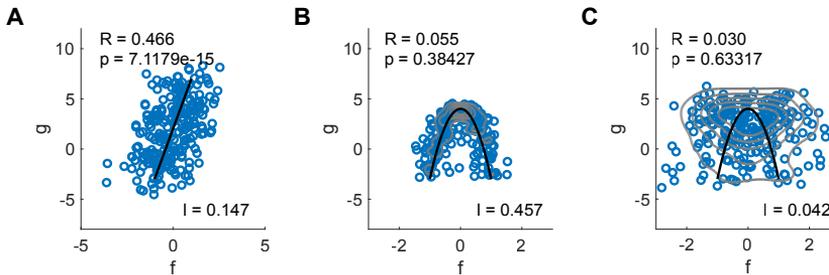
Figure 2.7: (A) Example of a linear relationship, and a cloud of points (each point might represent an experimental measurement at time n) generated by randomly drawing from that relation (shown in black) and adding noise (normal distribution with $\sigma = 1$). The Pearson correlation coefficient R is able to detect the relationship despite the noise. (B-C) Example of a non-linear relationship ($g = -7 * (f).^2 + 4$, shown in black), and a cloud of points generated by randomly drawing from that relation and adding noise (normal distribution with respectively $\sigma = .25$ and $\sigma = 1$ in panels B and C). Since the relationship is non-linear the Pearson correlation coefficient R is not able to detect the relationship. The gray lines are isolines that represent Kernel Density Estimation of the distribution of points. In all panels, I gives the mutual information $I(F; G)$ (based on a Kernel Density Estimation of the probability distribution) of which the non-zero value for panel B shows it is able to detect a dependence between f and g for non-linear relationships (see text). Panel C illustrates that non-linear relationships also quickly get harder to detect, even by eye.
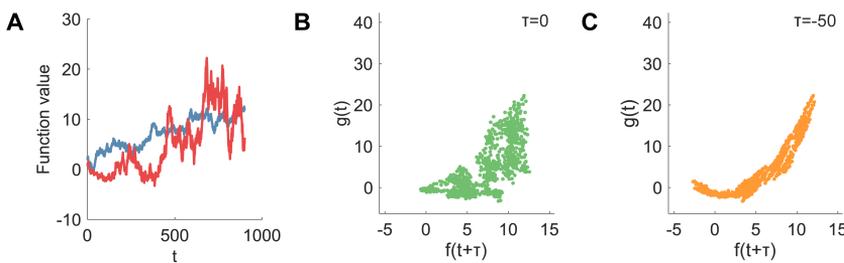
Figure 2.8: (A) Two example noisy signals, $f(t)$ and $g(t)$. $f(t)$ shown in blue is a random walk process, while $g(t)$ shows a random walk plus a contribution by $f(t-50)$, i.e. $g(t) = .8 \cdot w(t) + .2 \cdot f(t-50)^2$, where $w(t)$ represents a random walk like $f(t)$. (B-C) As expected, this imaginary scenario shows that plotting $f(t)$ against $g(t)$ reveals the correlation between f and g poorer compared to plotting $f(t)$ against $g(t-50)$.

or equivalently in continuous form:

$$I(X;Y) = \int_y \int_x p(x,y) \log_2 \left(\frac{p(x,y)}{p(x)p(y)} \right) dx\,dy. \tag{2.7}$$

(Note that the probability densities are always normalized to one.) It has also been shown that the mutual information can be found by

$$I(X;Y) = H(X) + H(Y) - H(X,Y), \tag{2.8}$$

with $H(X)$ being the entropy of a random variable X and $H(X,Y)$ the joint entropy[1]. When X and Y are independent, the mutual information will be zero, since

$$p(x,y) = p(x)p(y).$$

Conversely, the more X and Y depend on each other, the higher the mutual information. Figure 2.7 illustrates that this offers a way to quantify to what extent two parameters are related. For our (fake) example data, we have used the

```
kde2d
```

function written by Zdravko Botev [50], a kernel density method, to estimate the underlying distributions (i.e. the marginal distributions and the joint distribution). However, to apply this analysis to experimental data, one would need to statistically quantify the validity of the estimated probability distributions, or use a statistical method that provides an estimate of the mutual information directly.

2.4.5. OBTAINING CROSS-CORRELATIONS FROM EXPERIMENTAL DATA

The experimental nature of the data that we gather presents us with specific challenges regarding the cross-correlation analysis. In this section, we will discuss how we obtain cross-correlations and scatter plots from the experimental data.

EXPERIMENTS LEAD TO A BRANCHED LINEAGE TREE

In a typical experiment, we measure biological quantities over time in growing and dividing bacteria. Since we are interested in processes that have time scales similar to or longer than one bacterial life cycle, we measure over multiple generations of bacteria (depending on experimental conditions, the generation time of bacteria is typically 20 minutes to 5 hours). For microcolonies growing on gel pads, this leads to a branched data structure, as was previously described [17, 18], see also Figure 2.9.A. For microcolonies growing in the microfluidic device as described above (see also Figure 2.4.B), cells disappear from the analysis when they exit the well. This means that the datastructure will also contain lineages that end before the end of the experiment (Fig 2.9.A, branch 4).

---

[1]See also http://mathworld.wolfram.com/MutualInformation.html.

DEALING WITH REDUNDANCY IN THE BRANCHED LINEAGE TREES

The first challenge is to use the time series from each branch of the lineage tree to create a composite cross-correlation or scatter plot that represents all data. For each branch, cross-correlations and scatter plots can be determined by taking into account all pairs of points with a specific delay in between them. However, this would result in pairs of points being used more than once, since they appear in multiple branches. For example, the point pair labeled [1] in Figure 2.9.A appears in all branches, and would thus be overrepresented if no correction is applied. This can be addressed by only taking into account point pairs that are unique [15]. For the scatter plots, when we plot point pairs $f_i(n)$, $g_i(n + \tau)$ for all time points n and all branches i, we simply omit point pairs that are duplicates. For the cross-correlation, we follow a similar procedure. To determine the final composite cross-correlation, we employ a weighing scheme that grants less weight to point pairs that are redundant in the dataset. (This was also described earlier [15, 17, 18]). Specifically, the cross-correlation as defined in Equation 2.3 is adjusted to obtain a composite cross-correlation with contributions of points from multiple branches i:

$$S_{\text{composite,f,g}}(\tau) = \frac{1}{W_{\text{total},\tau}} \sum_i \frac{1}{N_i - |\tau|} \sum_{n=0}^{N_i - |\tau|} w_{n,i,\tau} \tilde{f}_i(n) \tilde{g}_i(n + \tau), \qquad (2.9)$$

with $w_{n,i,\tau}$ a weighing factor that corrects for the redundancy of the specific point pair $\tilde{f}_i(n)\tilde{g}_i(n + \tau)$ (see below for a more detailed discussion of the value of $w_{n,i,\tau}$). Furthermore,

$$W_{\text{total},\tau} = \sum_{n,i} w_{n,i,\tau}$$

is the sum of all weights for a specific τ value. (Because any branch i has $N_i - \tau$ pairs of points, the total weight changes per τ value.) $R_{\text{composite}}(\tau)$ is then again determined by normalizing with

$$\sqrt{S_{\text{composite},f,f}(0) S_{\text{composite},g,g}(0)}.$$

Also, in this formula, $\tilde{f}_i(n)$ is defined as

$$\tilde{f}_i(n) = f_i(n) - \langle f_i \rangle_n,$$

i.e. the mean of the value of $f(n)$ for all cells in the colony at that point in time (n) is determined and subtracted from the value of $f_i(n)$ to obtain the mean-subtracted value $\tilde{f}_i(n)$ for branch i. In the rest of this thesis, we refer to $R(\tau)_{\text{composite}}$ also simply as the cross-correlation, $R(\tau)$. The most straight-forward weighing simply aims to achieve what was also done for the scatter plots, i.e. make sure that duplicate point pairs are counted only once. Thus, the most straight-forward choice of $w_{n,i,\tau}$ is

$$w_{n,i,\tau} = 1/\lambda_{n,i,\tau} \qquad (2.10)$$

with $\lambda_{n,i,\tau}$ being the number of times this specific point pair was used. This method still leaves some redundancy in the data, as there might be point pairs that in which

one point of the pair was already used, but the other point is unique (this is illustrated at the label [2] in Figure 2.9). Note that the uniqueness of a point pair is determined completely by the point in the pair that was measured at the latest time point in the dataset. This issue is not straightforward to solve, but other choices of weights $w_{n,i,\tau}$ that try to address this issue are described in [17] and [18]. Additionally, one can consider additional weighing schemes, e.g.

$$w_{n,i,\tau} = 1 / \left(\lambda_{n,i} \cdot \lambda_{n+\tau,i} \right),$$

where $\lambda_{n,i}$ and $\lambda_{n+\tau,i}$ are to the number of times the specific points making up the pair are considered for the composite cross-correlation. Nevertheless, it is hard to choose the perfect weighing scheme, since the nature of the cross-correlation requires using pairs of points that are separated by a time delay, and ergo pairs might consist of two points that have a different degree of uniqueness. One could choose to throw away any pair that contains a redundant point, but then one would also throw away (valuable) information. In this work, to be consistent with the scatter plots, we use the weights as described in Equation 2.10, this is also consistent with [15].

DEALING WITH IMPARTIAL TIME SERIES

As mentioned, the experimental data of the microfluidic device also contains lineages that end before the end of the experiment (branch 4 in Figure 2.9.A illustrate this). Such lineages of course also produce time series that end before the experiment ends (Figure 2.9.B, bottom panel). The fact that branches are of unequal length can be dealt with in a straightforward manner when computing the composite cross-correlation function. Equation 2.9 allows for the fact that branch lengths N_i are different for each branch i. The weighing terms do not need to be adjusted, as $w_{n,i,\tau}$ and $W_{\text{total},\tau}$ can be simply calculated for impartial data series. Figure panels 2.9.B-C illustrate this procedure.

CONTROLS

Now that we have a way to determine cross-correlations for experimental data, an important question that remains is "is the signal that we observe real"? We are aware of two scenarios that might lead to a false positive signals. One scenario is an under-sampled dataset. Particularly when timescales of fluctuations are long, e.g. multiple generations of cells, and the size of fluctuations is high, a particular real temporal dynamic might dominate the dynamics of the dataset, and lead to a particular cross-corerelation. This cross-corerelation might then however not characterize the entire dynamics of the parameters of interest. Another scenario could result from a fluctuating mean signal. In principle, the aim is to measure in steady state, but experimental conditions might not be perfect and lead to a fluctuating mean signal, that transpires to the cross-correlation despite the subtraction of the colony mean. (Potentially, there might also be real colony-wide dynamics.) The latter scenario underlines the importance of also visually inspecting the signals of all lineages (Figure 2.10.A-C). Both these scenarios underline the importance of replicate experiments. Nevertheless, also for data from one microcolony, one can quantify parameters that help determine the validity of the calculated cross-correlation. To provide a control based on the data of one microcolony, we determine the cross-correlation for time series f_i and g_j, where f_i comes from an arbitrary branch
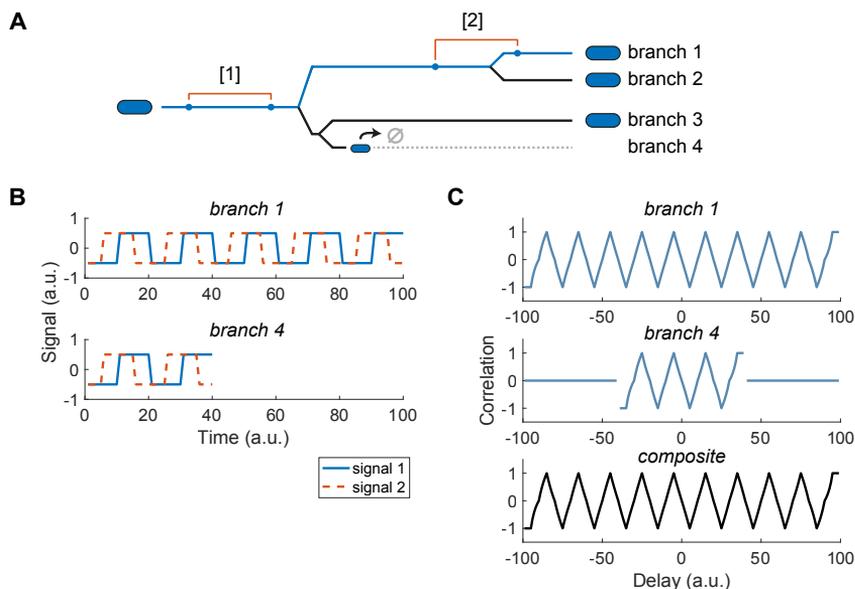
2

Figure 2.9: (A) As cells divide, the data from microcolony experiments will show a branched struc-
ture. This leads to redundancy in the data. For example, data recorded from the cell on the left
will show up in all branches (1-4) of the lineage structure in this cartoon. Additionally, in the mi-
crofluidic device described in this chapter, cells are washed away from the growth well, and thus
from the experimental observations. This is shown in branch 4 in this cartoon. (Note that in a
growth well, the number of cells will remain approximately constant, and lineages originate from
a few cells that remain at the bottom of the well; not shown here). (B) Cartoon representation of
two signals that might come from branches 1 and 4. A block function would produce a saw-tooth
correlation function (see panel C). (C) To determine cross-correlation for these cell branches that
are not present during the whole experiment, a weighing scheme is used that does not simply av-
erage the cross-correlation of the two separate branches, but instead goes over each pair of points
in the branch structure, and weighs the contribution of these two points to the composite cross-
correlation based on the redundancy. For example, the combination of points labeled [1] in panel
(A) will receive less weight than the combination of points indicated by [2]; the aim is to com-
pensate for the fact that points [1] will appear in the time series of all 4 branches and will thus be
considered four times. This weighing is complicated by the fact that for some pairs of point, one
point is more redundant than the other (this is the case for points [2] in panel (A)). See main text
for a more detailed description of the weighing procedure.

2

i and g_j from an arbitrary branch j; in other words, we combine time series from two different lineages, which should not correlate, and ergo cross-correlations should be 0. Specifically, for gel pad data all branches are of equal size, and whole time series f are combined with arbitrary time series g from another branch. For microfluidic data from the device described in this chapter, branches are not of equal length, and thus randomization of branch data $f_i(n)$ and $g_j(n)$ is performed for each point in time n; this might be slightly less rigorous as this further decorrelates the data. In both cases, we repeat this procedure 50 times, and determine the minimum and the maximum value of $R(\tau)$ resulting from these 50 repeats. This provides an estimated region of $R(\tau)$ where a non-zero signal might be observed that is likely false-positive.

Additionally, error bars can be estimated by dividing the data from one colony in multiple groups, and analyzing these groups independently. Lastly, as a sanity check, the cross-correlation determined from the scatter plots determined for all different delays τ can also be determined. The result of these procedures is shown in Figure 2.10.D.
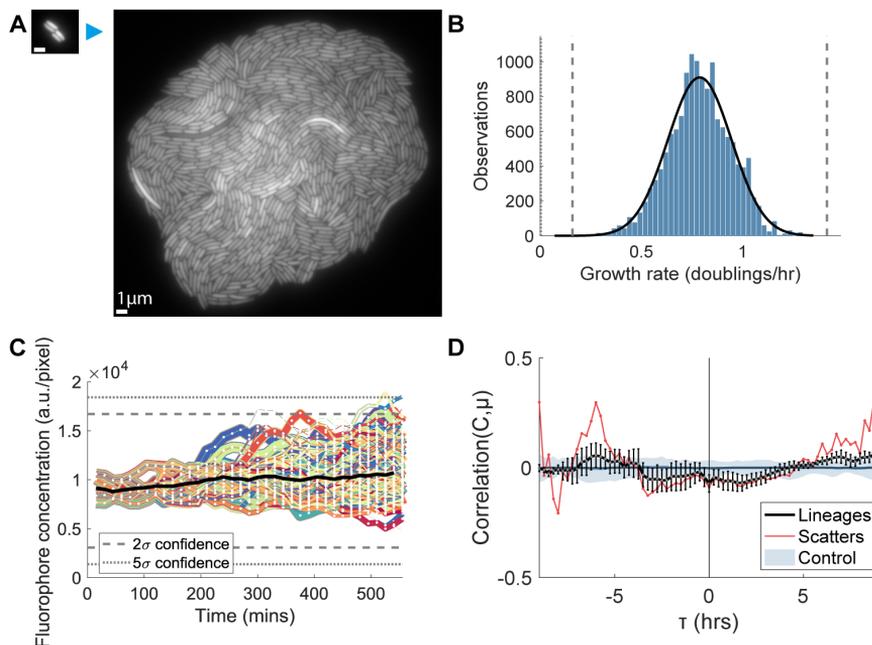
ACKNOWLEDGEMENTS

2

Figure 2.10: (A) Example of fluorescence images from a gel pad experiment, where typically a small colony (left) grows into a bigger colony (right). (B) Distribution of growth rates in the microcolony, the black line shows a fitted normal distribution, the dotted line shows values located at 2 standard deviations (σ) from the mean. (C) Fluorescent concentrations over time for all lineages from the microcolony shown in panel (A). Values over the lineages are shown on top of each other (lineages that are plotted behind others are given thicker lines). Note that values overlap at the beginning due to the branched structure of the data (see Figure 2.9). The dotted lines indicate values that are at a distance of 2 and 5 standard deviations (σ) from the mean, as determined by fitting a normal distribution. (D) The composite cross-correlation as determined between the instantaneous growth rate (μ) and fluorescence concentration (C) values over the lineages in this microcolony (black line). The data was subdivided in 4 groups, which allowed for the determination of standard errors. For each delay τ also a scatter plot was made from which the correlation coefficient was determined, which is shown here in red. (Note that this might deviate slightly depending on the used weighing procedure and the averaging over different sub-groups of data.) Additionally, as a control, growth rate data from the lineages was combined with fluorescence data from random other lineages, and by repeating this procedure 50 times correlation values were determined that should not be interpreted as meaningful deviations from zero.

3

SIZE LAWS AND SEPTAL RING DYNAMICS IN FILAMENTOUS ESCHERICHIA COLI CELLS

IN BRIEF

Wehrens, Ershov, et al. report a new size control mechanism in E. coli cells, which have elongated due to stress. Multiple division rings are continuously rearranged in response to growth and division to control daughter cell size when divisions resume. Divisions are spatially controlled by the Min system and temporally by the adder principle.

HIGHLIGHTS

- Long filamentous E. coli cells continuously assess and adjust division ring positions.
- Sudden division site suppression and formation is explained by the Min system.
- Division timing is consistent with the adder principle.
- These rules enable controlled entry in and exit from the filamentous state.

3
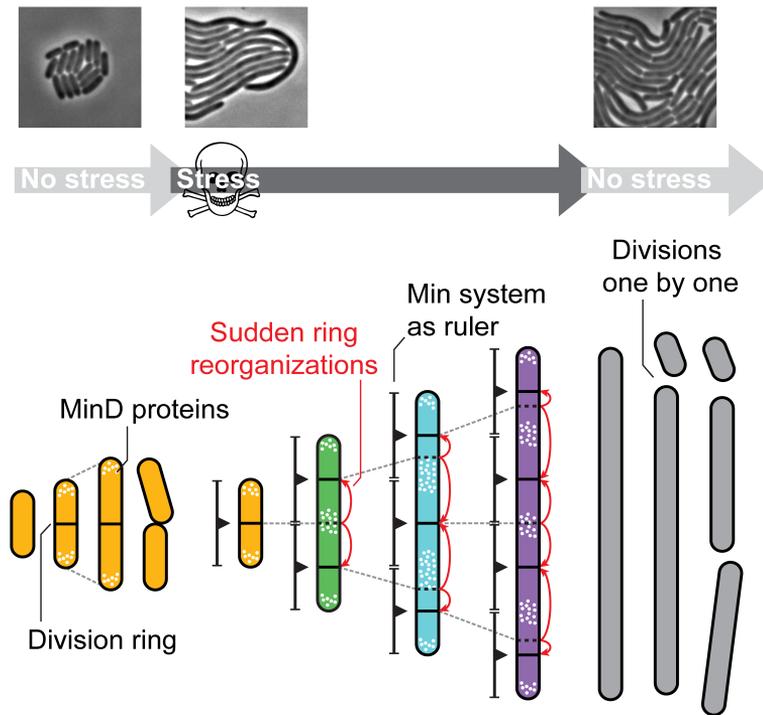
Graphical abstract.

SUMMARY

Our understanding of bacterial cell size control is based mainly on stress-free growth conditions in the laboratory [33, 52–60]. In the real-world however, bacteria are routinely faced with stresses that produce long filamentous cell morphologies [61–78]. E. coli is observed to filament in response to DNA damage [72–75], antibiotic treatment [61–64, 78], host immune systems [65, 66], temperature [67], starvation [70], and more [68, 69, 71]; conditions which are relevant to clinical settings and food preservation [76]. This shape plasticity is considered a survival strategy [77]. Size control in this regime remains largely unexplored. Here we report that E. coli cells use a dynamic size ruler to determine division locations combined with an adder-like mechanism to trigger divisions. As filamentous cells increase in size due to growth, or decrease in size due to divisions, its multiple Fts division rings abruptly reorganize to remain one characteristic cell length away from the cell pole, and two such length units away from each other. These rules can be explained by spatio-temporal oscillations of Min proteins. Upon removal of filamentation stress, the cells undergo a sequence of division events, randomly at one of the possible division sites, on average after the time required to grow one characteristic cell size. These results indicate that E. coli cells continuously keep track of absolute length to control size, suggest a wider relevance for the adder principle beyond the control of normally sized cells, and provide a new perspective on the function of the Fts and Min systems.

3.1. RESULTS

3.1.1. DIVISION SITE SELECTION RULES IN FILAMENTOUS CELLS

To investigate divisions in filamented Escherichia coli cells we used a microfluidic device that allows media exchange [32] (Fig. S3.5.A-C). We first grew the cells for 2-3 generations at 37 °C in minimal medium within the device, and then induced filamentation in one of three ways: exposure to tetracycline (TET), a temperature increase to 42 °C, or overexpression of SulA. While the molecular basis in the former cases is unclear, translation inhibition antibiotics have been reported to induce filamentation [78], and SulA is a division inhibitory protein [79]. As a result the cells grew to approximately 10-20 times the typical length without dividing (Fig. 3.1.A), while their width remained approximately constant (Fig. S3.5.D). Division resumed when the stressor was removed and ultimately returned to normal stress-free growth and division, unless the stress was too severe and the cells failed to recover (Fig. S3.5.E-G). The relative location of division events throughout this recovery process was characterized by $S = L_d/L_m$, where L_d and L_m are the daughter and mother cell lengths respectively (Fig. 3.1.B).

We found that S displayed a specific pattern when plotted against L_m (Fig. 3.1.D and S3.5.H-K), which was for instance not the case when plotted against the time elapsed since stress removal (Fig. 3.1.E and S3.5.L). The same S-L_m pattern was observed for the different levels and types of stress (Fig. S3.5.H-K). Several features were as expected. First, normally sized cells of a few microns long divided in the middle, with $S = 1/2$. Second, for increasing L_m, other division sites appear with S-values that remain constant within a certain length-window (Fig. 3.1.D). These features are consistent with a long-standing model of division site locations (see Fig. 3.1.C), which was supported by early

3
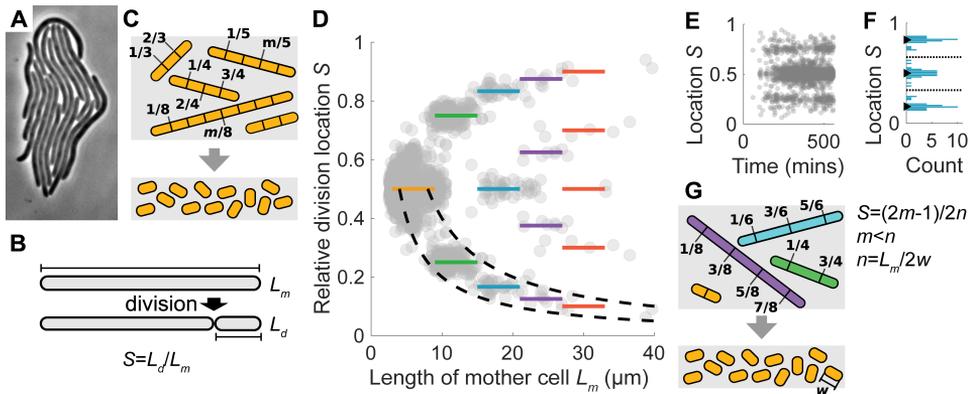
Figure 3.1: Division site rules in filamentous E. coli. (A) Colony of filamentous cells. (B) Relative division site S, daughter length L_d, and mother length L_m. (C) Possible division sites in classical scenario, with a minimal equidistant spacing, and hence more sites for longer cells. (D) Division sites (see panel B) observed in filamentous cells when switching from stress medium (1 μM tetracycline) to non-stress medium. Each point corresponds to one division (N = 4108). Data is symmetric around $S = 1/2$. Colored lines denote inferred approximate division location (see panel G). Dashed lines correspond to divisions producing daughter cells of 2 and 4 μm. (E) As panel (D) but as a function of the time of the division event. (F) The distribution of relative division locations for cells with a maternal size between 15-21 μm, suggesting a lack of site preference. Dashed black lines indicate bin boundaries. (G) Division site rules inferred from the data in panel D. n is the number of possible division sites, m indexes the possible sites in one cell, w is a characteristic length (about 3μm). See also Figure S3.5.

time-lapse microscopy experiments in the 70s [80], detection of invaginations in cell division mutants [81], and consistent with observed roughly equidistant division rings at multiple locations [82–84] and uniform cell wall growth [85]. In this model, during filamentous growth, the distance between division sites increases progressively, until this distance doubles and new sites emerge in between existing sites. Note that this scenario can be compared to normally dividing cells arranged along a line, as new division sites would then also emerge in between two previous division sites.

At the same time, other features differed substantially. Fig. 3.1.C suggests moderately filamented cells divide at S = 1/4, 1/2, and 3/4. However, for cells with $S = 1/4$ and $S = 3/4$, the site at $S = 1/2$ was repressed in most of that length-window (Fig. 3.1.D, green bars). Other sites appeared suppressed as well. For instance, the data did not show cells with two rings at $S = 1/3$ and 2/3, or with other odd fractions such as multiples of 1/5 (Fig. 3.1.D). We did not detect a significant spatial preference: the different possible sites within the length window 15-21 µm displayed a similar probability to divide (Fig. 3.1.F). On the other hand, the observed sites could be captured in specific rules (Fig. 3.1.G). The distance between two possible division sites equaled two normally-sized cells (rather than one as in Fig. 3.1.C), though the distance between a pole and the nearest division site was one normally-sized cell. More generally, the different possible sites within one length-window was described by $S = (2m - 1)/2n$, with m indexing the sites from 1 to the total $n = \text{round}(L_m/2w)$, and w = 3 µm the average length of unstressed cells. These rules also indicate discontinuous changes. For instance the S values of the green bars (1/4 and 3/4) disappear in the subsequent length regime with the blue bars (1/6, 3/6, and 5/6), and so on (Fig. 3.1.D). Thus, potential division sites appear not to be conserved upon changes in length, unlike the model indicated in figure 1C. To further probe these various unexpected findings, we analyzed the recovery process in time.

3.1.2. TIME AND LENGTH CHANGES BETWEEN DIVISIONS

To study the recovery process in time, we quantified cellular lengths along lineages after stress removal. The resulting traces displayed sudden drops denoting division events, as the long cells progressively converted into normally sized cells (Fig. 3.2.A and S3.6.A,D). Notably, we rarely observed multiple divisions occurring at the same time within one cell – after a single division event some time elapsed (for cells over 10 µm long recovering from tetracycline exposure, 89% of the interdivision times were 10 min. or more) before the next division took place. Thus, in cells that were long enough to fit multiple possible division sites, division occurred in just one site at a time.

The interdivision time was also found to decrease with increasing cell length; cells that were born longer divided faster (Fig. 3.2.B). However, the interdivision time appeared to level off at about 10 min, consistent with divisions in a minCDE null strain and cells having a limited 'division potential' [86]. This dependence on cell length was similar for the different filamentation triggers (Fig. 3.2.C and S3.6.B,E,G). The tetracycline and temperature data differed significantly in one bin only (two sample t-test, p = 0.05), while the SulA data set showed a somewhat smaller interdivision time. The latter may reflect the more downstream role and limited growth-rate effect of SulA. Interestingly, the length added between two divisions appeared constant and independent of birth size (Fig. 3.2.D, S3.6.C,F). This added length was similar to the length of normal new-

3
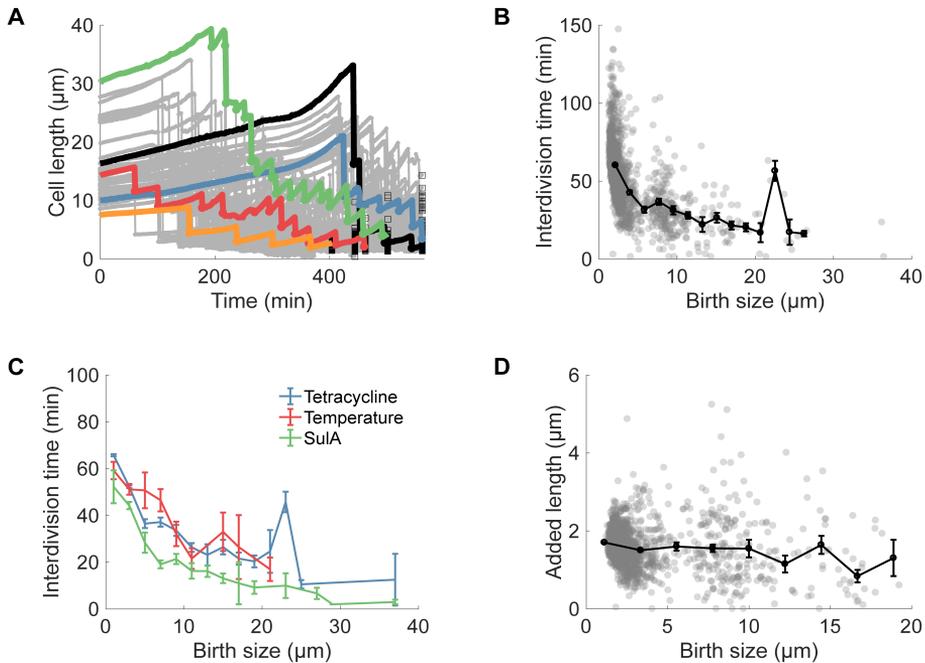
Figure 3.2: Time and length changes between divisions. (A) Measured cell lengths over time after switching from stress medium (1 μM tetracycline) to non-stress medium. Grey traces correspond to single cell lengths. Colored lines are lineages following the longest daughter, black line the shortest daughter. Black squares indicate end of measurement. Data is from five 1μM tetracycline recovery experiments (N = 4134 cells). (B) Interdivision time against mother birth size, for recovery from 1μM tetracycline. Black dots are averages, bars are standard deviations (N = 4108 division events). (C) Average interdivision time against mother birth size, for recovery from 1μM tetracycline (blue), 42 °C heat shock (red), and overexpression of the division-inhibitor SulA (green) (N=4108, N=404, and N=494 division events, respectively). Error bars show SEM. (D) Absolute length added between two divisions against mother birth size. Black dots are averages (N = 4108 division events.), error bars show SEM. See also Figure S3.6.

born cells (1.6 vs. 1.8 microns respectively, Fig. S3.6.H). Thus, strikingly, long and short new-born cells grow a similar absolute amount until the next division, while the shorter cell takes more time to produce that added length. The latter is consistent with larger cells producing cell mass faster because they contain more biosynthetic machinery (Fig. S3.6.I). Consistently, we find that the average area under the length versus time curve is proportional to the added length divided by the growth rate (Figure S3.6.J).

A constant added volume or length between divisions has been reported for normal stress-free growth, when cells are not filamentous and divide mid-cell [33, 57–59]. This "adder" principle can explain how cells in these conditions maintain a constant average size despite stochastic variability in birth size [33, 58, 59]. Here, the adder principle does not strictly govern size, as birth sizes of filamentous cells are affected more strongly by the position of division sites. The findings suggest that the time between divisions is coupled to the growth process. The added length in stress-free cells has been reported to be proportional to the number of origins of replication [87, 88]. We found nuclei to continue to multiply during filamentation in accordance with length increases, as observed by fluorescent labeling (Fig. S3.6.K), which suggests that the origins continue to increase as well. Note that the unstressed interdivision time here is about 60 mins (Fig. 3.2.C), which is associated with non-nested replication. Thus, while the number of nuclei differs several fold between normal and filamented cells, the added size between divisions remains approximately the same, suggesting that the number of origins does not set the added size in this stressed regime. We find that the cell volume to nucleoid ratio was approximately constant for cells of different length within the filamentous regime (Fig. S3.6.L). The ratio of cell volume to the number of replication origins was studied previously for non-filamentous cells [89, 90].

3.1.3. DYNAMIC REORGANIZATION OF DIVISION RINGS

The data revealed that the reductive divisions during recovery occur in concert with substantial cellular growth (Fig. 3.2.A and S3.6.M). To illuminate how continued growth affected division site positions, we imaged the division machinery by fluorescently labeling the essential cell division protein FtsA with sYFP2 [91, 92]. As expected, we observed a number of bands of fluorescence intensity along the cell axis indicating division rings [82–84] (Fig. 3.3.A). When tracked in time however, these division rings displayed an unexpectedly dynamic behavior, both during growth of the filament and upon division events. For instance, upon a division, a new ring appeared at a new location (Fig. 3.3.B, top arrow), within the smaller daughter. At the same time, one of the other rings that had been present in the mother cell, disappeared in the other daughter at the same moment (Fig. 3.3.B, bottom arrow). Indeed, most division events were immediately followed by the rapid reorganization of division rings (Fig. 3.3.B). Independently of division events, growing filamentous cells also showed sudden reorganization of ring positions (Fig. 3.3.C, S3.7). For instance, all four rings in one cell abruptly disappeared at the same time, while five rings appear in the next frame at different positions (Fig. 3.3.C, grey arrow).

Once formed, division rings often remained fixed at the same position for over 50 mins (Fig. 3.3.B). The sudden reorganization dynamics of these mature division rings indicated that such positional memory can be erased rapidly. In early stages of ring forma-

3
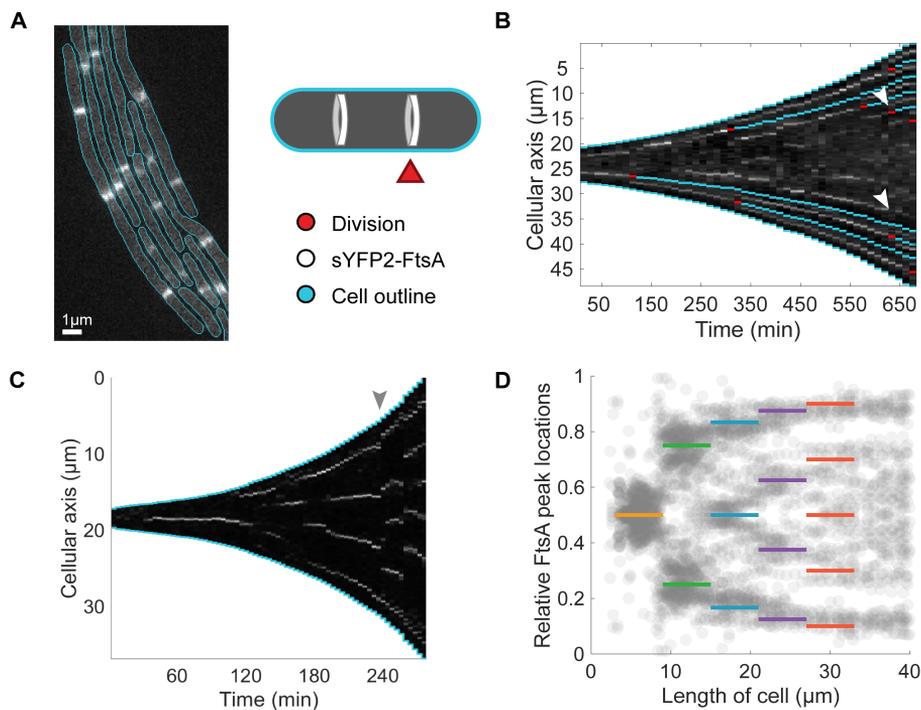
Figure 3.3: Dynamic reorganization of division rings. (A) Fluorescence image of sYFP2-FtsA signal in cells filamented by heat shock at 42 °C. FtsA is an essential component of the division apparatus. Blue lines indicate cell edge determined from phase contrast imaging. (B) Kymograph of sYFP2-FtsA intensity profiles along the long cellular axis, for cells recovering from 1µM tetracycline. Red indicates division events, blue the resulting separation between daughters. See main text for arrow. (C) Similar kymograph for filamentous cells during heat shock at 42 °C. (D) Relative locations of sYFP2-FtsA peaks along the cellular axis, during heat shock at 42 °C and subsequent growth at 37 °C. Colored bars correspond to inferred division site rules (Fig. 3.1.D-E). Peaks were identified using a Matlab peak finder algorithm (N = 1572 cell images, taken from 400 cell cycles). See also Fig. S3.7.

tion, FtsZ clusters have been reported to display positional dynamics in non-filamentous cells [93]. The observed ring dynamics also explained the observed 'suppression' of division events at certain sites: it for instance allows a central ring at $S = 1/2$ in a normally sized cell to disassemble when filamentous growth causes entry into the second length window with rings at $S = 1/4$ and $3/4$. Similarly, it explains when rings are added. Cells do not add multiple rings when the distances between existing rings has doubled, but rather add a single ring when the total cell length has become $2w$ longer, and correspondingly the entire pattern of rings changes. Indeed, more generally, the observed division ring positions (Fig. 3.3.D) followed the same spatial rules as the division event positions (Fig. 3.1.D,G). Note that one does not observe division rings at all the sites denoted by the spatial rules (Fig. 3.1.D,G), such as the middle region of the cell depicted in Fig. 3.3.B. Furthermore, as expected, we did not see division events unless a fluorescent band was observed first (see Fig. 3.3.B). In cells with multiple rings, division typically occurred at one site at a time (Fig. 3.3.B, see also Fig. 3.2.A-B). The fluorescence intensity of the bands did not correlate significantly with division probability (Fig. S3.7.B).

Thus, the dynamics of the division rings is important to obeying the spatial division rules. Growth and division events can 'move' lineages from one length-window to another, and hence require changes in the pattern of possible division sites. Note that for cellular growth within a length window, rings do not change their relative positions. Thus, the cells continuously assess the locations of their division machinery, and repositions the potential division locations, resulting in daughter cells of specific size.

3.1.4. MIN OSCILLATIONS CAN EXPLAIN DIVISION SITE RULES

The spatial position rules indicate absolute bounds for the distance between putative sites. Specifically, this distance is on average between 1.5 and $2.5w$, with w being approximately 3 μm. These observations indicate a mechanism able to measure absolute lengths. We surmised that this function is served in *E. coli* by proteins encoded at the min locus: MinC, MinD, and MinE. MinD can bind to the cell membrane, and the MinD and MinE reaction-diffusion system results in pole to pole oscillations of MinD along the cell membrane. In non-filamentous cells, the time-averaged concentration profile of MinD shows a minimum value localized at $S = 1/2$. Because MinC binds MinD, but also inhibits ring formation, non-filamentous cells form division rings mid-cell [94, 95].

To test whether MinD oscillations could contribute to the division rules in filamentous cells, we extended the reaction-diffusion model of Meinhardt and De Boer [96] (Materials and Methods). Strikingly, the simulations recapitulated all key features of the spatial position rules (Figs 3.1.D,G and 3.4.A): distinct length windows of equal size, discontinuous changes in all minima positions between adjacent windows, and absolute bounds for the distance between minima. A similar MinD pattern was obtained using model of Huang et al. [97] as implemented by Fange et al. [98, 99] (Fig. S3.8.A). To obtain further support for a role of the Min system in the position rules, we used a YFP-MinD fusion to assess the MinCD distribution. The resulting time-averaged fluorescence profiles validated the simulation results (Fig. 3.4.B).

Next, we characterized filamentation and recovery in minCDE null cells. Divisions no longer obeyed the division rules (Fig. 3.4.C). Division sites were more uniformly distributed, and for instance no longer peaked at $S = 1/4$ and $3/4$, nor showed the previously

3
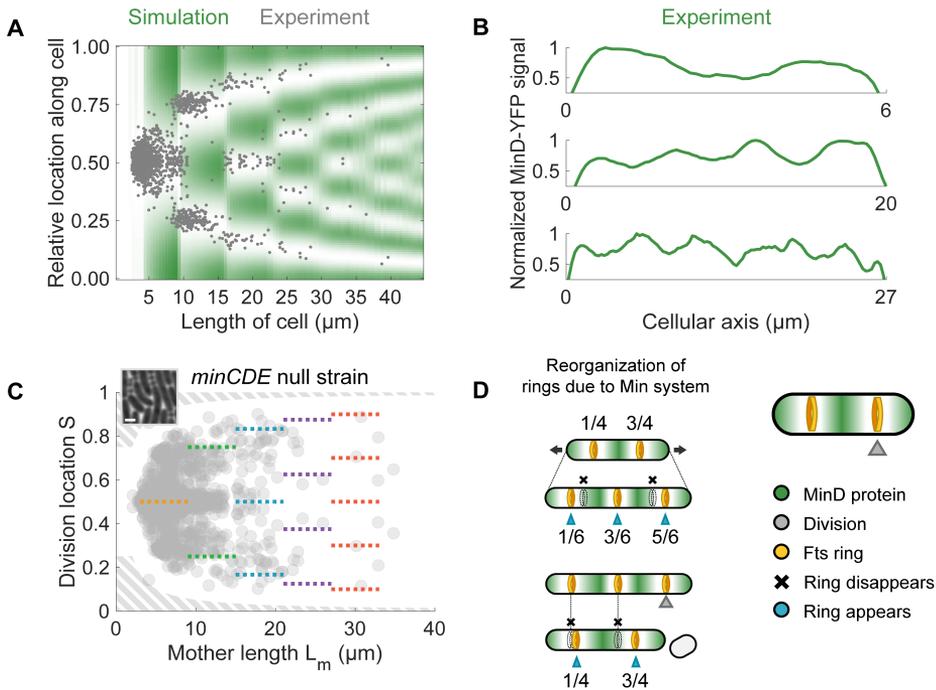
Figure 3.4: Min oscillations can explain division site rules. (A) Computed time-averaged MinD concentration profile along the longitudinal cellular axis for cells of different lengths, using the Meinhardt and De Boer [96] model. Green corresponds to high MinD concentrations. Grey dots are observed divisions (Fig. 3.1.D). Simulation results are scaled linearly along x-axis for best correspondence with the experimental data. (B) Experimentally obtained time-averaged intensity profiles of YFP-MinD fusion proteins for cells of different sizes. Top to bottom indicated 1, 3 and 5 minima. Low values at cell poles are artefacts due decreasing cellular width. (C) Division sites in minCDE null strains filamented by 2 μM tetracycline. Each point corresponds to one division (N = 1260). Data is symmetric around S = 1/2. Colored dotted lines denote division rules in wild type cells (Fig. 3.1.D). Minicells were observed (inset), but were not included in the analysis (grey shaded region). (D) Cartoon illustrating the observed division site plasticity. Changes in cell length, due to continued cell growth (top) or division events (bottom) produce discontinuous changes in the MinD profile, and corresponding reorganization of the pattern of division rings, and ultimately locations of division events. See also Figure S3.8, Table S3.1 and S3.2.

observed suppression at $S = 1/2$ or other S-values. The positions were not completely arbitrary, with $S = 1/2$ remaining frequent, suggesting the involvement of other mechanisms such as nucleoid occlusion. The interdivision time was similar as observed for wild-type cells, and consistent with earlier observations [86] (Fig. S3.6.N-P). While the length added between divisions remained independent of cell length, it was now more broadly distributed (Fig. S3.6.O). Moreover, the recovery produced numerous non-viable mini-cells [86, 100] that were arrested in their cell cycle and stopped growing. These findings further support the notion that the Min system is central to the spatial division rules in filamentous cells.

3.2. CONCLUSIONS

Cell size control of *E. coli* is actively studied in continuous laboratory cultures, where cells grow in the bacillary form (see e.g. [33, 52–58, 60]). Here, size-control typically concerns maintaining a constant cell size in the presence of stochastic variability, for instance in birth size and growth rate. *E. coli* is also known to adopt the longer filamentous form in response to diverse stressors. Here we surmised that size regulation mechanisms could also be relevant to enter into, maintain, and exit from filamentous states. We find that within growing filamentous cells, multiple non-constricting division rings remain at specific fixed relative positions when the cells remain within a specific length range, while all the rings abruptly change in number and position when exceeding this range, which can be caused by growth and division events (Fig. 3.4.D). These spatial rules are explained by the Min system, which is found to produce a pattern that is strikingly similar to sound standing waves, with minima at matching positions that change discontinuously when an additional minimum fits along the cell length. Hence, the Min system can be thought of as a ruler mechanism that measures absolute size, with division rings as tick marks that have a spacing corresponding to not one but two unstressed cells. Upon disappearance of the stressor, divisions occur in sequence, at just one of these multiple putative division sites at a time, with the interdivision time decreasing with cell length. The length added in between divisions is rather invariant with cell length, in a manner that is reminiscent of the on average constant added length for non-filamentous cells that experience small variations in birth size during normal growth, suggesting that this adder principle is more broadly relevant beyond normal growth. How these long cells divide just once and then add a fixed length remains an intriguing question. One may consider additional suppression of division at certain sites by nucleoid occlusion mechanisms [101, 102], though this would not naturally explain the suppression of division for a certain amount of growth. Another possibility is a limiting septal protein that must increase in number after division [103], though the multiple chromosomes and FtzA rings were observed. Overall, it is puzzling how cells that strongly vary in size, would be limited similarly to just one division. Taken together, the findings reveal a system of size sensing and division control in filamentous *E. coli* cells, and bring a new perspective to the functional role of Min and Fts dynamics.

Acknowledgements

3

3.3. STAR METHODS

3

REAGENT or RESOURCE	SOURCE	ID
Bacterial and Virus Strains		
E. coli , strain ASC555: wild type MG1655 (ilvG- rfb-50 rph-1)	AMOLF lab	ASC555
E. coli , strain ASC884: wild type W3110 (λ- IN(rrnD-rrnE)1 rph-1) with SulA plasmid.	Gift from Alexander Dajkovic [79]	ASC884
Escherichia coli, strain ASC777: ASC555 with sYFP2-FtsA plasmid.	Gift from Svetlana Alexeeva, Tanneke den Blaauwen lab	ASC777
E. coli , strain ASC784: ASC555 with pFX40 plasmid.	Gift from Cees Dekker lab	ASC784
E. coli , strain PB114 (also known as ASC1035 or PAL40): MG1655 (F- lambda- ilvG- rfb-50 rph-1 ΔminCDE, KanR)	Gift from Piet de Boer lab [104]	PB114 / PAL40 / ASC1035
E. coli , strain SJ182 (also known as ASC1106): MG1655 low motile (F- lambda- ilvG- rfb-50 rph-1 hupA::[hupA::mCherry FRT kanR])	Gift from Suckjoon Jun lab	SJ182 / ASC1106
Recombinant DNA		
SulA plasmid: pACT3 (Plac-sulA, CamR)	Gift from Alexander Dajkovic [79]	N/A
sYFP2-FtsA plasmid: pSA018 (PTRCdown-sYFP2-FtsA, AmpR)	Gift from Svetlana Alexeeva and Tanneke den Blaauwen	N/A
YFP-MinD plasmid: pFX40 (Plac::yfp-minD minE,AmpR)	Gift from Cees Dekker lab	pFX40

KEY RESOURCE TABLE. Bacterial strains and recombinant DNA used in this manuscript.

REAGENT or RESOURCE	SOURCE	ID
Software and Algorithms		
Metamorph 7.8.0.0	Molecular Devices	[a]
Matlab 9.1.0.441655 (R2016b)	MathWorks	[b]
Custom Matlab segmentation and analysis scripts	Tans lab	[c]
Meinhardt and De Boer simulation scripts (translated to Matlab syntax in Tans lab)	[96]	[d]
MesoRD simulations software	[99]	[e]
Other		
Microfluidic SU-8 master mold for PDMS device 1	[32], MicroChem	N/A
Epoxy mold for Mother Machine like PDMS device 2	Daan J. Kiviet, Martin Ackermann lab	N/A

KEY RESOURCE TABLE. Software and miscellaneous items used in this manuscript.

[a] https://www.moleculardevices.com/

[b] https://mathworks.com/products/matlab.html

[c] https://github.com/TansLab/Tans_Schnitzcells,
https://github.com/TansLab/Tans_filamentation,
https://github.com/TansLab/Common_libraries

[d] http://www.pnas.org/content/suppl/2007/11/23/98.25.14202.DC1/p10.html

[e] http://mesord.sourceforge.net/

3.3.1. CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Sander Tans (Tans@amolf.nl).

3.3.2. EXPERIMENTAL MODEL AND SUBJECT DETAILS

For tetracycline stress experiments, wild type strain MG1655 (rph-1 ilvG- rfb-50) was used (ASC555, Key Resource Table). For experiments involving SulA we used wild type strain W3110 (λ- IN(rrnD-rrnE)1 rph-1) with pACT3 plasmid containing Plac-sulA, a kind gift of Alexander Dajkovic [79] (ASC884, Key Resource Table). For experiments involving temperature recovery and sYFP2-FtsA, we used wild type strain MG1655 with plasmid pSA018 containing PTRCdown-sYFP2-ftsA (ASC777, Key Resource Table), a kind gift from Svetlana Alexeeva and Tanneke den Blaauwen (University of Amsterdam). For experiments involving MinD dynamics, we used strain MG1655 with plasmid pFX40, containing Plac::yfp-minD minE (AmpR) (ASC784, Key Resource Table), a kind gift from the Cees Dekker Lab (Delft University of Technology). For experiments with minCDE null mutants we used strain MG1655 with the minCDE gene deleted (ASC1035, Key Resource Table), a kind gift from Piet de Boer [104]. For experiments with labeled nucleoids [105], we used an MG1655 strain with hupA::[hupA::mCherry FRT kan] (ASC1106, Key Resource Table), a kind gift from the Suckjoon Jun lab (University California, San Diego). All cell

lines were stored at -80 °C in freeze mix stocks. Before experiments, they were inoculated in TY medium and grown at 37 °C into exponential phase, then transferred to culture tubes with M9 media, grown overnight to reach exponential phase next day, and transferred to microfluidic setup or gel pad the next morning where they were supplied with M9 minimal medium. The microfluidic setup or gel pad was then imaged under the microscope. Freeze mix contained 0,7% g/ML Peptone (Bacto™, BD Biosciences) and 24% V/V glycerol (Merck) dissolved in sterile dH2O. TY medium contained 1% gr/mL Tryptone (Bacto™, BD Biosciences), 0,5% gr/mL Yeast Extract (Bacto™, BD Biosciences) and 0,5% gr/mL NaCl (Merck) dissolved in sterile dH2O. M9 minimal medium contained 47.7 mM Na2HPO4, 25 mM KH2PO4, 9.3 mM NaCl, 17.1 mM NH4Cl, 2.0 mM MgSO4, 0.1 mM CaCl2; all the chemicals were provided by Merck. M9 medium was supplemented with 0.2 mM uracil and 0.1% g/mL lactose (tetracycline, minCDE null mutant and labeled nucleoids experiments) or 0.1% g/mL maltose (temperature, SulA and MinD experiments); all provided by Merck. Tetracycline stock solutions (1mM in ethanol) were made from tetracycline powder (Merck) and stored at -20 °C for not more than 4 weeks, final concentrations were 1μM, 2μM and 10μM for tetracycline experiments and 2μM for minCDE null mutants and labeled nucleoids experiments. IPTG (Merck) stock solutions (1mM in water) were stored at -20 °C.

3.3.3. METHOD DETAILS

EXPERIMENTS WITH MICROFLUIDIC DEVICE 1

For experiments with wild type cells (ASC555) exposed to and recovering from tetracycline, microfluidic device 1 was used, see also Fig. S3.5.B. This device [32] consisted of cover slip, a polyacrylamide gel membrane (thickness, 500 μm) and a polydimethylsiloxane (PDMS) flow cell whose channel (3 cm * 3 mm * 91 μm) contained evenly spaced square pillars (400 μm, spaced by 600 μm) to ensure a uniform pressure on the membrane. The polyacrylamide gel membrane was formed by mixing 1.25 mL 40% acrylamide (Bio-Rad), 3.7 mL deionized sterile water, 50 μL 10% ammonium persulfate (Sigma) and 5 μL TEMED (Bio-Rad). 450 μL of the mixture was poured in a mold and the solution was left to polymerize for about 1.5 h. After polymerization, the gel was cut in a piece of 4 x 1.5 cm and stored in a flask with sterile water. The master PDMS mold consisted of one layer patterned by negative phototransparency masks on a silicon wafer. This layer was deposited using SU-8 (MicroChem). The PDMS flow cell was fabricated by molding silicone elastomer (Sylgard 184, Dow Corning) to this master mold. PDMS was mixed in a 1:10 (v/v) ratio of catalyst and resin, poured into the master mold, degassed for 1 h in a desiccator and cured in an oven at 75 °C for 1 h. To perform experiments 1uL (OD ≈ 0.005) of the desired culture was pipetted on a cover slip and covered by the polyacrylamide membrane and then the microfluidic device. The device was then connected to two syringe pumps (ProSense, NE-1000 and NE-300) by polyethylene tubing of 0.58 mm internal diameter (Smiths medical International Ltd.). The flow was controlled by a manual valve (Hamilton, HV 4-4). The culture medium flow rate during the experiments was 60 μL/min.

3

Experiments with microfluidic device 2

Another PDMS device was used for experiments involving the minCDE null mutant (ASC-1035) and labeled nucleoids strains (ASC1106). The device was developed by Daan J. Kiviet in the Martin Ackermann lab. It is similar to the device described in ref. [33], but it has wider microcolony wells. It contains a 200 μm wide main flow channel, splitting into two 100 μm wide flow channels (both 23.5 μm high). Perpendicular to these flow channels are 5 times repeated 0.75 μm high cavities (also known as "wells", where microcolonies of cells grow during the experiments) with widths of 1x 80μm, 1x 60μm, 2x 40 μm, 3x 20 μm, 3x 10 μm + 3x 5 μm, and depths of 60 μm, 30 μm, 50 μm and 40 μm. The PDMS devices were made by casting them into an epoxy mold, a gift from Daan J. Kiviet and the Martin Ackermann lab. To produce the PDMS device, polymer and curing agent (Sylgard 184 elastomer, Dow Corning Corp.) were mixed in ratio of 1 mL of curing agent to 7.7 g of polymer (we found this deviation from the recommended 1:10 ratio provided a better rigidity of the PDMS). This mixture was cast into the epoxy mold. Air bubbles were removed from the mixture either by putting the mold and casting in a desiccator for 30 minutes, or by leaving the mixture for several hours before casting. The mold and casting were then put in an 80 °C oven for 1-12 hours. Subsequently, the mold was removed from the casting, and holes were punched for the liquid in- and outlets. The casting was cut into a smaller size using a scalpel to remove rough, raised or uneven edges. Then the PDMS casting was covalently bound to a clean glass cover slip by treating the PDMS and glass surface with a portable laboratory corona device (model BD-20ACV, Electro-Technic Products, Inc.) (5-10 sweeps of approx. 5 seconds for each surface from approx. 5-10 mm distance). The casting was gently tapped using a gloved finger to improve contact between the PDMS and glass surfaces. Consecutively, the device was baked for another 1-12 hrs and stored for a couple of weeks before the experiment was started. To perform an experiment, 2 mL culture of *E. coli* was grown to high OD (>1) in a 10ml Falcon culture tube on a rotator at 37 °C. The concentration of bacteria was further increased 30x by spinning down 1ml of the sample in an Eppendorf tube at 2300 RCF, removal of supernatant and resuspension. To inoculate the device, first, 1 μL of sterile 0.01% Tween (dH2O) solution was slowly pipetted into the PDMS device, after which 1 μL of the concentrated culture was introduced in the device. When bacteria had penetrated the growth wells, the device was connected to polyethylene tubing, pumps and a valve controller similar to the other microfluidic device. Superfluous bacteria in the flow channels were removed by the culture medium flow. The flow rate during these experiments was 16 μL/min.

Experiments with gel pads

Experiments with the SulA strain (ASC884) and the sYFP2 labeled FtsA strain (ASC777) were performed on polyacrylamide gel pads. To produce polyacrylamide pads [10], a mold was created by placing two 25 mm x 76 mm x 1 mm silanized microscopy glass slides (Thermo Scientific) on top of each other. The top glass slide contained a 18 mm x 52 mm rectangular hole, and the two slides were sealed together with high vacuum grease (Dow Corning). Polyacrylamide mix (1.25 mL 40% acrylamide, 3.7 mL deionized sterile water, 50 μL 10% ammonium persulfate, 5 μL TEMED) was poured into the cavity and covered by a silanized coverslip. The mix was placed at room temperature for half

an hour to allow polymerization, and then cut into gel pads of approx. 5 mm x 5 mm x 1 mm which were stored in sterile dH2O. To perform experiments, a gel pad was soaked in the desired medium, placed in the cavity of a clean two-glass slide setup identical to the mold (except that glass slides were not silanized), inoculated on top with 1 uL (OD ≈ 0.005) of the desired bacterial culture, covered with a glass cover slip, and mechanically sealed with a metal clamp to avoid drying of the sample (see also Fig. S3.5.C).

SINGE CELL MICROSCOPY

In tetracycline experiments, strain ASC555 (see Key Resource Table) was grown in the microfluidic device first in clean M9 medium, then M9 medium with 1μM tetracycline (or 2μM or 10μM tetracycline for supplemental datasets), and then clean M9 medium again (see experimental model and subject details for detailed information on growth media). For temperature experiments, strain ASC777 was grown on M9 medium soaked polyacrylamide gel pads subsequently at 37 °C, 42 °C and 37 °C; sYFP2-FtsA expression was induced with 3.5μM IPTG. For SulA experiments, SulA expression was induced in strain ASC884 with 200μM IPTG during O/N growth in culture tubes with M9 medium, and cells were transferred to polyacrylamide pads soaked in clean M9 medium next day. For YFP-MinD experiments, ASC784 cells were grown in culture tubes with M9 medium at 37 °C. Filamentation was induced by 1 μM tetracycline and expression of YFP-MinD was induced by addition of IPTG (20μM final concentration) half an hour before imaging. 2 μL of culture was then imaged under the microscope between a glass slide and a cover slip. For minCDE null mutant experiments, cells were grown and filamented in M9 media with 2μM tetracycline, while divisions still occurred. For nucleoid labeling experiments, cells were grown and filamented in M9 media with 2μM tetracycline. For all time lapse experiments, phase contrast images were acquired at 1-2 minute intervals. Additionally, during the temperature experiment, fluorescent pictures were taken every 4 minutes, with an exposure time of 200ms. For nucleoid visualization, additional fluorescent images were taken every 5 minutes, with an exposure time of 25 ms. For the YFP-MinD experiments, only fluorescent images were taken at maximum acquisition rate with a 500 ms exposure time (i.e. with a rate of approximately 2 frames per second).

IMAGING AND IMAGE ANALYSIS

Cells were imaged with an inverted microscope (Nikon, TE2000), equipped with 100X oil immersion objective (Nikon, Plan Fluor NA 1.3), cooled CMOS camera (Hamamatsu, Orca Flash4.0), xenon lamp with liquid light guide (Sutter, Lambda LS), GFP, mCherry, CFP and YFP filter set (Chroma, 41017, 49008, 49001 and 49003), computer controlled shutters (Sutter, Lambda 10-3 with SmartShutter), automated stage (Märzhäuser, SCAN IM 120 x 100) and an incubation chamber (Solent) allowing precise 37 °C temperature control. An additional 1.5X lens was used, resulting in images with pixel size of 0.041 μm. The microscope was controlled by MetaMorph software. Series of phase contrast images were analyzed with a custom Matlab (Mathworks) program originally based on Schnitzcells software [16], which allows for automated segmentation of cells growing in a colony. The number of segmented and analyzed colonies was: 5 (1 μM tetracycline), 3 (2 μM tetracycline), 3 (10 μM tetracycline), 5 (SulA), 2 (temperature), 1 (nucleoid), 5 (Min deletion). In the Min deletion experiments, mini-cells were observed (inset Fig. 3.4.C)

but not segmented because of their abnormal size and dynamics. For all experiments, some cell cycles could not be monitored completely because the cells grew outside the field of view or because the experiment stopped, and were hence excluded from the analysis. To follow the cells over time, the images were manually corrected where necessary, and tracked to create a lineage branch-like structure. Each cell's length (polynomial fitted to a cell's curved segmentation region, or the segmented region's skeleton length for the temperature and sulA datasets) is computed for each frame. To determine relative division locations S, daughter cell lengths were divided by the maternal cell length (defined as the summed length of the two daughters). Growth rates were determined by fitting an exponential function to recorded cell lengths over multiple frames. For the experiments involving labeled FtsA or nucleoids, the fluorescent intensity along the cellular axis was determined using 1 pixel wide slices perpendicular to the morphologically computed skeleton. Peaks were then identified using the function peakfinder.m written by Nathanael C. Yoder; only peaks with a signal above a threshold level of 400 A.U. were considered (against an estimated background signal of approximately 50-100 A.U.). In Figure 3.3.A, the signal intensity outside cells was decreased for visualization purposes.

Simulations with the Meinhardt and De Boer model

We simulated the Min system behavior using the differential equations and rate constants described in Table S3.1, which were developed by Meinhardt and De Boer [96]. We extended the range of bacterial lengths simulated. This simulation numerically solves the differential equations, with stochastically fluctuating reaction constants, to calculate the protein numbers in each length element of a cell for multiple reaction species. We ran simulations for bacterial lengths of 15 to 100 a.u. for 106 iterations and recorded the system state every 1000 iterations. We then calculated the time-averaged MinD protein number profiles for each bacterial length.

Simulations with the KC Huang model using MesoRD

We simulated Min system behavior using the software MesoRD [98], "a tool for stochastic and deterministic simulation of chemical reactions and diffusion in 3D and planar 2D spaces". The model of the protein interactions is described in Table S3.2, and based on a stochastic adaptation [99] of the K.C. Huang model [97]. Ref. [99] also provides the reaction scheme in Systems Biology Markup Language (SBML). We used a diffusion constant of 8.2 µm2/s as measured from cytoplasmic diffusion of GFP, which has a mass similar to MinD (26.9 kDa and 29.4 kDa respectively) [106], the other parameters indicated in table S2 are taken from [99]. We ran 100-300 s simulations for a range of cellular sizes (with a compartment size of $5 \cdot 10^{-8}$, and recording the system state every second) and determined time-averaged concentration profiles.

3.3.4. QUANTIFICATION AND STATISTICAL ANALYSIS

The number of data points for each experiment can be found in the figure captions. To compare interdivision times, two sample t-tests were used (p = 0.05), as described in the main text and caption of Figure S3.6. Error bars always show SEM.

3.3.5. DATA AND SOFTWARE AVAILABILITY

Analysis and plotting was performed using custom Matlab scripts, which can be found at:
`https://github.com/TansLab/Tans_Schnitzcells`,
`https://github.com/TansLab/Tans_filamentation`, and
`https://github.com/TansLab/Common_libraries`.

3

3.4. Supplemental Data

Reaction scheme	Rate constants
Dynamics of membrane associated FtsZ: $$\frac{\partial F}{\partial t} = \rho_F f \frac{F^2 + \sigma_F}{1 + k_F F^2} - \mu_F F - \mu_{DF} DF + D_F \frac{\partial^2 F}{\partial x^2}$$ Dynamics of freely diffusible cytosol FtsZ: $$\frac{\partial f}{\partial t} = \sigma_f - \rho_F f \frac{F^2 + \sigma_F}{1 + k_F F^2} - \mu_f f + D_f \frac{\partial^2 f}{\partial x^2}$$	$\mu_F = \rho_F = 0.004$; $k_F = 0$; $\sigma_F = 0.1$; $\mu_{DF} = 0.002$; $D_F = 0.002$; $\sigma_f = 0.006$; $\mu_f = 0.002$; $D_f = 0.2$.
Dynamics of membrane associated MinD: $$\frac{\partial D}{\partial t} = \rho_D d \left(D^2 + \sigma_D \right) - \mu_D D - \mu_{DE} DE + D_D \frac{\partial^2 D}{\partial x^2}$$ Dynamics of freely diffusible cytosol MinD: $$\frac{\partial d}{\partial t} = \sigma_d - \rho_D d \left(D^2 + \sigma_D \right) - \mu_d d + D_d \frac{\partial^2 d}{\partial x^2}$$	$\mu_D = \rho_D = 0.002$; $\sigma_D = 0.05$; $\mu_{DE} = 0.0004$; $D_D = 0.02$; $\sigma_D = 0.0035$; $\mu_d = 0$; $D_d = 0.2$.
Dynamics of membrane association of MinE: $$\frac{\partial E}{\partial t} = \rho_E e \frac{D}{(1 + k_{DE} D^2)} \frac{(E^2 + \sigma_E)}{(1 + k_E E^2)} - \mu_E E + D_E \frac{\partial^2 E}{\partial x^2}$$ Dynamics of membrane associated MinD: $$\frac{\partial e}{\partial t} = \sigma_e - \rho_E e \frac{D}{(1 + k_{DE} D^2)} \frac{(E^2 + \sigma_E)}{(1 + k_E E^2)} - \mu_e e + D_e \frac{\partial^2 e}{\partial x^2}$$	$\mu_E = \rho_E = 0.0005$; $k_{DE} = 0.5$; $\sigma_E = 0.1$; $k_E = 0.02$; $D_E = 0.0004$; $\sigma_e = 0.002$; $\mu_e = 0.0002$; $D_E = 0.2$.

Table 3.1: Related to Figure 3.4. Reaction rules used in the Meinhardt and De Boer [96] simulations, see main text for a description.

Reaction scheme	Rate Constants
$MinD_{cyt}^{ATP} \xrightarrow{k_d} MinD_{mem}$	$k_d = 0.0125 \, \mu m^{-1} s^{-1}$
$MinD_{cyt}^{ATP} + MinD_{mem} \xrightarrow{k_{dD}} 2 MinD_{mem}$	$k_{dD} = 9 \cdot 10^6 \, M^{-1} s^{-1}$
$MinE + MinD_{mem} \xrightarrow{k_{de}} MinDE$	$k_{de} = 5.56 \cdot 10^7 \, M^{-1} s^{-1}$
$MinDE \xrightarrow{k_e} MinD_{cyt}^{ADP} + MinE$	$k_e = 0.7 \, s^{-1}$
$MinD_{cyt}^{ADP} \xrightarrow{k^{ADP \to ATP}} MinD_{cyt}^{ATP}$	$k^{ADP \to ATP} = 0.5 \, s^{-1}$

Table 3.2: Related to Figure 3.4. Reaction rules used in the Huang et al./Fange et al. [97, 99] stochastic MesoRD model of the Min system, see main text for a description.
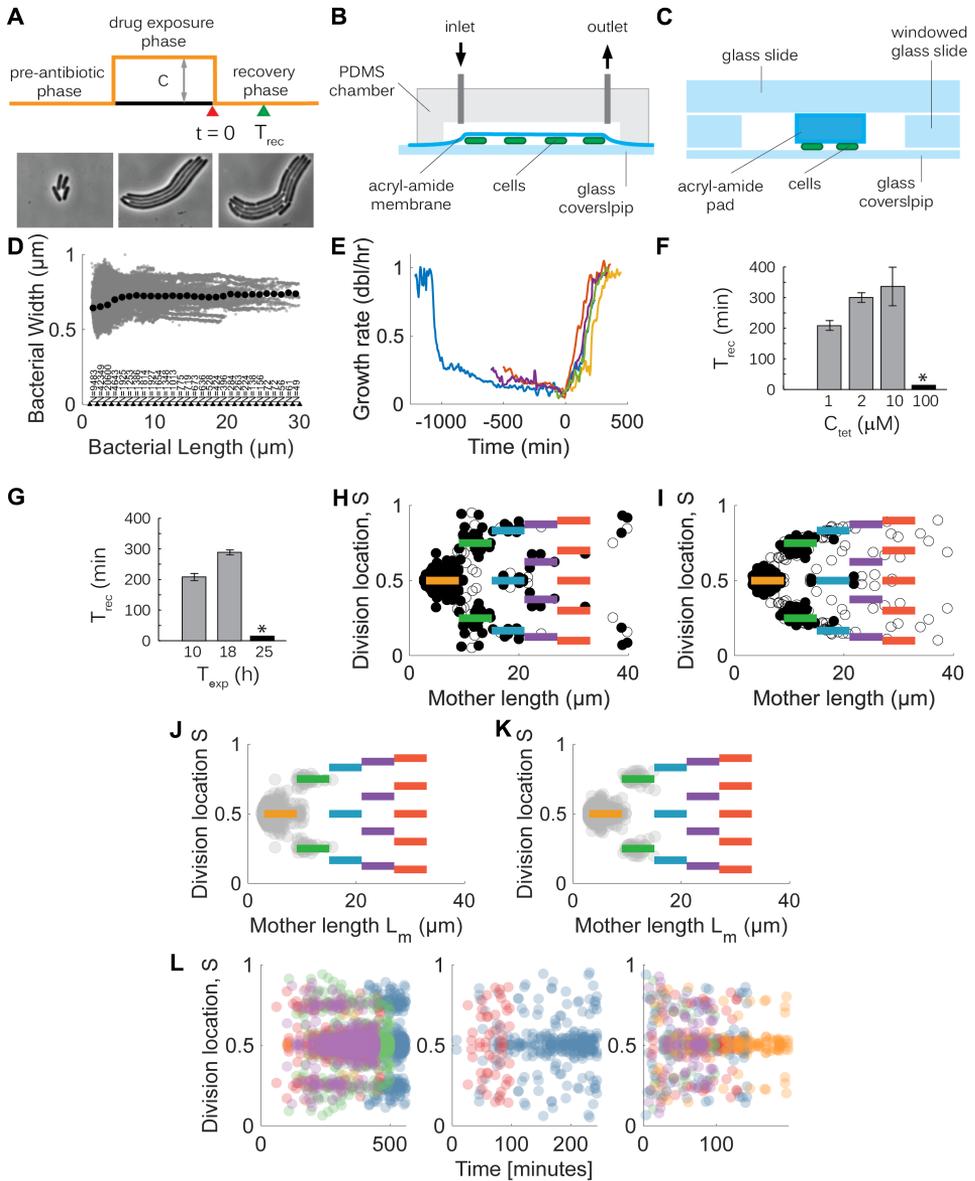
3

Figure 3.5: Division sites in filamentous *E. coli*. Related to Figure 3.1. (A) Three phases of the filamentation experiments: the pre-exposure phase during which cells grow normally; the exposure phase with filamentation-inducing stress; the recovery phase when stress is removed. Time $t = 0$ is the moment when stress is switched off; time T_{rec} is time when the growth rate recovers to the half of its initial pre-exposure value. (B) Schematics of the flow cell set-up. Medium is exchanged through inlets, and diffuses to cells through the acryl-amide membrane, which also holds the cells in place. (C) Schematics of the pad set-up, during which the medium is constant. The acryl-amide pad is soaked in the medium of interest and is kept in a windowed glass slide sandwiched between a normal glass slide and a glass coverslip. (D) Cellular width. Grey dots show bacterial width, determined as the cell area divided by the cell length, black dots the average width per length-bins. Red line is an estimate for the area divided by length, assuming a width of 0.74 μm and rounded caps of the rod-shaped bacteria, using that a bacterium of size L and width W has an area of $W \cdot L$ μm^2 minus $((0.74 \, \mu m)^2 - \pi (0.74 \, \mu m/2)^2)$. (E) Growth rate of cells before, during, and after tetracycline exposure. Growth rates were determined by fitting time evolution of single cell lengths to $L_0 2^{\mu t}$, and averaging over multiple cells in five datasets. (F) The time to recover to half the growth rate after removal of tetracycline for different tetracycline concentrations C_{tet} (see Fig. S1D). The star indicates that cells did not grow or filament at this tetracycline concentration. Error bars are SD. (G) The time to recover to half the growth rate after removal of tetracycline for different tetracycline exposure times T_{exp}. The star indicates that cells did show filamentous growth, but failed to recover from the exposure and/or lysed. Error bars are SD. (H) Filamentation induced by growing cells at 42C. When temperature was decreased to 37C, divisions resumed and relative division locations are displayed for each division. When the time since the last division is < 20 minutes, division events are marked by open circles. n=404 division events are shown. (I) Filamentation induced by overexpression of the division-inhibitory protein SulA. n=494 division events are shown. (J,K) Filamentation induced by 2 μM (J) or 10 μM (K) tetracycline. (L) Each dot shows the relative division location within a single cell against the time it took place. Different panels respectively correspond to the three conditions: recovery of filamentation after tetracycline exposure, 42C heat shock, or overexpression of division inhibitory protein SulA. Colors correspond to different datasets. (Respectively n=4108, n=404 and n=494 division events are shown.)
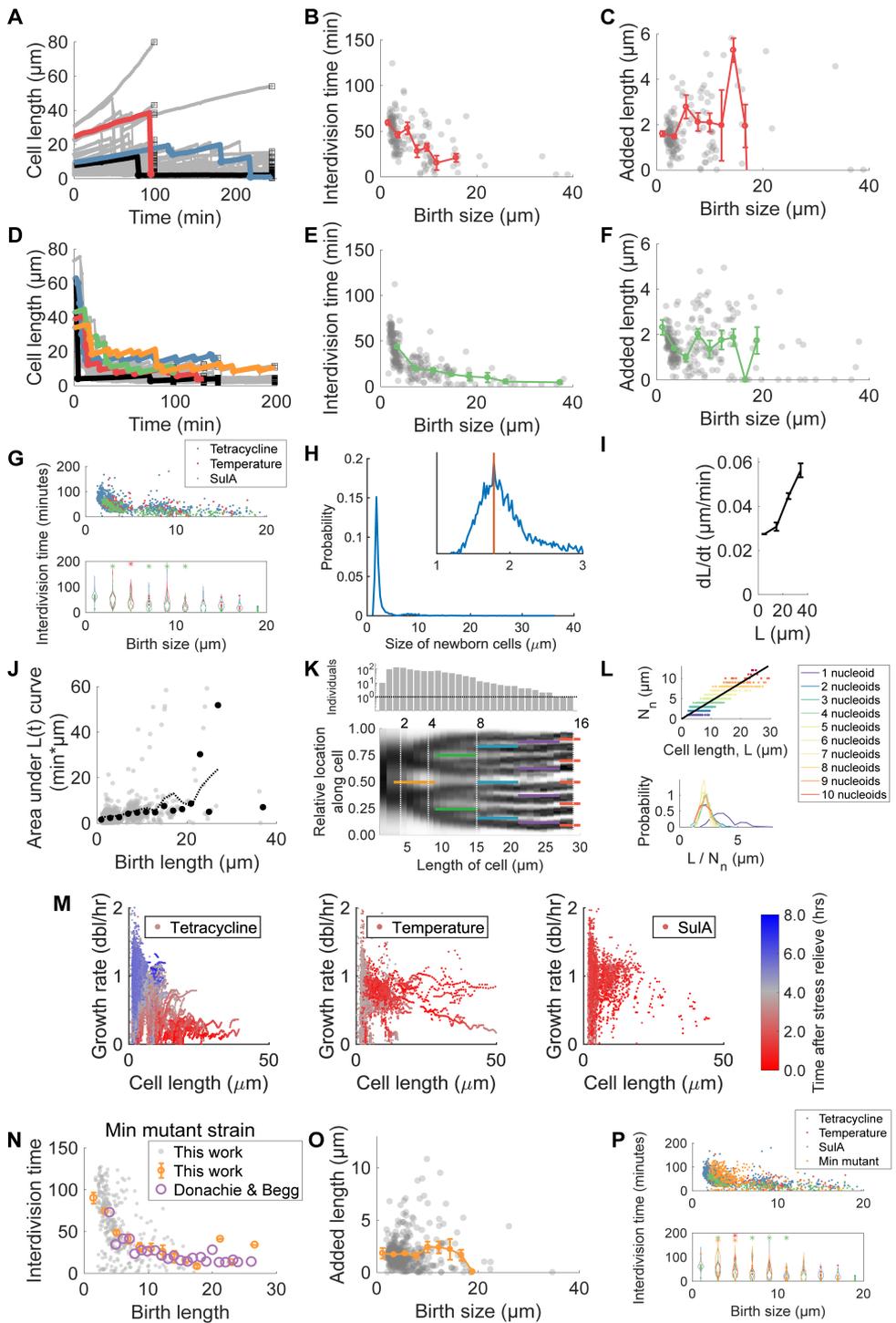
3

Figure 3.6: Timing, length changes, and nucleoid imaging. Related to Figure 3.2. (A,D) Single cell lengths versus time during recovery from temperature (A) or SulA (D). Grey dots represent the length of a single cell at a specific point in time. Colored lines trace example cell lineages, where drops in length correspond to divisions. Black squares indicate where measurements ended. (Respectively N=409 and N=499 cells were analyzed.) (B,E) Time between two subsequent divisions against birth size, for recovery from temperature (B) or SulA (E). The colored dots show average interdivision times (± SEM, respectively n=404 and n=494 division events). (C,F) Dots corresponds to a single cell, and shows how much length that cell added in between two divisions, versus birth size; again either for temperature (C) or SulA (F). Lines indicate averages (± SEM). (G) The top panel: data from Fig. 2B, Fig. S2B and S2E shown for comparison. Bottom panel: Corresponding Kernel Density Estimates per length bin. Stars indicate significant difference (two-sample t-test, p=0.05) compared to the tetracycline data set. For bins with n<11 points, single points are shown. (H) The distribution of newborn cell sizes observed in the tetracycline recovery experiment. The inset shows the distribution for the window 1-3 µm, with a smaller bin width. The red line indicates the maximum value at 1.78 µm (N=4108). (I) Derivative of cell length over time (dL/dt) for different lengths, as determined from every two consecutive frames, for in the tetracycline dataset, showing larger cells add more length per unit time. Error bars are SEM. (J) Grey dots: area A under the cell-length vs. time data, for the tetracycline dataset. Black dots: average values. These data agree with expectations for the adder behavior, in which A depends on added length ΔL and growth rate μ as $A = \Delta L/(\mu \cdot log(2))$ (see dotted line, using ΔL of 1.6 µm). (K) Nucleoid location. Black denotes signal from a HupA-mCherry fusion construct that labels the chromosome, plotted against cell size. Data is for growing filamenting cells. Fluorescence profiles along the cell axis were averaged within cell length bins of 1 µm. n = 5382 fluorescent profiles from N_c = 286 unique cells. Top panel: number of unique cells per bin. White dotted lines: 2^m multiples of cell birth length (1.8 µm), at which the number of nucleoids is expected to double if the concentration of nucleoids would be constant. (L) The number of nucleoids N_n (Matlab peakfinder algorithm) plotted against cell length L (n = 5382 observations). The ratio between these quantities remains approximately constant within the filamented regime, as shown by the line (which has a slope corresponding to 2.25µm/nucleoid). This is also reflected by the probability distributions $P_{N_n}(L/N_n)$, which show overlapping distributions of the cell lengths divided by the number of nucleoids, per number of nucleoids observed. (M) Growth rates (db/hr) vs. cell length, color coded for time after removal of stress. Panels indicate recovery from tetracycline, high temperature (42C), SulA overexpression, respectively. Number of cells: N=4134, N=409 and N=499. Number of data points n=129722, n=11049 and n=7304, respectively. (N) Interdivision time vs. birth length for *minCDE* null strains. Orange dots are averages, bars are SEM. Purple points: data from 1996 Donachie and Begg paper [86], showing data sets are consistent. (O) Added size between divisions for *minCDE* null strains. Orange dots are averages, bars are SEM. (P) As panel G, but in addition, now also data from the *minCDE* null strains is shown (orange).
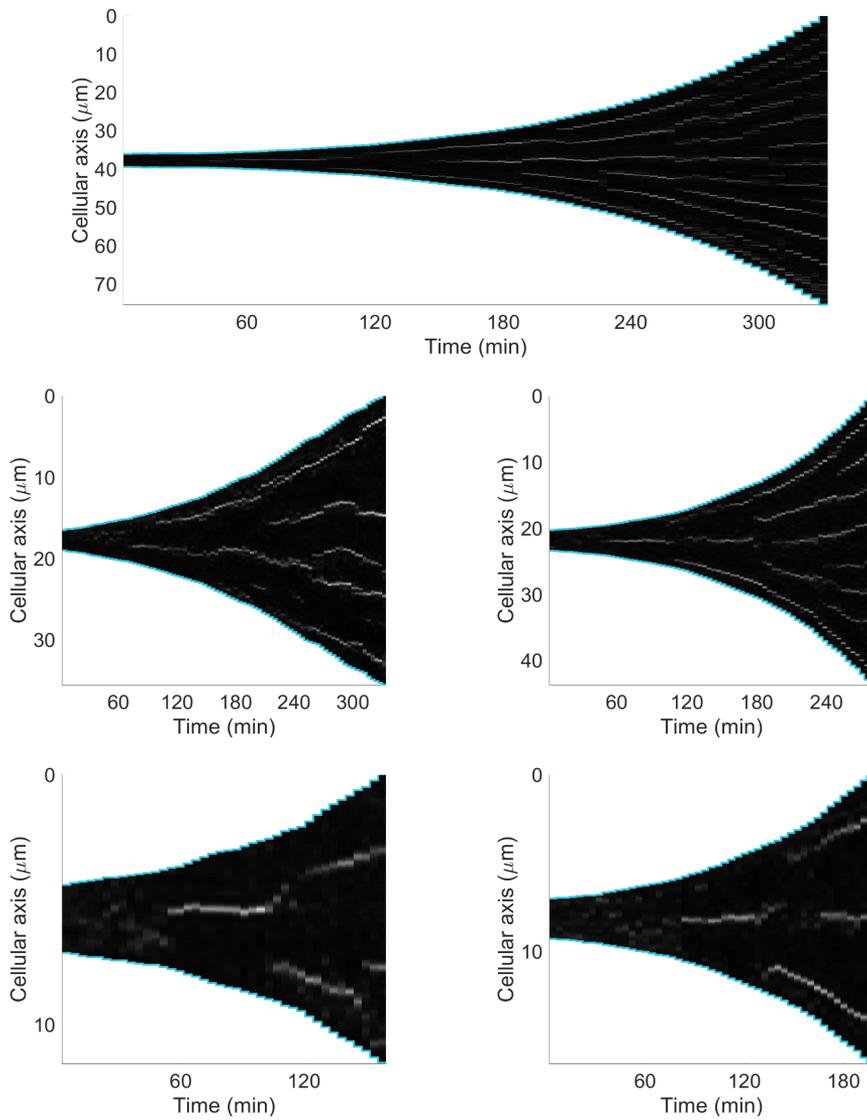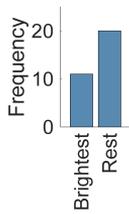
Figure 3.7: Division ring imaging. Related to Figure 3.3. (A) Fts ring localization in cells fila-
mentous at 42C. Kymographs of sYFP2-FtsA intensity along the cellular axis over time. Each panel
corresponds to an individual cell in between two divisions. The top three panels show cells grow-
ing during exposure to 42C, the bottom two panels show cell born small after temperature was
lowered to 37C. (B) Intensity of sYFP2-FtsA ring and probability of division at this ring. About 35%
of observed divisions occurred at the brightest FtsA ring, showing no clear preference for dividing
at the brightest ring.
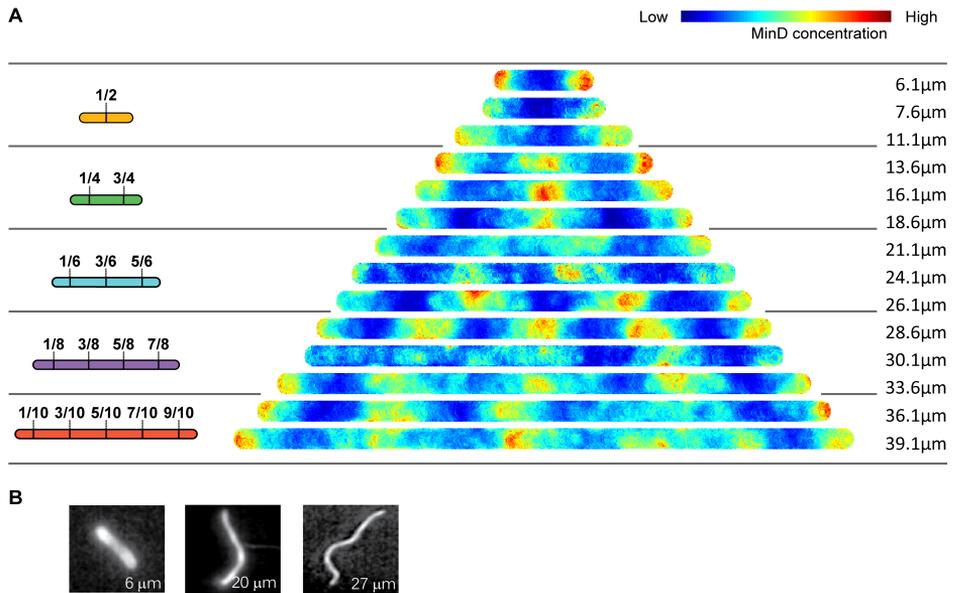
Figure 3.8: Min patterns. Related to Figure 3.4. (A) Results of stochastic simulations using the Huang et al. model implemented by Fange et al. [97, 99]. Heat maps show the membrane concentration of MinD for cells of different lengths. Data is consistent with division rules (left). (B) YFP-MinD fluorescence for cells filamented in the presence of 1 μM of tetracycline. Three cells of different lengths representing different division windows are shown. Cells of different lengths are shown, corresponding to division regimes with respectively 1, 3 and 4 division sites.

4

STOCHASTICITY IN CELLULAR METABOLISM AND GROWTH: APPROACHES AND CONSEQUENCES

HIGHLIGHTS

- Single cell data shows that metabolism undergoes stochastic fluctuations.
- They produce fluctuations in cellular growth, gene expression, and other phenotypes.
- They can be detrimental to populations, but also beneficial.
- They raise new questions on the interplay between different enzymes and fluxes.

ABSTRACT

Advances in our ability to zoom in on single cells has revealed striking heterogeneity within isogenic populations. Attention has so far focussed predominantly on underlying stochastic variability in regulatory pathways and downstream differentiation events. In contrast, the role of stochasticity in metabolic processes and networks has long remained unaddressed. Here we review recent studies that have begun to overcome key technical challenges in addressing this issue. First findings have already demonstrated that metabolic networks are stochastic in nature, and highlight the plethora of cellular processes that are critically affected by it.

4.1. STOCHASTICITY AND METABOLISM

Elucidating the role of molecular stochasticity in metabolic processes is a central issue in cellular physiology. It is key to understanding cellular homeostasis, and could help explaining heterogeneous phenotypes ubiquitously observed across all domains of life, ranging from persistence to cancer [108, 109]. Stochasticity in metabolism could underlie bet-hedging strategies, in which distinct sub-populations anticipate future environmental change [110, 111]. On the other hand, metabolic stochasticity could limit optimal growth and require regulatory mechanisms to ensure homeostasis [112]. More generally, as metabolism ultimately drives all cellular processes, fluctuations and instability could impact a myriad of phenomena ranging from the cell cycle to differentiation events. So far however, stochastic variability is commonly considered to have negligible effects in metabolic networks, as reflected by current theoretical models [113]. Indeed, metabolic fluctuations may be insignificant because of averaging over the many reaction events underlying metabolism in cells, chemical equilibration, metabolite secretion, or a lack of limiting steps within metabolic pathways [113–120].

At the practical level, quantifying any type of metabolic fluctuations comes with its own specific challenges. In contrast to regulatory proteins within signalling networks, which can be tagged fluorescently, metabolites are difficult to visualize at the single-cell level. Metabolites can be quantified by single-cell mass spectrometry [121], but so far not dynamically in time. Spectroscopic methods can follow metabolite abundance in time, but only for specific highly abundant molecules such as lipids [122]. FRET and fluorescent sensors hold a lot of promise, but remain limited to some metabolites and cannot yet quantify stochastic fluctuations [123–127].

Recently, important progress has been made in developing novel approaches that circumvent these limitations. In this review, we will examine these new efforts, their first findings, as well as related theoretical modelling. We will also cover recent work that is addressing the impact metabolic variability has on other cellular phenomena.

4.2. ENZYME EXPRESSION GENERATES METABOLIC NOISE

Early single-cell experiments showed how the expression of transcription factors fluctuate and propagate to downstream genes [15, 128, 129]. Similarly, such expression noise in key metabolic enzymes could generate variations in the flux of the reaction they catalyse, even if reaction-event noise averages out [130]. Moreover, if these flux variations propagate down-stream along the pathway, they could produce variations in the rate of cellular growth. A recent study by Kiviet et al [10] was based on this premise. While such an approach presents the challenge of quantifying enzyme expression and cellular growth with high accuracy, it avoids the need to measure fluctuations in metabolite concentrations.

Growth was quantified by following the size of individual cells by time-lapse microscopy. Specifically, using the known overall shape of E. coli - a rod capped with half-domes - its length could be determined to below the diffraction limit, which may be compared to how fluorophores are positioned in super-resolution microscopy [10]. Currently, a range of different single-cell image analysis approaches are available [36–43], including ones utilizing machine learning [44, 131, 132]. Cellular growth has also been

quantified by measuring cellular dry mass [133], and by using AFM-like cantilevers [134], as will be discussed more exhaustively below.

The data on the instantaneous cellular growth rate appeared correlated with the expression of metabolic enzymes [10]. However, such correlations could signal that growth fluctuations perturb expression, rather than the other way around. Time dependent correlation analysis can be used to address this issue [15, 128] (Fig. 4.1). This approach showed that the correlations were on average stronger after a certain delay, consistent with enzyme production fluctuations happening first, and growth fluctuations happening some time later (Fig 1a). In line with the idea that enzyme (expression) fluctuations affect the flux of the reaction they catalyze, this delay was observed only for genes that were considered limiting, such as gltA and icd in acetate media, and pfkA and icd in lactose media.

Interestingly, even when considering non-limiting genes, the expression rate was still strongly correlated with growth – however the correlations were now instantaneous and did not show a delay (Fig. 4.1b). It suggested that more generally, proteins are expressed significantly faster in cells that transiently grow faster, which is actually not unreasonable given that some cells grow twice as fast others for almost a full generation, and expression needs diverse metabolites. Put differently, fluctuations in growth-controlling factors, which may be anything from ribosomes to ATP, are also a source of gene expression noise [135]. In turn, metabolic fluctuations may thus affect processes that are controlled by gene expression, such as differentiation events [136, 137]. Metabolic noise can be compared to other noise sources such as transcription factors [9] and the cell cycle [138], which can also affect more than one gene or process and hence may be considered as extrinsic noise sources. A picture thus emerges of a system as a cycle of reciprocally interacting sources of extrinsic noise: metabolic fluctuations simultaneously affecting the expression of multiple genes, including transcription factors, polymerases, and metabolic enzymes, and conversely, noise in the latter resulting in fluctuations in metabolic fluxes. At the same time, the precise relations between noisy signals, and hence their ultimate mechanistic origin remains largely unresolved. For instance, it is unclear whether different pathways fluctuate independently, or alternatively, whether observed fluctuations result from a continuous dynamic interplay between them. Overall, the data so far shows that expression and growth are tightly intertwined, not only in terms of their mean levels when comparing different media [139], but also dynamically within constant external conditions.

4.3. (MIS)MATCHING PATHWAYS

The notion that metabolic pathways are stochastic raises questions about the dynamic interaction between them. For instance, it is thought that cells co-regulate functionally related genes to balance their overall input and output fluxes [130, 140]. In yeast, genes related to either stress response, mitochondria or amino acid biosynthesis were found to fluctuate jointly in response to general regulators [141]. Mismatches between (parts) of the cellular pathways can have large effects. Specifically, it was observed that metabolic imbalance within glycolysis can amplify non-genetic variability within the population [142]. When the upper and lower parts of this central pathway are not well matched, gly-

4
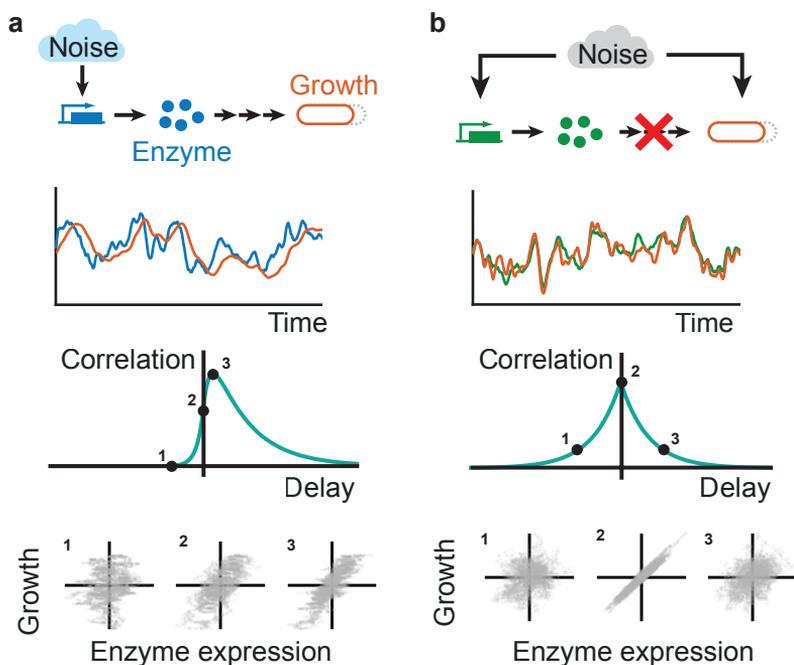
Figure 4.1: Fluctuations from enzyme expression to metabolism, and from metabolism to enzyme expression. Expression measurements of a single metabolic enzyme and growth rates in individual cells can be used to reveal metabolic stochasticity. Two key modes of noise transmission have been observed, which can act both individually and jointly, and may interact. (a) Noise in the expression of a single enzyme (blue trace), result in fluctuations in metabolic flux that are transmitted through the metabolic network and affect growth with some time delay (orange trace). The delay can be quantified by cross-correlation analysis. The cross-correlation curve illustrates that on average, current enzyme expression correlates better with growth some time later, as illustrated by the expression-growth scatter plots. Note that the sources of expression noise here are not only intrinsic, or caused by molecular processes specific to one gene. They also include extrinsic or transmitted noise from other processes, such as transcription factor, polymerase, or metabolic factors such as amino acid abundance, which may affect expression but not growth. Noise sources that affect both expression and growth are discussed in panel b. (b) Noise sources within the metabolic network that perturb both expression (green trace) and growth (orange trace). Fluctuations in components that affect both expression and growth, such as ATP and other central metabolites, could define such sources of noise. In contrast to panel a, the cross-correlation here is symmetric because expression and growth respond approximately equally fast to the fluctuations. Note that the resulting expression noise may affect growth (panel a), or may not (this panel) - for instance because the expressed enzyme is not metabolically active or because it is abundant and hence does not limit growth.

colytic intermediates can accumulate while ATP levels are reduced, thus strongly affecting cellular physiology. Expression variability has also been suggested to drive changes in flux partitioning [143]. These studies underscore the importance of further dissecting how cells coordinate different cellular processes in the face of the random fluctuations of its components, and which regulatory mechanisms they employ.

4.4. METABOLISM AT THE CENTER

Metabolism and growth ultimately power all cellular activity. A fluctuating or unstable metabolism thus could have wide-ranging effects. For instance, perturbations of metabolic homeostasis may cause fluxes to collapse and metabolite pools to deplete, which in turn can induce persistence [109]. Metabolic heterogeneity has been suggested to affect the synchronization of metabolic oscillations observed in dense yeast populations, and hence the communication between cells [144], while a recent study revealed a coupling between metabolic oscillations and the cell cycle in yeast [145]. Strikingly, it has recently been reported that slow-growing yeast sub populations display downregulated ribosomal activity and upregulated stress response genes, increased RNA polymerase error rates and indications of DNA damage, which may be explained by oxidative stress [146].

One may also expect that metabolic and growth fluctuations impact cell size. Bacteria grow in exponential fashion - increases in growth rate could thus produce large increases in cell size, which could be further amplified and diverge in subsequent cycles because larger cells effectively grow faster. Some answers to how cells deal with this issue are already emerging. First, the timescale of growth fluctuations in E. coli was found to be just below that of the cell cycle for a range of growth media [10]. Cells thus inherit faster growth for just one or two generations, which limits amplifying effects. Second, while the molecular mechanism is unclear, it has been found that cells compensate for growth variability [33, 57, 58, 147–150]. Cells that grow faster on average have a smaller interdivision time, thus yielding similar sizes at division as slow-growing cells (Fig. 4.2a). Moreover, faster-growing cells were also found to initiate DNA replication earlier, providing a further indication of underlying regulatory compensations [57, 147]. These findings support the suggestion that the cells compensate for growth variability by measuring size rather than time.

4.5. BENEFITS OF METABOLIC FLUCTUATIONS

Stochasticity of growth and expression is directly observed within individual cells, but it can also affect the composition of the population in non-trivial ways. This issue has been studied theoretically and in experiments [27, 151, 152]. Counter-intuitively, analysis showed that growth rate distribution along a single a lineage is not necessarily equal to the distribution within the population at a single time point [27]. The cause however is actually quite simple: faster growing phenotypes produce more offspring, and hence become overrepresented within the population (Fig. 4.2b). The effects are most striking when the mean concentration of a growth-controlling enzyme is suboptimal, as gene expression noise and resulting growth noise can then increase the growth rate of
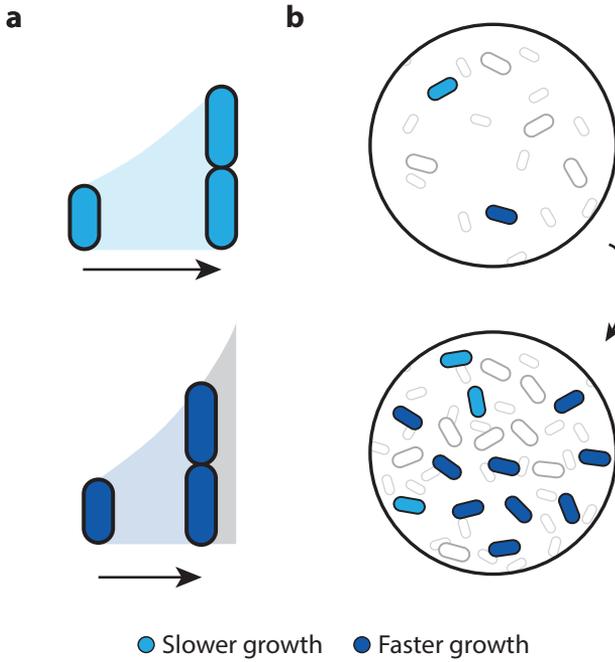
Figure 4.2: Impact on cell cycle and population structure. (a) Cell cycle compensations. Recent work has shown that spontaneously faster growing cells initiate DNA replication earlier, and divide earlier, than slower-growing cells in the population. Such compensations limit the effects of heterogeneity in growth rate on cell size. (b) Effects on population structure. Faster-growing and faster-dividing cells increase their frequency within the population. As a result, growth noise can result in population growth rates that are higher than the average cellular growth rate within a lineage.

the population as a whole [151]. Such sub-optimal regulation of enzyme expression has been observed experimentally (e.g. [153]), and in one direct study, population growth rates were found to be almost 10% faster than the average single-cell growth rate [27]. A similar study in yeast showed a 4-7% increase in growth rate for the population as a whole [152]. Additionally, an artificial reduction of gene expression noise in catabolic networks decreased heterogeneity in cellular division times [152], consistent with noise in metabolic enzymes controlling growth [10]. The advantage of fluctuating gene expression in variable environments was studied earlier in a synthetic system, in which bistable switching allowed cells to be prepared for environmental change [154]. The idea of "stochastic sensing" has been addressed theoretically [110] and observed in metabolic networks [32, 155, 156]. It has been proposed that the regulatory control of metabolic genes constrains the space of possible random metabolic phenotypes, and hence come with entropic energy costs [157]. Overall, noise in metabolic systems thus may not exclusively limit optimal growth, but can also be beneficial. This point is further illustrated by observed evolutionary adaptation towards more heterogeneous phenotypes [146, 158–161].

4.6. AN EXPANDING ARRAY OF EXPERIMENTAL APPROACHES

Tracking cell size and fluorescence has already led to surprising insights to the dynamics of cellular physiology. Novel approaches will open up additional possibilities. Fluorescence methods have been used to detect the synthesis of single proteins in eukaryotic cells [162]. The growth rates of eukaryotic cells are difficult to measure using time-lapse microscopy, given their complex three-dimensional shapes. A recent technique overcomes this problem, by quantifying how the cell volume reduces the abundance of fluophores in the surrounding medium [163]. The accuracy of gene expression measurements is also improving. Single proteins could be visualized in E. coli cells by slowing down their diffusion [164]. Measuring metabolite concentrations would allow direct access to fluxes. Concentrations of FAD and NADH can be measured using auto fluorescence [144, 165], while FRET sensors have already been developed for calcium [123, 124], ATP [125, 126] and cAMP [127]. Additionally, it is possible to obtain single cell Raman spectra, which allow for determination of concentrations of certain abundant metabolites [122]. Together, these novel and existing approaches will be central to arrive at a dynamic view of physiology at the single-cell level.

4.7. CONCLUDING REMARKS

In this review, we have discussed recent studies that have revealed the stochastic nature of metabolism and its interplay with gene expression and other cellular processes. The results press the notion of cells as autocatalytic and stochastic systems engaged in a dynamic equilibrium, with metabolism and enzyme expression as two fluctuating and interdependent processes. One may expect other processes to be in similar dynamic equilibria, and it will be intriguing to decipher how the result can be stable and robust. In recent decades, growth has not been considered as an important piece of the cellular puzzle. This new wave of experiments is revising this view, and re-affirms metabolism and growth at the center of cellular activity and dynamics.

CITATIONS OF OUTSTANDING INTEREST

- Kiviet et al. (2014) [10]: Stochastic fluctuations in the concentration of a single enzyme can correlate with future cellular growth rates, indicating transmission of noise through cellular networks. Fluctuations in growth rate can also affect enzyme expression, showing the interdependence of these parameters.

- Hashimoto et al. (2016) [27]: Experimentally measured discrepancies between single cell growth rate distributions and population growth rate distributions can be understood theoretically, and show that single cell growth noise can be beneficial to population growth.

- Cerulus et al. (2016) [152]: Experiments show that population growth can benefit from growth noise on the single cell level, and stochastic catabolic gene expression can contribute to single cell growth noise.

CITATIONS OF INTEREST

- Van Valen et al. (2016) [44]: Machine learning algorithms are applied to segment time-lapse movies of single cells.

- Walker et al. (2016) [138]: Volume increase and gene duplication during replication during the cell cycle both affect gene expression, leading to non-genetic cellular heterogeneity.

- Van Dijk et al. (2015) [146]: Sub-populations of slow-growing cells are analyzed and found to have distinct phenotypes that also have higher mutation rates.

- Adiciptaningrum et al. (2015) [147]: Replication and division timing compensate for both growth rate variability and cell size variability in single cells.

4

- Towbin et al. (2017) [153]: Metabolic enzyme expression is optimized for most conditions, but not all.

- Schreiber et al. (2016) [158]: Phenotypic heterogeneity is increased in response to nitrogen limitation, and benefit thereof is experimentally shown.

- Morisaki et al. (2016) [162]: A smart combination of fluorescent techniques allows probing of single-molecule translation events and mRNA translation kinetics.

4.8. ACKNOWLEDGEMENTS

5

CRP RESPONDS DYNAMICALLY TO INTERNAL NOISE

5.1. INTRODUCTION

5.1.1. ENVIRONMENTAL AND STOCHASTIC INPUT TO REGULATORY NETWORKS

The world is unpredictable. A bacterial cell observes both its extracellular environment and its intracellular environment, and adjusts its gene expression to optimize its chances of survival in this world. To control gene expression a bacterium — like any other cell — relies on chemical interactions between proteins and small molecules [166–169]. These interactions, also referred to as "biochemical networks", allow cells to deal with unpredictable situations both in the intracellular and extracellular environment. Changes in the extracellular environment might require the cell to adjust gene expression accordingly. In other words, the cell responds to "environmental inputs". On the other hand, even without environmental changes, the stochastic nature of chemical reactions can lead to fluctuations of proteins and metabolites over time within the cell [9, 10]. Such fluctuations might present regulatory networks with "stochastic inputs". The architecture of the biochemical network might enable the cell to deal with these fluctuations, or even use them to its advantage (see also chapter 4). To what extent networks respond to stochastic inputs, and what the implications of dealing with such inputs are for network architecture is a largely unaddressed question. In this work, we investigate the link between environmental and stochastic regulation for the first time in a native cellular control circuit. Specifically, we ask: are regulatory interactions that are known for their role in adaptation to environmental changes also activated by concentration fluctuations in the intracellular environment due to noise?

5.1.2. STOCHASTIC AND ENVIRONMENTAL INPUT IN A MODEL SYSTEM: CRP

To investigate this question, we looked at the regulation of metabolic enzyme expression by the cAMP receptor protein (CRP). Activated CRP stimulates metabolic enzyme

expression, but CRP activation is inhibited by the metabolites that are produced by metabolic enzymes [170]. It was recently discovered that when cells are grown in environments with different sugar sources, this sensing of the metabolite concentrations by negative feedback is the regulatory interaction which allows cells to express the appropriate amount of metabolic enzymes for most sugar sources [153]. This can involve large changes in proteome composition [140, 170]. However, when a population of cells is growing in a constant environment with a constant carbon source, there is still a large heterogeneity observed in growth rates and metabolic enzyme expression levels [10] (see also chapter 4). Observations by Kiviet et al. suggest that single cell enzyme concentrations fluctuate stochastically over time, which in turn results in fluctuations in metabolite fluxes and eventually in growth rate [10]. Thus, single cell metabolite concentrations might fluctuate over time due to stochastic fluctuations . That would imply that also the CRP system receives a large variation of inputs in single cells, even when there are no changes in the extracellular environment. We here set out to investigate whether the CRP system indeed experiences and acts on these different stochastic inputs.

5.1.3. THE CRP SYSTEM

Before turning to how we manipulate and measure the dynamics of the CRP system, we would also like to provide some context on this considerably researched system [171–174]. CRP might be called a master regulator of metabolic enzyme expression in *Escherichia coli*[1], since it is thought to control the expression level of all catabolic genes in concert [140, 170, 175]. CRP controls 378 promoters, among which 70 transcription factors [176–179]. Amidst these targets are all enzymes in the TCA cycle, see also figure 5.12. Such catabolic enzymes are of major importance to the cell, as they convert large carbohydrate molecules into smaller metabolites, and generate energy for the cell during this process in the form of ATP [180]. The concerted regulation of catabolism related genes leads to large scale cellular changes, as illustrated by experimental observations. One study showed that a CRP reporter (based on the lactose gene) showed a tenfold change in activity in response to different growth media [170]. A second study showed that in response to artificial limitations on cellular carbon import, cells adjusted the fraction of the proteome dedicated to catabolism from 10% up to to 25% [140].

CRP is allosterically activated by the small signaling molecule cyclic adenosine monophosphate (cAMP), which is produced from ATP by the enzyme adenylate cyclase (CyaA). The cell can also control cAMP concentration by active degradation, which is catalyzed by the cAMP phosphodiesterase (CpdA) enzyme[2]. cAMP production by adenylate cyclase is thought to be inhibited by α-ketoacids, such as oxaloacetate (OAA), α-ketoglutarate (α-KG) and pyruvate (PYR) [170]. (Previously, also the phosphorylated enzyme IIA of the phosphotransferase system was thought to activate adenylate cyclase [171, 173, 174], but it has now been suggested that the role of α-ketoacids feedback is bigger [170].) Thus, the CRP-cAMP regulation is wired such that metabolites from the TCA cy-

---

[1]CRP was previously called catabolite gene activator protein, or CAP.

[2]It is sometimes mentioned in literature that cAMP is also actively exported from the cell by the membrane channel protein TolC, but the primary publication that made this claim has been retracted [181].

cle and glycolysis pathway provide negative feedback on metabolic enzyme expression, see also figure 5.1.A. The effect of CRP regulation can be quantified by the relationship between the cAMP concentration and cellular growth rate, which was done for different sugar sources [153]. For each sugar, there appeared to be a specific optimal concentration of cAMP, consistent with the idea that too low enzyme concentrations limit metabolite fluxes, and too high concentrations draw resources from other cellular processes [153, 170, 182–185]. This leads to a concentration-growth relationship with a clear optimum, which we will therefor call the optimum curve, see also figure 5.1.B. When cells use different sugars as carbon source, they respond by adjusting the concentration of metabolic enzymes [170]. As mentioned, it is thought that the CRP system is responsible for finding the optimal concentration of enzymes for most sugar sources, which is thought to be achieved by the negative feedback from metabolites [153].

5.1.4. CRP WITH AND WITHOUT FEEDBACK

To find out whether the CRP system responds to stochastic input we decoupled the responses to stochastic and environmental inputs. To achieve this, we employed two strains that had the appropriate mean CRP activity for the sugar source they were grown on, but where in one strain the feedback loop could not respond to stochastic input it might receive. The two strains we used are a wild type MG1655, and a *cyaA, cpda* null mutant [153] (see also table 5.1 in the methods section). As mentioned, *cyaA* enables cAMP production and *cpda* cAMP degradation, which means that the *cyaA, cpda* null mutant was unable to modify its cAMP levels, and had a crippled feedback loop that could not respond to both environmental and stochastic input signals. We therefore call this strain the ΔcAMP strain. To repair this strain's ability to express the right amount of metabolic enzymes with regard to their environment, we provided cAMP extracellularly in the cells' growth medium. This bypassed the feedback loop and directly set the correct CRP activity for the sugar they were grown on [153], but left the feedback loop unable to respond to stochastic input signals, see also figure 5.1.C. This experimental design allowed us to compare a cell that has the correct mean CRP activity and *could* respond to stochastic input signals, with a cell that had the correct mean CRP activity but *could not* respond to stochastic input signals. We then used single cell time lapse microscopy and fluorescent labeling to determine single cell growth and CRP dynamics. We first used these single cell techniques to see whether the CRP system could respond on timescales at which the stochastic fluctuations take place. Secondly, we used cross-correlation analyses to quantify the dynamical growth and CRP behavior of the two strains described above. These analyses showed that the CRP system can indeed operate on timescales that are comparable to time scales of stochastic fluctuations, and moreover that the dynamics of metabolism and growth change remarkably when its regulation is not able to respond to stochastic inputs. These observations suggested that regulatory interactions also respond to stochastic inputs. This implied in turn that perhaps our notion of steady state behavior should be updated, as the cellular state might be constantly changing in a way that is facilitated by regulatory interactions.
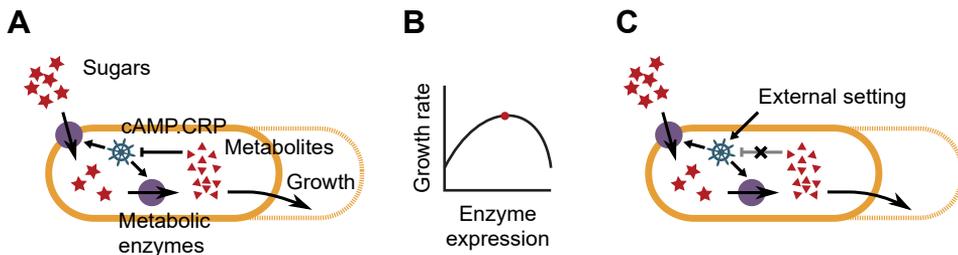
5

Figure 5.1: (A) This diagram depicts the regulation of metabolic enzyme expression in *E. coli.* Metabolic proteins import sugars and convert these large carbohydrate molecules to smaller metabolites, a process in which the cellular energy source ATP is produced. The metabolites are also a cellular resource and are building blocks for other cellular components, and are thus needed for cellular growth. They also inhibit cAMP production, thereby inhibiting CRP activation by cAMP. CRP is a master regulator (hence represented by a small ship's wheel): the concentration of activated CRP sets the metabolic enzyme concentration. (B) This cartoon illustrates the general relationship between enzyme expression (which for metabolic enzymes is controlled by cAMP) and growth rate. In a constant environment, too little expression will limit metabolic fluxes, and too much expression draws cellular resources from other cellular processes. Therefore, there is an optimal concentration at which the growth rate is highest, as indicated by the red dot. (C) In our *cyaA, cpda* null mutant, the cell is unable to modify cAMP concentration and metabolites are thus unable to set the CRP activity (indicated by the black cross) nor metabolic enzyme expression. Instead, we supply cAMP in the medium such that we can artificially control metabolic enzyme expression. This external setting of the cAMP concentration makes that these mutants are unable to respond to intracellular fluctuations of metabolites.

5.2. Results

5.2.1. The CRP system is able to respond to dynamic signals

Our hypothesis is that stochastic fluctuations in metabolic enzyme concentration result in metabolite concentration changes, which in turn lead to a response by the CRP system. However, the response to such stochastic fluctuations — which can occur on timescales shorter than the cellular division rate — might be limited because the fraction of the proteome regulated by CRP is so large and because the cell only has a limited capacity to produce proteins or to reduce their concentration. Therefore, we first wanted to see whether the CRP system was at all capable of responding to a quickly changing input signal, such as stochastic fluctuations might generate. In this section we show that the CRP system indeed responds considerably at sub cell cycle timescales to changes in the cAMP signal it receives.

To probe the response of the CRP system to a quickly changing signal, we used our ΔcAMP strain that does not respond to intracellular metabolic feedback. Instead of the intracellular signal, we provided it with an artificial cAMP signal by providing high (2100 μm) or low (43 μm) concentrations of cAMP in the cellular growth medium. These concentrations were chosen to lay respectively above and below the optimal cAMP value, which we determined at 800 μm cAMP (supplementary figure 5.13). Additionally, to assess the response of the CRP system we equipped these cells with a chromosomally inserted reporter construct that reads out the expression level of metabolic enzymes [153]. This reporter construct uses a lac operon promoter from which the lacI binding site has been removed (see supplementary figure 5.14). The promoter is fused to an mVenus fluorescent reporter protein which allowed us to assess the expression level of metabolic enzymes as induced by CRP. We therefor call this the metabolic reporter. We also introduced a second reporter to our cells: a constitutive promoter fused to an mCerulean fluorescent protein, see also figure 5.14. This will allow us to compare protein expression regulated by CRP with protein expression not regulated by CRP. We will refer to this reporter as the constitutive reporter.

We grew the ΔcAMP cells in our microfluidic device (see methods section and chapter 2) under the microscope, and first alternated between the low and high concentrations of cAMP at 1 hour intervals ("fast pulses"), and then alternated between the same concentrations at 5 hour intervals ("slow pulses"). The metabolic reporter responded as expected, as its concentration went up when the concentration of cAMP was high, and down when the cAMP concentration was low. During the fast pulses, the range of reporter concentration was narrower (with population averages between 215 a.u. and 258 a.u) than during the slow pulses (where population average concentrations ranged between 144 a.u. and 407 a.u.), indicating cellular metabolic concentrations could not reach an equilibrium within a one hour-long pulse. Our data also allowed us to calculate protein production rates, which we based on the slope (determined from a linear fit) of the total fluorescence of the metabolic reporter at time points t_{i-1}, t_i and t_{i+1} (i.e. the value at time point t and the previous and subsequent values), divided by the cellular area at t_i (see methods section). Consistent with the observations on the concentrations, we saw that also the production rate of the metabolic reporter went up and down with the cAMP concentration (supplementary figure 5.15.A). Additionally, we saw

that the constitutive reporter concentration went down as the CRP reporter went up, and vice versa (supplementary figure 5.15.B-C). This observation is consistent with the idea that when certain large groups of genes are up-regulated, other genes are expressed less [170]. In general, these observations confirm that we can control the CRP activity by controlling the external cAMP concentration and that the reporter adequately reflects this in our dynamical experiment. Importantly, in the fast regime, we see that the reporter signal responds at time scales that are below the cell cycle time, which ranges between approximately 1-2 hours.

This led to our next question: what is the effect of these changes in metabolic enzyme concentration on the cell? To gauge this, we looked at how cellular growth rate responded to the fluctuations in cAMP signal. Figure 5.2.B shows that also the growth rate varied substantially, again suggesting the cell is susceptible to fluctuations in the cAMP regulatory molecule. In the regime where the cAMP concentration changes hourly, we see that low concentrations of cAMP correspond with lower growth rates, and that high concentrations of cAMP correspond with higher growth rates (figure 5.2.B, supplementary figure 5.16.B). This indicates that during the fast pulses, the range of metabolic enzyme concentrations remained below the optimal value. Additionally, during the fast pulses, the data points right before and after the time of switching cAMP concentration show large changes in growth rate, whilst there is only 20 minutes in between two datapoint and CRP controls a substantial fraction of genes in the proteome. This indicates that the metabolic fluxes (and therefor growth rates) might be very sensitive to the metabolic enzyme concentration in this range of enzyme concentrations.

During the first slow pulse (with a low concentration of cAMP), the growth rate appears to decrease almost monotonically. However, both during the second and third pulse (with a high and low concentration of cAMP respectively), the growth rate first rises, and then decrease again. This can also be understood in the context of the optimum curve (figure 5.1.B), as this probably indicates the concentration of enzymes now reaches values both below and above the optimum.

To further investigate the behavior of the cells during the pulsing experiment, we plotted a time trace of the growth rate against the metabolic reporter. Some parts of this curve, especially those stemming from the slow pulses (figure 5.2.C, arrows 8-11), can be related to the optimum curve. For example, parts indicated by arrows labeled 8 and 11 (figure 5.2.C), seem to follow such a curve. Also the part labeled by arrows 9 and 10 seem to follow an optimum curve, although it is not overlapping with the parts labeled 8 and 11. The lack of overlap indicates that the dynamics are more complicated than simply tracing back and forth on a path described by a fixed function of two parameters, like for example an optimum curve. This is further emphasized by the puzzling behavior of the trace that stems from the fast pulses (arrows 1-7, figure 5.2.C). During fast pulses, growth rates seem to change considerably in comparison with the change in metabolic reporter, and moreover the overall slope of the trace seems to be negative. This is striking, as we previously determined — based on the supplied cAMP concentration and observed growth rates — that during fast pulses the concentration of enzymes remained lower than the optimal concentration. In contradiction, the negative slope of the trace in figure 5.2.C now seems to suggests the opposite, namely that the concentration of metabolic enzymes is above the optimum. An explanation for this disagreement,

and for the fact that the trace doesn't adhere to a single path, could be that expression of the metabolic promoter and expression of metabolic enzymes is not completely synchronous. For example, there could be a delay between expression from the reporter gene and the expression of other CRP-controlled genes. Given that not all promoters that respond to CRP are equal, both in gene sequence and location in the chromosome, it would not be surprising if the behavior of the CRP-responsive promoters is more rich than a simple response in unison.

Nevertheless, taken together, the observations from the pulsing experiment show that we can control and monitor CRP-induced expression by extracellular cAMP in a dynamical manner, and that the cell responds to cAMP pulses that occur on sub cell-cycle, which indicates that it might also respond to stochastic fluctuations.
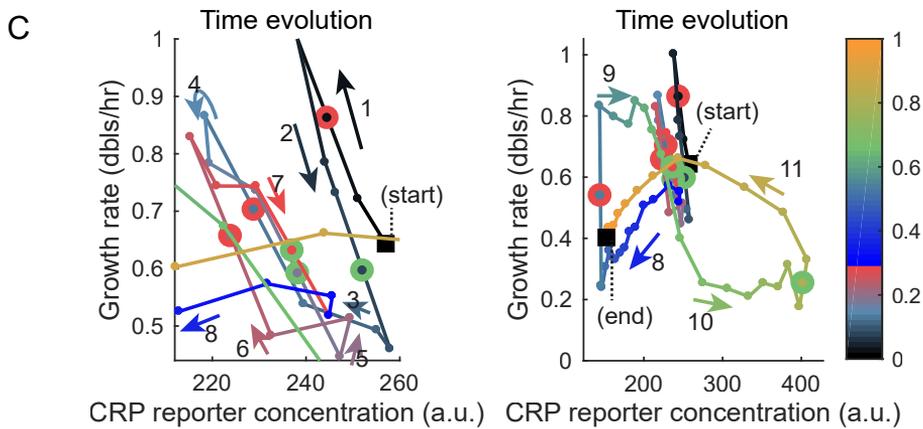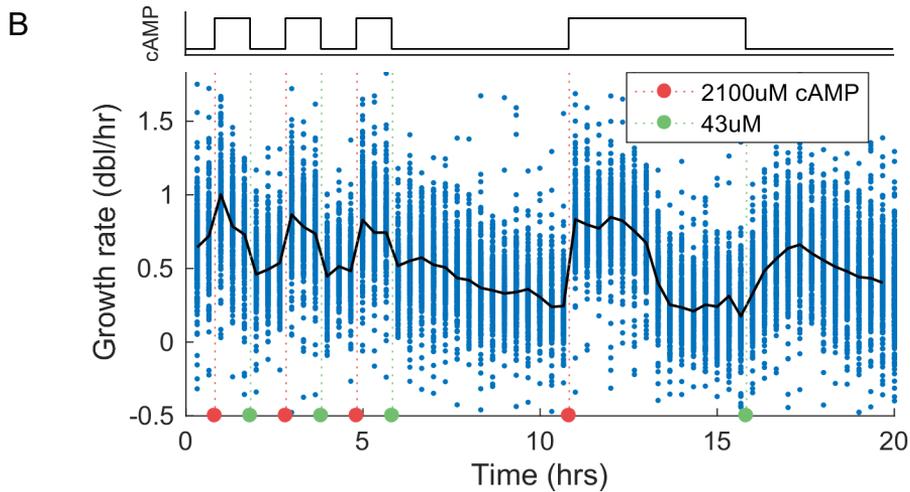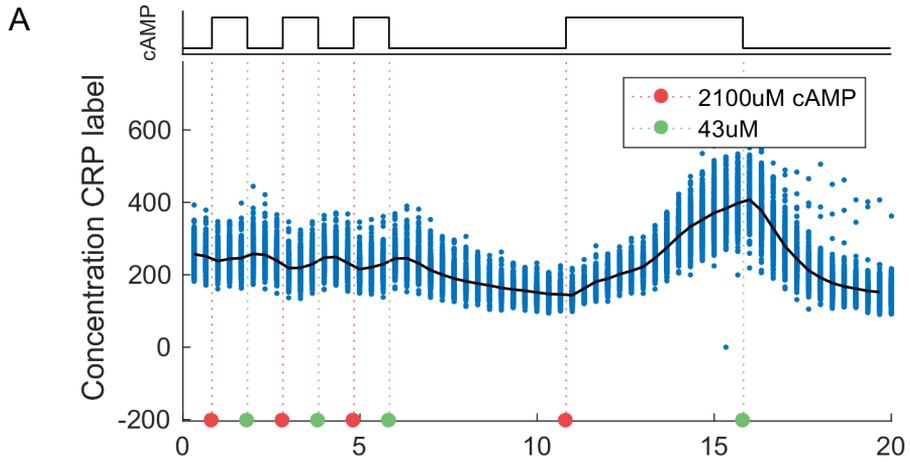
5

Figure 5.2:   The response of the CRP system to external variations. (A-B.) The concentration of the metabolic reporter (A) and growth rate (B) over time during the cAMP pulsing experiment. Blue dots correspond to single cell observations, and the black line is the population average. The concentration of cAMP in the growth medium is plotted above the graph on the same time scale, for convenience, the times of switches are also indicated by dashed lines that end in coloured circles on the x-axis. Red indicates a switch to a cAMP concentration of 2100 μM, and green a switch to 43 μM. (C) These two panels show how the relation between the growth rate and metabolic reporter concentration evolves over time, where the colors of the trace and the numbered arrows indicate the progression of time. The left panel shows the same data as the right panel, but zooms in on the part of the trace where switches occur every hour. The regime with fast pulses (every hour) is indicated by the colours black, light blue and pink, whilst the regime with slow pulses (every five hours) is indicated in blue, green and yellow. Only the population average is shown (see supplemental figure 5.20 for a few examples of single cell traces). Black squares indicate the start and end of the experiment. Big red and green circles indicate the switches towards high or low cAMP concentrations respectively (these circles are placed at the centers of line segments connecting two data points), small circles matching the color of the line segments indicate data points. In between each data point are 20 minutes. See supplemental figure 5.15 for time traces of other parameters measured during this experiment.

5.2.2. THE EFFECT OF STOCHASTIC FLUCTUATIONS ON CRP REGULATION

THE FEEDBACK LOOP AFFECTS POPULATION DYNAMICS

Since the CRP system seemed sensitive to changing input, we wanted to further probe how cells growing in a constant environment respond to intracellular stochastic fluctuations. We wanted to isolate and/or manipulate the regulatory response to these fluctuations. We aimed to remove any potential response of the regulatory system to stochastic fluctuations, while maintaining the mean expression level required in the cellular environment. This was done by growing our ΔcAMP strain on a gel pad with minimum medium supplemented with lactose and 800 μM cAMP, the cAMP concentration that we determined to be optimal (supplementary figure 5.13). On the other hand, in single cells, the CRP feedback system will be unable to respond to internal fluctuations in the metabolite concentrations. Using our metabolic reporter construct, we were able to assess the CRP dynamics in the feedbackless cells. This experimental design gave us a way to potentially decouple the regulatory response to stochastic intracellular changes from the regulatory response to environmental changes, and thus give us a way to study whether there is regulatory activity in response to stochastic changes.

We compared CRP expression dynamics in these cells without feedback against strain ASC990, which are wild type *E. coli* cells, except that they have chromosomal inserts with the metabolic and constitutive reporter constructs (see table 5.1). For simplicity, we refer to this strain as wild type. The left panel in figure 5.3.A shows a scatter plot that relates the concentration of metabolic reporter to the growth rate for a single cell experiment with wild type cells on a gel pad. The scatter cloud appears to lay on a straight line with a negative slope. The ΔcAMP cells without feedback on the other hand seem to show a relationship between metabolic expression and growth rate which lays on a flat line, instead of a negative slope (figure 5.3.A, green cloud in right panel). The contrast be-

tween the wild type cells and ΔcAMP cells suggests that the CRP regulatory interactions play a role in shaping stochastic fluctuations in the cell. To probe whether the changes were not due to generic changes in cellular dynamics, we also investigated the constitutive reporter that we introduced in our strains. As mentioned earlier, this probe was aimed to report for protein expression dynamics of proteins not regulated by CRP. Figure 5.3.B shows that the expression-growth relationship of this constitutive reporter does not change between cells with or without feedback (blue and green clouds). This supports the attribution of the change in dynamics observed in 5.3.A to the CRP regulation responding to stochastically fluctuating metabolite concentrations. We will later try to understand the nature of these changes using simulations (see section 5.2.3).

THE AVERAGE CELL

First, our "feedbackless" ΔcAMP cells allowed us to investigate a second question: how do the stochastic deviations *from* mean behaviour relate to the mean behaviour itself? To further probe the behavior of cellular populations, we performed two additional single cell experiments where we grew them at 80 or 5000 μM cAMP respectively, in addition to the previous experiment where we grew them at 800 μM cAMP.

Reassuringly, for the metabolic reporter and growth rate, the population average relationship appears to follow an optimum curve for these three conditions (i.e. growth in medium with 80, 800 or 5000 μM cAMP), see figure 5.3.A. This is consistent with the optimum curve we observed in bulk experiments (figure 5.13) and optimum curves observed in earlier experiments [153]. Strikingly, we see that also for the constitutive reporter, the population average relationship with growth rate seems to resemble an optimum curve, but mirrored along the x-axis (figure 5.3.B). This observation prompted us to think about the relationship between the concentrations of the two reporters, and hypothesize that an increase in metabolic protein expression leads to a mandatory equal decrease in constitutive protein expression and vice versa. Such an effect might result from limitations on the total protein budget of the cell, and is consistent with previously described growth laws [170]. This would predict that the the sum of the concentrations of metabolic reporter and constitutive reporter remains constant, which indeed appears to be the case (supplementary figure 5.7, bottom left panel); see also supplementary note I. This allowed us to predict a trend line for the population average relationship between constitutive reporter and growth rate, which appears consistent with our data (figure 5.3.B). Besides the concentrations of reporters, we can also quantify the production rates of the reporters in single cells. These show striking population average relationships with growth rate (supplementary figure 5.22). However, in supplementary note I, we show that the population average production rates (both for the metabolic and constitutive reporter) can be predicted well by the concentration, growth rate, and a simple mathematical model. This model was also used to generate algebraic curves that describe the relationship between production rate and growth rate, that appear to be consistent with the population average relationships between these two quantities (supplementary figure 5.22). We also plotted the relationship between the concentrations of the two reporters, and the relationship between the production rates of the two reporters (supplementary figure 5.23); the population average curves for these two relationships can also be determined from the fact that the sum of the concentrations remains constant

and by using aforementioned mathematical model, respectively. Thus, the relationships between the population average parameters can be captured well by a few rules and a simple model. In the next paragraph, we will discuss the behavior of single cells, which appears to deviate from the relationships that we observed for population averages.

THE NOT SO AVERAGE CELL

We saw in the previous paragraph that the population average behavior of the CRP system can be captured well by a simple model. However, the single cell behavior of both the wild type and the ΔcAMP cells seems to deviate from the mean behavior, as shown by the scatter clouds that show the concentration and growth rate for the single cells for both these strains (figure 5.3). The single cell data of the wild type cells simply show a negatively sloped relationship between CRP concentration and growth rate, whilst the feedbackless single cells seem to follow a flat line of constant growth rate. The single cell data points from the non-optimal cAMP concentrations (80 or 5000 μM cAMP) do show a positive and negative slope respectively (5.3.A), consistent with the concentrations being below and above the optimal concentration, but overall the clouds from the three conditions do not seem to lay on a continuous concave curve.

   This deviation from the population average behavior across conditions is even more pronounced in the concentration-growth relationship of the constitutive reporter: whilst the population averages from the different conditions follow a concave curve, the scatter clouds for all conditions — the wild type strain, and the ΔcAMP strains at, above and below the optimum — all show a negative slopes (figure 5.3.B).

   Taken together, these observations suggest that the dynamical behavior of single cells does not follow relationships that exist for average quantities between different conditions. Additionally, the fact that ΔcAMP cells that are unable to respond to internal metabolite fluctuations show substantially different behavior from those that can, indicates that the CRP regulation network is actively responding to inputs when the cell resides in a constant environment. To further understand the behavior of the single cells, we turn to cross-correlation analyses in the next paragraph.

5.2.3. THE CRP RESPONSE TO STOCHASTIC FLUCTUATIONS

We want to understand the effect of fluctuations in cellular parameters on regulatory interactions and other parameters (such as growth) at the single cell level. To understand these dynamic relationships time is an important component. In this paragraph we dissect the time dynamics with cross-correlations to show that metabolic fluctuations would transfer to cellular growth rate if it were not for the existence of the negative feedback CRP regulation.

CROSS-CORRELATIONS REVEAL MAJOR CHANGES DUE TO REMOVAL OF FEEDBACK

We start by considering the effect that single cell metabolic concentration fluctuations have on growth rates. We use the cross-correlation $R_{C_M,\mu}(\tau)$ to determine the relation between the metabolic reporter concentration (C_M) and past and future growth rates μ. More precisely, as explained in chapter 2, the cross-correlation $R_{C_M,\mu}(\tau)$ calculates the average correlation between C_M at time t with μ at time $t+\tau$. The colored lines in figure 5.4.A-B show the cross-correlation $R_{C_M,\mu}(\tau)$ for wild type (green) and feedbackless
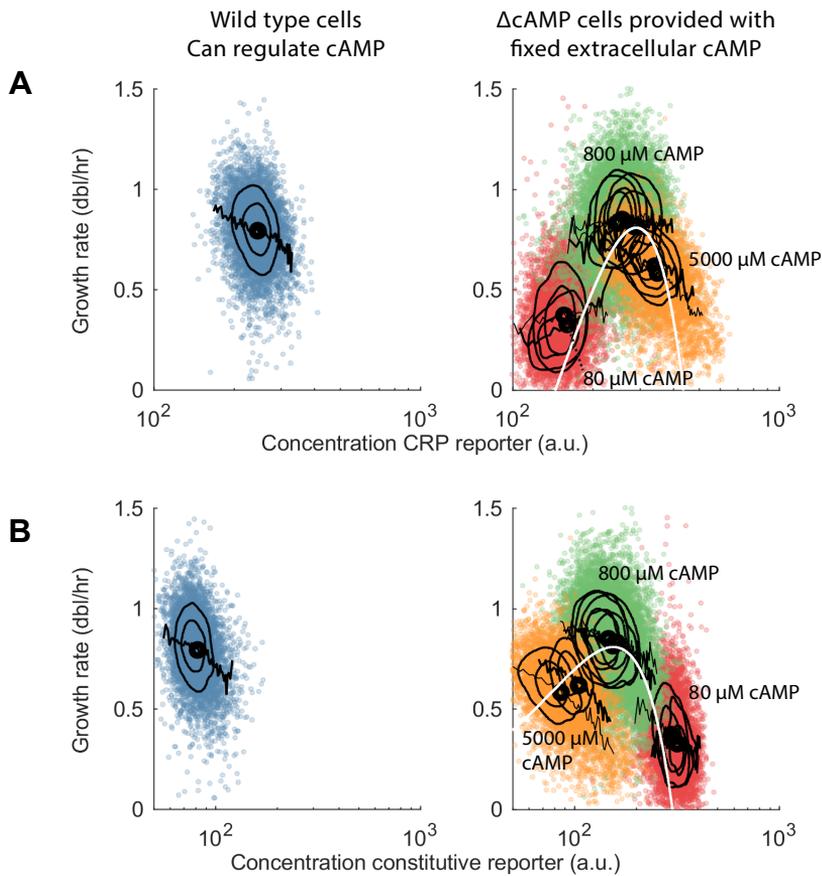
Figure 5.3: Artificial removal of negative feedback regulation changes the growth-metabolism relationship and reveals its dynamic role. (A) Colored dots show single cell growth rate values plotted against respective single cell concentrations of metabolic reporter, which is a proxy for the concentration of metabolic enzymes. The left panel shows that wild type cells which deviate from the population average (indicated by a black circle in the middle of the cloud) show a negative correlation between metabolic reporter and growth rate. The green cloud corresponds to externally supplied cAMP levels that lead to wild type growth rates (800 μM cAMP), the red and orange dots correspond to cAMP concentrations that lead to diminished growth rates (80 and 5000 μM respectively). The black lines show the average growth rate for cells that are binned according to concentration, and the black isolines reflect kernel density estimates of the probability distribution (using the Matlab function kde2d [50]). A comparison between the shape and slope of the cloud of the wild type cells with the shape and slope of the ΔcAMP cells' cloud, suggest that these two strains show different metabolic dynamics. The white trend line is a second order polynomial fitted to the population average values. (B) As panel A, except that this panel shows the relationship between growth and the concentration of a constitutive reporter. The white line is a prediction based on the optimum curve displayed by the metabolic reporter and the observation that the sum of the metabolic and constitutive reporter concentrations remains constants.

ΔcAMP cells (blue). From these curves it is immediately apparent that the dynamic metabolic concentration-growth relationship is very different between cells that possess the endogenous CRP feedback and those that do not. Firstly, as expected, the R values at $\tau = 0$ — which reflect the correlation in the scatter plots shown in figure 5.3 — are different. But secondly and more importantly, the concentration-growth relationship of the wild type cells shows a strong negative correlation at negative delays, whilst the ΔcAMP cells show correlations that are moderately positive for most delays and approximately equally strong at positive and negative delays. Thus, in wild type cells fluctuations in single cell growth rate negatively correlate with future metabolic reporter fluctuations, whilst in ΔcAMP cells growth rate and metabolic reporter concentration fluctuate in concert, but are less correlated. This indicates there is a difference in dynamics between the wild type cells with feedback and the feedbackless cells. We will later use a model to further interpret these dynamics (see the next section).

However, before further interpreting these dynamics, we will look at the relation between protein production rates and growth, as a previous study showed this can give further insights in the dynamics of metabolism and growth [10]. We calculated cross-correlations for the relation between the metabolic reporter production rate (p_M) and the growth rate μ, both for the wild type and ΔcAMP strain. For both strains, the cross-correlations show positive values, but the correlations at positive delays are more pronounced in the ΔcAMP strain (figure 5.4.A-B). This shows that in single ΔcAMP cells, there is more correlation between protein production and future growth rate. This analysis again indicates there is a difference in dynamics between the wild type cells with feedback and the feedbackless cells.

Next, we examined the behavior of the constitutive reporter, to investigate whether the difference in dynamics between wild type and feedbackless cells are indeed due to CRP regulation and not due to an overall change in protein expression-growth dynamics. We generated cross-correlations $R_{C_Q,\mu}$ and $R_{p_Q,\mu}$, which correlated the concentration C_Q and production p_Q of the constitutive reporter (indicated with a Q) with growth rate μ. These cross-correlation curves share similar features for both the wild type and the ΔcAMP cells. The $R_{C_Q,\mu}$ curves for these two strains both show negative correlations at negative delays, and the $R_{p_Q,\mu}$ curves of these two strains both show only very small correlations (5.4.C-D). The fact that these cross-correlations are similar suggests that the constitutive reporter has similar dynamics both in wild type cells as in ΔcAMP cells. This contrasts with the cross-correlations of the metabolic reporter, which are very different for the wild type and ΔcAMP cells. This is consistent with our hypothesis that the dynamics between metabolic enzyme expression and growth are affected by the feedback loop in the CRP system, and that dynamics of other protein expression are not affected by this feedback interaction. This is in turn consistent with the idea that the feedback loop is performing an active role in the wild type cells in response to stochastic signals.

In general, these observations highlight that the dynamic interactions between cellular parameters go beyond instantaneous relationships that are captured by scatter plots, but instead show richer interactions that act over delays. To further interpret these relationships we turn to a minimal model of the cell.
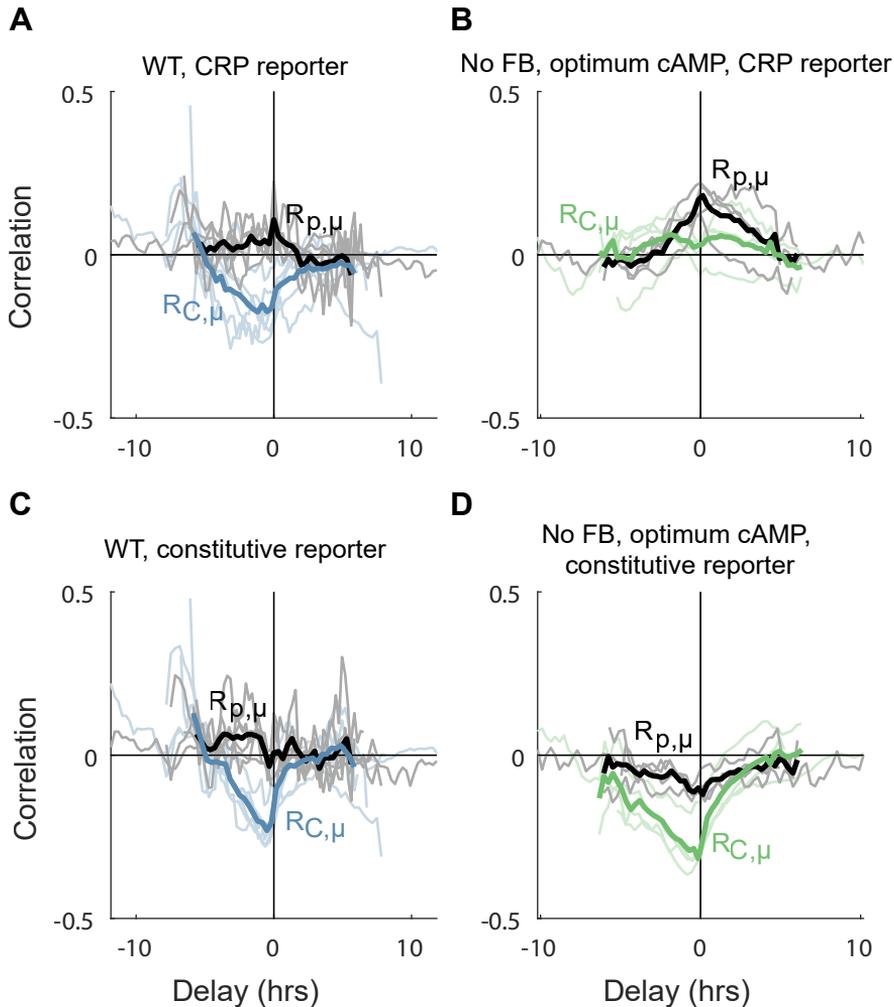
Figure 5.4: Without feedback, metabolic dynamics change and transmit to growth. (A-B) Cross-correlations $R_{C_M,\mu}(\tau)$ between single cell metabolic reporter concentration M and growth rate values μ (colored lines) and cross-correlations $R_{p_M,\mu}(\tau)$ between metabolic production rate p_M and growth rate values μ (black lines). Correlations for wild type cells are shown in A, while correlations for feedbackless ΔcAMP cells (supplemented with 800 μM cAMP) are shown in B. (C-D) Cross-correlations $R_{C_Q,\mu}(\tau)$ between single cell constitutive reporter concentration Q and growth rate values μ (colored lines) and cross-correlations $R_{p_Q,\mu}(\tau)$ between constitutive production rate p_Q and growth rate values μ (black lines). Correlations for wild type cells are shown in C, while correlations for feedbackless ΔcAMP cells (supplemented with 800 μM cAMP) are shown in D. For all panels, the correlation (which is normalized) is plotted on the y-axis, and reflects the average correlation between two parameters between time points t and $t + \tau$, the delay τ is plotted on the x-axis (in hours). The faded lines indicate cross-correlations from different microcolonies, the darker lines their averages. The cross correlations are calculated from cell lineages as described in chapter 2. Cross-correlations for cells that were grown at non-optimal cAMP concentrations are shown in supplementary figure 5.24.

A MINIMAL MODEL HELPS TO INTERPRET THE CROSS-CORRELATIONS

In the following sections we present a model that suggests the experimental data is consistent with the hypothesis that without feedback, metabolism-growth dynamics would be dominated by fluctuations that originate in metabolic protein production and metabolism itself, but that with feedback, the transmission of these fluctuations is suppressed.

To establish our model, we drew on previous models by Dunlop et al. [15], Kiviet et al. [10] and Towbin et al. [153]. The Kiviet et al. model used coupled stochastic linear differential equations to elucidate the the dynamics between growth μ, enzyme production p and enzyme concentration C. We also use these parameters, and — based on Towbin et al. — we additionally use the parameter x to represent the metabolite concentration. In our adapted model we model each of these four parameters by its own differential equation. Similar to the Kiviet model, the influence of any parameter X on any other parameter Y is mathematically modeled by coupling coefficients $T_{Y \leftarrow X}$, such that

$$\dot{Y} = (..) + T_{Y \leftarrow X} \cdot X.$$

Furthermore, also like Kiviet et al. we added terms that introduced noise into the system, as well as dampening terms, to allow us to introduce stochastic fluctuations into our model. Using such coupling terms and noise sources, we created a model for the parameters μ, p, C and x that we think represents the biology of the cell, see figure 5.5.A. In this model, the parameter x allowed for a concrete interpretation of metabolic dynamics, as growth μ and production of proteins p can be influenced by the metabolite concentration x (trough transmission coefficients $T_{\mu \leftarrow x}$ and $T_{\mu \leftarrow x}$ respectively). In turn, the protein concentration is set by production rate p and dilution rate μ, i.e.

$$\dot{C} = p - \mu C.$$

When a protein is enzymatically active, protein concentration fluctuations might influence the metabolite concentration x through transmission coefficient $T_{x \leftarrow C}$. Finally, this model also allows for a conveniently simple interpretation of the metabolite feedback onto the CRP regulation: it might be implemented as a negative contribution to the transmission coefficient $T_{p \leftarrow x}$. We propagated this model numerically to simulate different modes, which we can use as reference to understand our experimental data. See supplementary note II for a more detailed description of the model.

THE MODEL CONNECTS EXPERIMENTAL CROSS-CORRELATIONS TO TYPES OF DYNAMICS

Previously, the Kiviet et al. model presented three biologically relevant modes of how fluctuations transmit from one parameter to the next [10]. We first used our model to reproduce these modes, which are the dilution mode, the catabolic mode and the common mode. Using our model, we simulated for each mode the dynamics between production rate (p), concentration (C), metabolite concentration (x), and growth rate (μ), and calculated the appearance of the associated cross-correlations $R_{C,\mu}(\tau)$ and $R_{p,\mu}(\tau)$ (see supplemental notes II for details). In the first mode, called the dilution mode, generic fluctuations in the cellular growth rate (i.e. growth rate fluctuations not attributed to metabolism or fluctuations in the concentration of the enzyme of interest) are the largest. This leads to dynamics in which dilution by volume growth dominates the cross-correlations.

Fluctuations in growth rate are thus followed by negatively correlated fluctuations in concentration, as expressed by the negative values of $R_{C,\mu}(\tau)$ at negative τ values. Since volume growth does not interact with protein production, the correlation between our other pair of observables p and μ was found to be zero for all τ values. Both the $R_{C,\mu}(\tau)$ cross-correlations and the $R_{p,\mu}(\tau)$ cross-correlation for the dilution mode are shown in figure 5.5.B. In the second mode, the catabolic mode, fluctuations in p dominate the system, and are propagated via C and x to μ. This leads to positive values of $R_{C,\mu}(\tau)$ and $R_{p,\mu}(\tau)$ at positive τ values, as shown in figure 5.5.C. Such a mode implies that concentration fluctuations can have cell-wide consequences, even affecting the growth rate of the cell. Finally, the common mode relates to propagation of noise that arises in the metabolite concentration as a result of fluctuations arising in metabolic processes. This mode represents a situation where such fluctuations have simultaneous effects on growth and protein production, which led to positive values of $R_{C,\mu}(\tau)$ for both positive and negative τ values, whilst correlations with concentration lagged slightly behind. Cross-correlations for the common mode are shown in figure 5.5.D. These modes can be used a reference to interpret our own experimentally obtained cross-correlations. See supplementary notes II for parameter values and further details.

AN ACTIVE ROLE OF FEEDBACK IN CONSTANT ENVIRONMENT

We first compared the cross-correlations of these three modes (figures 5.5.B-D) with the cross-correlations of the constitutive reporter in wild type and ΔcAMP cells (figures 5.4.C-D). The constitutive cross-correlations, both for wild type and feedbackless cells, appeared to match with the dilution mode. In this mode, protein fluctuations are not correlated with future growth correlations, which is associated with proteins that do not have a big effect on metabolism or growth rate. This is consistent with a constitutively expressed fluorescent protein reporter that does not interact with cellular processes. This mode also provided an explanation for the negative slope in the scatter plots that show the single cell relationship between constitutive reporter concentration and growth rate (figure 5.3.B): the negative slope could originate from the response of the metabolic reporter to stochastic growth rate fluctuations.

Next, we looked at the dynamics of the metabolic reporter, which were distinctively different between the wild type cells and the feedback-less ΔcAMP cells (figure 5.4.A and 5.4.B respectively). The wild type cells showed cross-correlations that also seem similar to the dilution mode. This implied that in wild type cells (which still have feedback regulation), no fluctuations are propagated from metabolic expression to growth. On the other hand, the ΔcAMP cells without feedback, showed different cross-correlations. These were harder to directly connect to one of the modes, as the pair of $R_{C,\mu}(\tau)$ and $R_{p,\mu}(\tau)$ curves did not exclusively fit a single of the three presented modes. When we however simulated a combination of all of the three presented modes, this showed cross-correlations that were similar to the $R_{C,\mu}(\tau)$ and $R_{p,\mu}(\tau)$ curves we saw for the metabolic reporter in the ΔcAMP cells without feedback (compare figures 5.4.B and 5.5.E, see supplementary notes II for simulation details). This suggests that fluctuations in ΔcAMP cells might originate at three spots: in the production of metabolic enzymes, in metabolism and in processes related to cellular growth. At $\tau = 0$ the combined mode predicts a strong positive slope between the concentration and growth rate, and a smaller

positive slope between production rate and growth rate. The latter is indeed observed (see figure 5.22.A), and the slope between concentration and growth rate is indeed smaller, but not clearly positive (figure 5.3.A). Note however that the correlation at zero delay ($\tau = 0$) appears to be an exception, as at other delays there is a positive correlation between the metabolic reporter and growth rate.

The model also provided a way to to analyze the effect of the feedback in the CRP system. Following the topology that is described in literature, where CRP-induced genes are inhibited by metabolites (see introduction), we added a negative component to the $x \rightarrow p$ coupling; i.e. we decreased the value of $T_{x \leftarrow p}$. Strikingly, we saw that when we added feedback to the mixed mode presented in figure 5.5.E, it began to resemble the dilution mode — as presented in figure 5.5.F. This is indeed consistent with the cross-correlation curve that we saw for the wild type cells that still possess the feedback regulation 5.4.A. Thus, since the possession of the feedback loop is the key difference between the wild type and ΔcAMP strain, this suggests we can attribute the change in dynamics between ΔcAMP and wild type cells to the negative feedback. Our model shows that the change in dynamics is consistent with a scenario where wild type cells experience fluctuation from many sources — as shown by the mixed mode for the ΔcAMP cells — but that transmission of these fluctuations is negated by the negative feedback loop. This also gives an interpretation for the shape of the scatter clouds observed earlier where the metabolic reporter concentration was plotted against growth rate for wild type cells (figure 5.3.A). The negative slope might result from growth rate fluctuations, that have become dominant in the system due to the reduction of transmission of fluctuations that originate in the metabolism.

Thus, although this model is rather simple and a biological cell much more complicated, the cross-correlations are consistent with modeling results in which the CRP-regulation by negative feedback plays an active role even in a constant environment. Specifically, it suggests that the CRP regulation might reduce the transmission of metabolic fluctuations, and prevent them from having cell-wide effects, such as on the cellular growth rate.
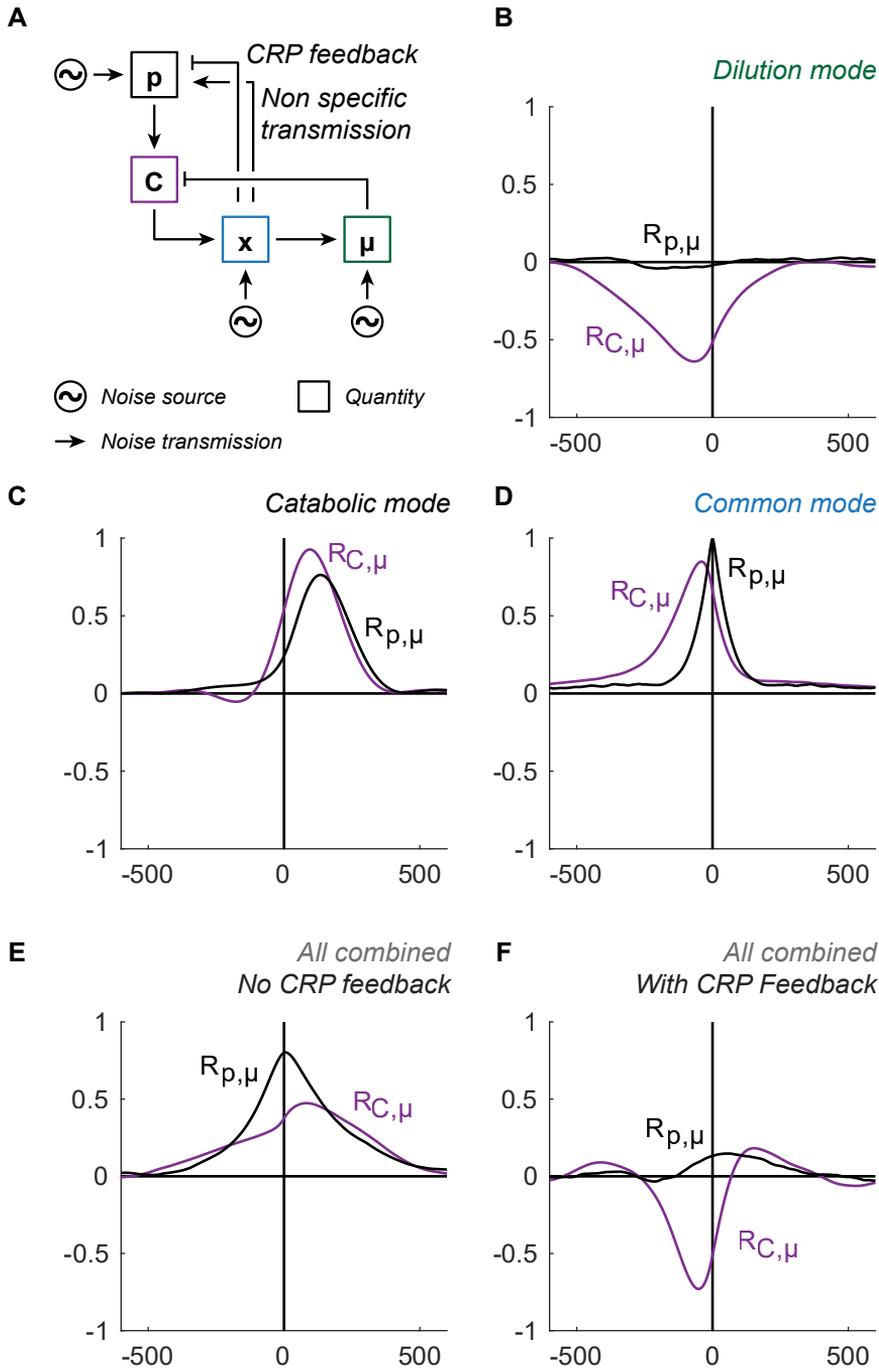
Figure 5.5: A simple model explains how feedback filters out noise transmission. (A) We used a coupled stochastic differential equations model with noise sources and dampening terms (also called Ornstein-Uhlenbeck processes). Here, p represents the production rate of any protein with concentration C that might influence the cells metabolic processes — represented here by an x that refers to the metabolic system involved that may give feedback to the production rate — and μ represents the growth rate of the cell. The production rate is set by the overall performance of the metabolism (hence the arrow from x to p), inhibitory feedback by metabolites (hence the inhibitory arrow) and noise on the production process. The concentration is set by a combination of production and dilution terms (hence the positive and inhibitory arrow pointing towards it from these quantities). The cell's metabolic performance x is set by the concentration of proteins C and noise from other cellular sources. This performance x then sets the growth rate μ, which also experiences noise from other cellular sources. All parameters that are displayed in boxes are modeled explicitly by differential equations. Arrows indicate interactions. Circles with a twiddle in it represent noise sources. See main text and supplementary note II for equations. (B) When the noise source on μ is largest, and transmission of noise only occurs through the arrow going from μ to C, this is called the dilution mode. This mode represents a case where fluctuations in the protein concentration C do not have a large effect on the cell's metabolism. This might be because the protein has no metabolic function, but it could also be that the protein does play an important cellular role but the cell is insensitive to fluctuations in the protein. (C) In the catabolic mode, noise on the production rate p is largest, and this is transmitted from p to C, from C to x and finally from x to μ. This leads to a delayed positive correlation between protein expression and growth. (D) In the common mode, the noise source on x is the largest, and this affects both production and growth simultaneously, leading to the symmetric $R_{p,\mu}(\tau)$ peak. (E) When the categories represented by panels B-D are combined, all kinds of dynamics are possible. This panel shows a combination of the three modes with an emphasis on the dilution and catabolic modes, leading to a broad $R_{C,\mu}$ correlation and a taller $R_{p,\mu}$ correlation. (F) By adding feedback to the situation in panel E (effectively decreasing the strength of the $x \to p$ interaction), these positive correlations can be suppressed, reverting the dynamics to something that is more similar to the dilution mode.

5.3. CONCLUSIONS

In this chapter, we investigated how the CRP regulatory networks responds to environmental inputs versus stochastic inputs. A negative feedback loop, which inhibits CRP activity (through cAMP) when metabolite concentrations increase, is known to be responsible for adjusting metabolic enzyme expression to the growth medium of the cell. We hypothesized that this regulatory interaction may respond also to metabolite fluctuations that occur in the cell due to noise.

We first subjected the CRP system to a quickly changing cAMP input signal. To artificially control the input signal, we used ΔcAMP cells that do not respond to internal regulation, but instead responded to the cAMP concentration that we provided in the cellular growth medium. We also introduced a reporter construct to gauge the metabolic enzyme expression levels. The cellular response to an input that alternated between a high and low signal showed that it takes hours before cells reach a new steady state, but much less than an hour before a change in metabolic expression or growth rate can be detected. We thus concluded that cellular networks have the potential to react at sub

cell-cycle timescales to changing inputs such as stochastic fluctuations in metabolite concentration.

Next, we aimed to uncover the regulatory dynamics, if any, of the CRP system to stochastic input signals. We grew ΔcAMP cells in 800 μM cAMP, the optimal concentration, which sustained growth rates that compare to wild type growth rates. This resulted in cells that expressed the right amount of metabolic enzymes, but whose CRP regulation could not respond to stochastically changing metabolite concentrations. By comparing metabolic expression versus growth scatter plots of these ΔcAMP cells with scatter plots of wild type cells (which could respond to both environmental and stochastic inputs) we saw that metabolism-growth dynamics was markedly different between the two conditions. Since the absence of stochastic input to the CRP system resulted in a change in metabolism-growth dynamics, this suggested that the CRP system responds to stochastic fluctuations in metabolite concentrations.

To further investigate what these changes in dynamics mean, we not only compared metabolic expression values with growth rates that were both obtained at the same point in time, but also calculated the correlation between metabolic expression with past, current and future growth rates. This was done by calculating cross-correlations. Using cross-correlations we saw that correlations in the feedback-less ΔcAMP cells and wild type cells were different over a wide range of delays. This strengthened our idea that stochastic fluctuations have a profound effect on the CRP regulatory system dynamics.

In order to understand the meaning of the difference in dynamics, we used a stochastic differential equation model. This model suggested that the positive correlations we observed between metabolic reporter expression and growth in feedbackless ΔcAMP cells are consistent with dynamics that are influenced by fluctuations in enzyme expression, metabolism itself and volume growth. The negative correlations between metabolic concentration and past growth rate observed in wild type cells are consistent with dynamics dominated by fluctuations in volume growth only. This difference between wild type and ΔcAMP cells, i.e. cells with and without feedback respectively, suggested to us that the negative feedback regulation might actually prevent fluctuations in metabolism from propagating throughout the cellular biochemical network.

5.3.1. DISCUSSION AND OUTLOOK

Given the idea that fluctuations might be prevented from propagating, we also investigated whether the coefficient of variation (CV) of the growth rate was higher for ΔcAMP cells compared to wild type cells. As supplementary figure 5.21 shows, it is indeed very slightly higher, but more data is required to make statistical claims.

Nevertheless, we observed that the CRP metabolic regulation responds to stochastic fluctuations in cellular concentrations. We know that the CRP regulation also functions to adjust enzyme expression in response to the extracellular environment. The negatively sloped relationship between metabolic enzyme expression and growth that we observe (figure 5.3), brings to mind the negatively sloped C-line [170] that can be used to understand growth in different conditions. The idea behind the C-line and associated growth laws is that some carbon sources provide a higher energy yield, which allows cells to express less metabolic enzymes and spent more resources on other cellular processes, leading to a higher growth rate. Since cells in our experiments all grow

in the same environment in the same sugar, such an explanation however does not apply to our observations. We argue that the trend we observed should be viewed in the context of the underlying growth-metabolism dynamics, instead of the context of the system's response to perturbations in the cellular environment. This might also apply more broadly. The response of biochemical networks to the change in one parameter (e.g. a change in sugar source), is often well understood. However, stochastic fluctuations might occur in all cellular components, meaning that many input parameters to regulatory networks are constantly changing. This is perhaps why it is hard to understand the fluctuations using a mechanistic description of a regulatory network (it is for example hard to apply equations in supplementary note III and understand the stochastic dynamics), but instead it is possible to describe the mechanics with a coarse grained linear model as we did (figure 5.5).

Another question one might ask is: why do not all single cells simply grow at the rate of the fastest growing cell in the population? Our observations indicate that fluctuations can have large effects on the cellular state, as illustrated by the correlation between metabolic expression and the growth rate. It might simply be too difficult for the cell to control such fluctuations and set a constant high growth rate. On the other hand, a versatile cellular population might have an evolutionary benefit (see also chapter 4). It could be that therefor there is no evolutionary pressure for cells in a population to all become fast growing and similar individuals.

To further probe these questions, one could device additional experiments. Firstly, an interesting alternative to also create a feedbackless strain, would be to create a construct with an inducible promoter that expresses constitutively active CRP protein. A challenge in this experimental approach might be that the known constitutive forms of CRP have a rather low activity, or are still mildly responsive to cAMP [186, 187]. Secondly, it might be interesting to probe different conditions, and e.g. change the sugar source in the growth medium. For certain sugar sources, like glucose, we expect similar behaviour as observed in this manuscript, but some sugar sources might be more challenging for the cell to handle, and alter the dynamics. For example, xylose is a sugar source for which it was observed that concentrations need to fall in a very narrow concentration range to observe growth at all [188]. Growing cells in such a sugar source might provide an opportunity to observe a bigger effects of feedback removal. Or, other interesting alternate sugar sources are pyruvate, glycerol and galactose. These are sugar sources where the metabolic enzyme expression is known to be sub-optimal, since the negative feedback loop only provides a "rule of thumb" to the cell, which is not adequate for these sugar sources [153]. This might lead to different dynamics also. Follow-up experiments could also involve further investigating the role of the α-ketoacids, which could be introduced in experiments to disturb the feedback loop. More broadly speaking, CRP is a versatile regulatory protein.As pointed out by the *Ecocyc* database [171], it is also involved in osmoregulation [189], stringent response [190], biofilm formation [191], virulence [192] nitrogen assimilation [193, 194], iron uptake [195], competence [196], multidrug resistance to antibiotics [197], and expression of CyaR sRNA [198]. An open question remains what the effect of fluctuations in CRP activity is on these processes. And finally, there are many other regulatory systems that are known for their regulatory role in responding to environmental changes. It will be interesting to investigate whether all of these systems

are also actively responding to stochastic input signals. If so, that would mean that the often used term "growing in steady state" requires thorough revision.

5.4. METHODS

ASC number	Shorthand	Description	Source
ASC838		Wild type MG1655 strain obtained from Benjamin Towbin, Uri Alon lab (also known as strain bBT12 and CGSC number 8003). This is the basis for all Towbin et al. strains. Known mutations: λ-, Δfnr-267, rph-1. (No resistance modules.)	[153]
ASC839		*cyaA, cpda* null mutant. Obtained from Benjamin Towbin, Uri Alon lab (also known as strain bBT80). Based on ASC838. (No resistance modules.)	[153]
ASC841	p_s70	MG1655 wild type strain with modified lac promoter fused to GFP (sigma 70 reporter) on pSCS101 plasmid. (Kanamycin resistant.)	[153]
ASC842	p_CRPr	MG1655 wild type strain with modified lac promoter fused to GFP (CRP reporter) on pSCS101 plasmid. (Kanamycin resistant.)	[153]
ASC990		Wild type strain, except for Δ(galk)::s70-mCerulean-kanR and Δ(intc)::rcrp-mVenus-cmR. (Kanamycin and chloramphenicol resistant.)	VS
ASC1004	ΔcAMP	Strain based on ASC839 (ΔcyaA Δcpda), introduced Δ(galk)::s70-mCerulean-kanR and Δ(intc)::rcrp-mVenus-cmR. (Kanamycin and chloramphenicol resistant.)	VS

Table 5.1: Strains used in this work. ASC stands for AMOLF strain collection. VS indicates this strain is produced at AMOLF by technician Vanda Sunderlikova.

Strains See also table 5.1. Strains ASC838, ASC839, ASC841 and ASC842 were a kind gift from the Alon lab. The sequence of the modified lac promoter which reports for CRP activity (p_CRPr) is:

CGTCAGGAGGAGAGGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCT
CACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGT
TGTGTGCATGGATAAGTAGCTAGGAATTTCACACTGCAAACAGCT.

Which is the LacZ promoter with LacI1 site reshuffled (created by Towbin et al. using synthetic oligos [153]). The sequence of the lac promoter modified to report for constitutive fluctuations (p_s70) is:

CGTCAGGAGGAGAGGGGCAGTGAGCGCAACGCAATCAGATCAAATGTGTC
GTTTCCATAGGCACCCCAGGCTTGACACTTTATGCTTCCGGCTCGTATAA

TGTGTGCATGGATAAGTAGCTAGGAATTTCACACTGCAAACAGCT.

This sequence has reshuffled CRP binding sites and a reshuffled lacI binding site. The sigma 70 site was changed to a consensus binding site [153]. Figure 5.6 highlights some important features of the LacZ promoter and indicates where the synthetic promoters differ from the native promoter. To be able to use the p_CRPr and p_s70 reporters in the same strain, the reporters in strains ASC841 and ASC842 were used to create new reporters using the promoters created by Benjamin Towbin but fused mCerulean and mVenus fluorescent proteins. These reporters were chromosomally inserted in the ASC838 and ASC839 strains using a lambda red protocol, resulting in the strains labeled ASC990 and ASC1004, respectively. Strains that were obtained or made for the purpose of this project but not used for experiments presented here, are listed in supplementary table 5.2. All strains listing ref. [153] as source were kindly supplied by the Alon lab.

Pulsing experiment. Strain asc1004 was grown O/N at 30 °C and 10X concentrated by spinning the cells down at 2300 RCF, removing supernatant and resuspension in a table top centrifuge. Cells were introduced into microfluidic device 2 (see chapter 3) with a syringe, after which the device was placed under the microscope in a 37 °C temperature chamber and we supplied TY medium (flow rate 8 μl/min) for a few hours, whereafter we switched to M9 minimal medium plus 0.2 mM uracil, 0.1 % lactose, 0.01 % tween and 300 μM cAMP (flow rate 7 μl/min). After this cells were grown in the same medium but supplemented with 0.001 % tween and sequentially 1hr of 2100 μM cAMP ("high") and 1hr of 43 μM cAMP ("low"), which was repeated 5 times (totaling 10hrs), and then 5 hrs low, 5 hrs high, and 5 hrs low (all at a flow rate of 8 μl/min). Times at which the valve switches were recorded and in the analysis corrected by adding the arrival delay of 58 minutes (in this particular experiment that delay was not yet optimized). Fluorescent images were taken every 20 minutes, using a CFP and YFP filter set (chroma models 49001 and 49003 respectively), both with exposure times of 150ms. A selection of this sequence was analyzed and displayed. Data in the figure 5.3 is based on experiments in which the same physical xenon arc light bulb was used to measure fluorescence for each experiment. Figure 5.4 is based on experiments in similar conditions, but contains more experiments, which also include experiments with a different fluorescent light bulb.

Gel pad experiments. A detailed description of the protocol for gel pad experiments including a list of the involved chemicals can be found in chapter 3. Briefly, polyacrylamide gel pads were soaked (at 37 °C and on a shaker) trice in the desired growth medium for a period of 30-90 minutes right before the experiment. In this case, M9 minimal medium was used supplemented with lactose (0.01 % g/mL) and uracil (0.2mM), and only in the last wash step also with Tween20 (0.001%). Also, the medium was supplemented with the desired concentration of cAMP (Sigma Aldrich) if applicable. 1 μl culture of the applicable strain was grown O/N in the same medium (OD 0.005; or diluted to that OD if necessary from exponential growth phase) and inoculated on the gel pad. Data acquisition was then performed as described in chapters 2 and 3.

More detailed information on the computer analyses. For a description of the computer analyses and more information see chapters 2 and 3. However, some details are also worth mentioning here. Since concentrations, production rates and growth rates play a large role in this chapter, it is good to indicate how they are calculated. Concentration (a.u./px) is defined as the mean fluorescence signal (a.u.) over all pixels of the
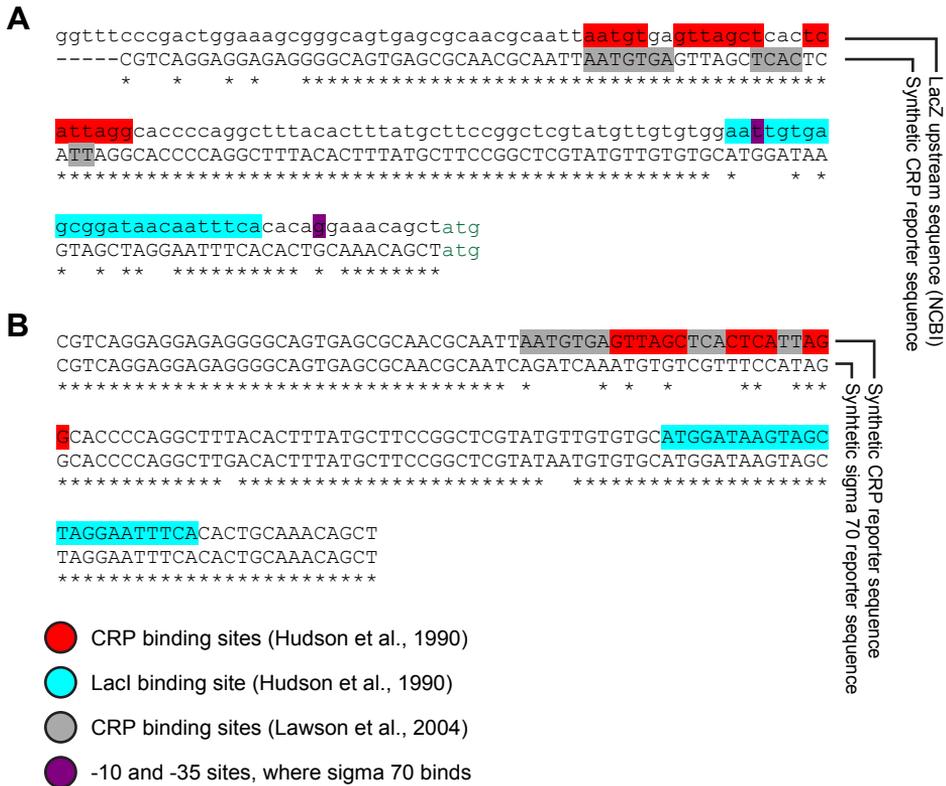
Figure 5.6: Promoter constructs used in this manuscript for metabolic and constitutive reporters. This figure displays additional details about the promoter sequences used for the reporters in this study. These promoter sequences were designed by Towbin et al. [153]. This image supplies additional information about them. (A) This panel compares the upstream sequence of the LacZ coding sequence (and the start codon, shown in green) as retrieved from the NCBI database (accession: NC_000913; region: 363231..366305; version: NC_000913.3) with the synthetic CRP reporter promoter sequence by Towbin et al [153]. For reference, the CRP binding sites as reported in refs [199] and [200] are highlighted, as are the LacI binding site and the -10 and -35 upstream locations. (B) This panel shows the promoter for the CRP reporter construct compared with the sigma 70 (constitutive) promoter sequence by Towbin et al. [153].

bacterial cell. To calculate the production rate (a.u. px^{-1} min^{-1}), first the sum of the fluorescence signal (a.u.) over all pixels that make up a cell is calculated. For each frame n where a fluorescence image was taken, the production rate is determined as the slope of a linear fit through three points $n-\delta n$, n, and $n+\delta n$, where δn is the interval at which fluorescence pictures are taken. This production rate is then subsequently divided by the total number of pixels of the cell in frame n. The growth rate (units dbl/hr when the symbol μ is used and units /min when the symbol λ is used) is determined for each frame n by fitting an exponential curve through frames $n-\delta n/2$ until $n+\delta n/2$ (or sometimes a smaller range). Note that to determine scatter plots and correlations, only frames where fluorescence images were taken are considered. This means that the choice for a growth rate fitting window of width $\delta n+1$ ensures that all data is used, but no point is used twice

5.5. **ACKNOWLEDGEMENTS**

5

SUPPLEMENTARY NOTES

I. UNDERSTANDING STEADY STATE VALUES OF METABOLIC AND CONSTITUTIVE PRODUCTION RATES, CONCENTRATION, AND GROWTH RATES

MOTIVATION

The purpose of this chapter is to understand single cell stochastic deviations away from steady state. However, understanding the steady state relationships themselves between metabolic and constitutive protein production, concentration and the cellular growth rates helps us place the stochastic fluctuations away from the steady state values in the right context. We explain our understanding of this matter in this supplementary.

MODEL AND FITTING OF MODEL TO DATA

To better understand steady state relationships it is convenient to write down the differential equations that describe the (supposed) time evolution of the system, and consider the steady state condition where the time derivative is zero. Let us here start with an equation for the metabolic reporter. Generally, the change in metabolic protein concentration C_M (a.u/(px)) over time is set by a production and a dilution term:

$$\dot{C}_M = p_M - C_M \lambda = \phi_M(\text{cAMP})\lambda - C_M \lambda, \tag{5.1}$$

where p_M is the production rate (a.u/(px min)) of metabolic protein and λ (/min) the growth rate of the cell. (We use the symbol λ here for growth rate instead of μ to be able to distinguish the units used; we use μ for growth rates which are expressed in doublings/hour.) We can determine all these three parameters from our single cell experiments. Following Towbin et al. [153], we additionally propose $p_M = \phi_M(\text{cAMP})\lambda$, where ϕ_M (a.u/(px)) is a function of the concentration of cAMP and scales proportionally to it; the fractional production ϕ_M sets the production rate p_M as a fraction of λ. To understand the interaction between the metabolic reporter expression and the constitutive reporter expression we were inspired by the linear regulation model posed by You et al. [170], and pose

$$\phi_Q = T - \phi_M, \tag{5.2}$$

where ϕ_Q is the fractional production of the constitutive reporter, and T the total fractional production of both C_M and C_Q, which we consider a constant independent of the cAMP concentration. (Note that this ignores interaction with other cellular protein sectors such as ribosomes.) This implies that

$$\dot{C}_Q = p_Q - C_Q \lambda = \phi_Q \lambda - C_M \lambda = (T - \phi_M(\text{cAMP}))\lambda - C_M \lambda, \tag{5.3}$$

where C_Q is the concentration of the constitutive reporter, p its production rate and ϕ_Q its fractional production rate. Based on equations 5.1-5.3 and the steady state assumption $\dot{C}_M = \dot{C}_Q = 0$ we can make some predictions. Firstly, this suggests that the production rates (a quantity that we determine directly from our experiments) divided by the growth rates should be equal to the concentration for both the reporters. The top two panels in figure 5.7 show that this is indeed the case. Secondly, it implies that the sum

of the concentration of the two reporters should be equal, since $\phi_M(\text{cAMP}) = C_M$ and $\phi_Q(\text{cAMP}) = C_Q$ and thus

$$T = C_M + C_Q. \tag{5.4}$$

This relationship is confirmed in the bottom left panel of figure 5.7. One might have expected that ribosomal expression also factors into the relationship between C_M and C_Q — e.g. $T = C_Q + C_M + R$ with R the ribosomal concentration. Surprisingly, equation 5.4 implies that this is not the case here; R either remains constant or R does change with C_M but does not affect C_Q. (Note that the observations in [170] are in wild type cells; so the observations here do not necessarily contradict their observations.) In any case, the steady state condition implies that $\phi_M(\text{cAMP}) = C_M$ (equation 5.1), and thus that we can set the concentration of C_M directly by adjusting the supplied cAMP concentration. In other words: $C_M \propto \text{cAMP}$. Furthermore, Towbin et al. observed a concave relationship between the CRP concentration and growth rate λ, which they call the "O-curve" (where the O stands for "open loop" [153]) and we call the optimum curve. We indeed observe a similar relationship, as shown in figure 5.7. Towbin et al. derive a functional form for this relationship, but for simplicity we have fitted a 3rd degree polynomial to this O-curve, i.e.

$$\lambda = f_\lambda(C_M) = aC_M^2 + bC_M + c, \tag{5.5}$$

where a, b and c are fitting parameters. Using the relationships we have seen so far (figure 5.7), we can predict the relationships between the parameter pairs λ-C_Q, p_M-p_Q, p_M-λ, and p_Q-λ. Firstly, the relationship between λ and C_Q is defined by:

$$\lambda = f_\lambda(T - C_Q), \tag{5.6}$$

which is indeed supported by the top left panel in figure 5.8. Secondly, the relationships between p and λ can be parameterized by C_M and C_Q:

$$\begin{aligned}(p_M, \lambda) &= \left(C_M f_\lambda(C_M), f_\lambda(C_M)\right) \\ (p_Q, \lambda) &= \left(C_Q f_\lambda(T - C_Q), f_\lambda(T - C_Q)\right);\end{aligned} \tag{5.7}$$

which can be found using the steady state condition and 5.1. These equalities in equation 5.7 are corroborated by the bottom left and bottom right panels in figure 5.8. Thirdly, we find that p_M and p_Q can be related through parameterizing (p_M, p_Q) by C_M:

$$(p_M, p_Q) = \left(C_M f_\lambda(C_M), (T - C_M) f_\lambda(C_M)\right). \tag{5.8}$$

This relationship is indeed confirmed in the top right panel of figure 5.8. Note that no additional fitting was performed in figure 5.8.

CONCLUSION

In conclusion, three ingredients determine the steady state relationships between metabolism, constitutive protein expression, and growth in this chapter: (1) The interaction between protein production and dilution that set the concentration, (2) the concave relationship ("O-curve") between metabolism and growth (figure 5.7 bottom right panel) and (3) the fact that the total concentration metabolic and constitutive expression remains constant. For a more extensive treaty and mathematical description of the steady-state CRP regulation the reader is referred to Towbin et al. [153].
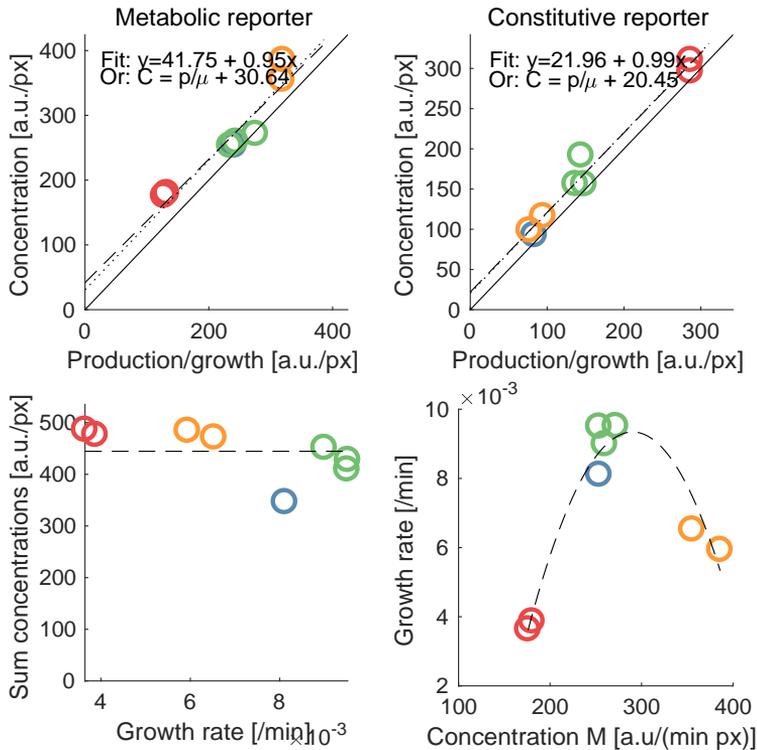
Figure 5.7: Fitted relationships between concentration, production rate and growth rate. Circles are based on population means of microcolonies of wild type cells (blue circles) or ΔcAMP cells supplied with 80, 800 and 2000 μM cAMP (red, green and yellow circles respectively), which were grown on gel pads with lactose-supplemented minimal medium; this data was also used for the scatter plots (figures 5.3, 5.22, 5.23). (Top) Concentration is predicted to equal production rate divided by growth rate. Though there is a minor offset to the line x=y (solid black line) as shown by the fits (dashed and dotted lines), this prediction seems approximately correct. (The dotted line only is based on the average offset, whereas the dashed line is a polynomial fit.) (Bottom left) This panel shows that the sum of the metabolic and constitutive reporter concentrations remains approximately equal. The mean sum is 445 a.u./px. (Bottom right) This panel shows the concave relationship between growth and the metabolic reporter concentration. The dashed line shows a second degree polynomial fit to the data points, which describes the optimum curve. The fitted parameters (equation 5.5) are $a = -4.38 \cdot 10^{-7}$, $b = 2.55 \cdot 10^{-4}$ and $c = -0.0276$.
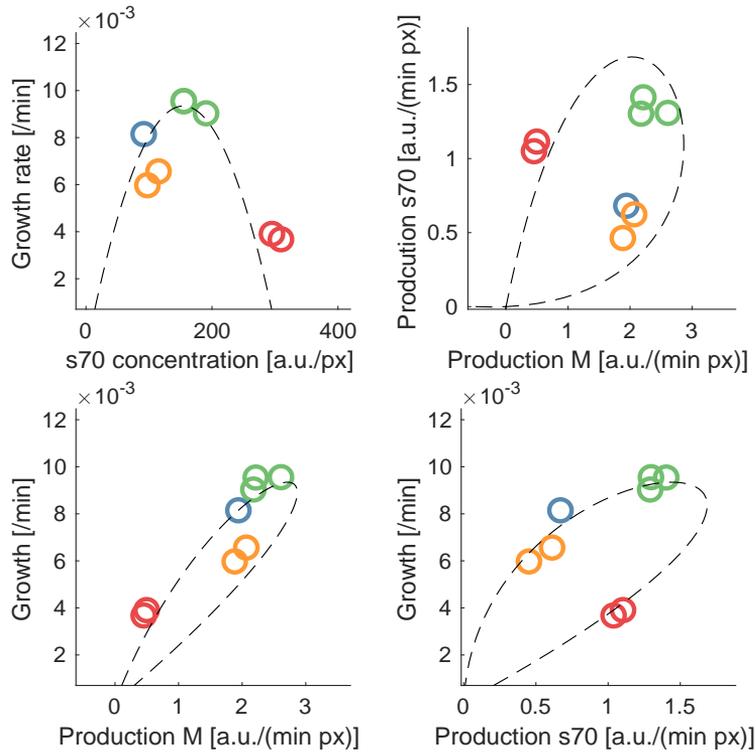
Figure 5.8: Observed and predicted (not fitted) relationships. As figure 5.7, circles are based on population means of microcolonies of wild type cells (blue circles) or ΔcAMP cells supplied with 80, 800 and 2000 µM cAMP (red, green and yellow circles respectively), which were grown on gel pads with lactose-supplemented minimal medium; this data was also used for the scatter plots (figures 5.3, 5.22, 5.23). (Top left) The observed and predicted (equation 5.6) relationship between the constitutive reporter production and growth rate. (Top right) The observed and predicted (equation 5.7) relationship between the constitutive and metabolic production rates. (Bottom left) The observed and predicted (equation 5.7) relationship between the metabolic production rate and growth rate. (Bottom right) The observed and predicted (equation 5.8) relationship between the constitutive production rate and growth rate.
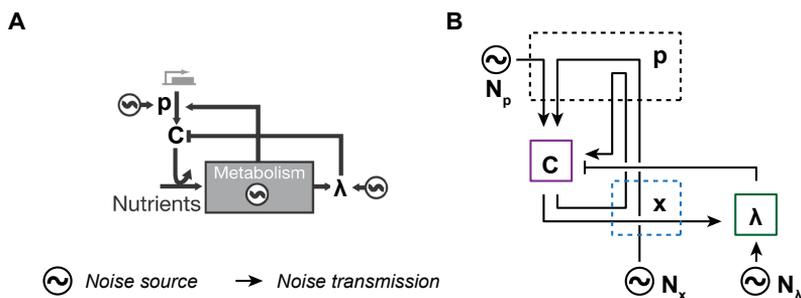
Figure 5.9: (A) Original diagram as printed in Kiviet et al. [10], except that we updated the symbols to match our notation. p is production rate, C is the concentration of the protein of interest, and λ the growth rate. (B) More technical diagram that describes the relations between parameters that are described by ODEs (boxes with solid outlines), and noise sources N_p, N_x, and N_λ, as they were modeled in Kiviet et al. C stands for the concentration of the protein of interest, λ for growth rate. Arrows indicate how ODEs or functions that describe the parameters of interest are coupled. Arrows that go trough the dashed box labeled p correspond to terms that not only couple the two parameters connected by the arrow, but also make up the production rate parameter p (see also main text). Arrows that go through the dashed box labeled x are arrows which are thought to be biologically connected to the metabolism. Compared to the model used in this manuscript, which describes all four parameters p, C, λ and x by ODEs (see also figure 5.5.A), Kiviet et al. model less parameters by ODEs [10].

II. ORNSTEIN-UHLENBECK FLUCTUATIONS MODEL

INTRODUCTION

In these notes I discuss a model which models how noise originates and transmits from and between observables respectively. I will discuss the model proposed by Philippe Nghe in the work by Kiviet et al. [10], which is more succinct than the model used in this chapter.

Parameters that are considered in general are the number of enzymes (C), the rate at which these enzymes are made (p), and the growth rate of the cell (λ)[3]. There is also the parameter x, which can be interpreted as a general noise source or the concentration of metabolites. Typically, enzymes disappear by dilution due to growth. Furthermore, there are noise sources, which add noise to these parameters. Depending on how they are implemented these source appear as N_x, N_λ and N_p or Γ_x, Γ_λ and Γ_p, the subscripts indicate to which parameter the noise source adds noise.

The goal of this model is to interpret the cross-correlations between λ, C and p — $R_{C,\lambda}(\tau)$, $R_{p,\lambda}(\tau)$ — that are obtained from experimental data.

IMPLICIT NOISE EQUATIONS

A straight forward way to model a system with these parameters, is writing down an ordinary differential equation (ODE) for each parameter involved. The Kiviet et al. model

---

[3]In this chapter, we use both the symbols λ and μ to indicate growth rate, λ when the units are /min and μ when the units are doublings/hour.

takes a somewhat different approach (it is more concise), which will be discussed later. The ODEs below relate to the cartoon in Fig. 5.5.A and describe the dynamics per parameter:

$$\dot{x} = -\frac{(x - x_0)}{\tau}$$
$$+ c_x \cdot \Gamma_x$$
$$+ T_{x \leftarrow C} \cdot c_x' \cdot (\frac{C}{C_0} - 1) \tag{5.9}$$

$$\dot{\lambda} = -\frac{(\lambda - \lambda_0)}{\tau_\lambda}$$
$$+ c_\lambda \cdot \Gamma_\lambda$$
$$+ T_{\lambda \leftarrow x} \cdot c_\lambda' \cdot (\frac{x}{x_0} - 1) \tag{5.10}$$

$$\dot{p} = -\frac{(p - p_0)}{\tau_p}$$
$$+ c_p \cdot \Gamma_p$$
$$+ T_{p \leftarrow x} \cdot c_p' \cdot (\frac{x}{x_0} - 1)$$
$$+ R_{p \leftarrow x} \cdot c_p' \cdot (\frac{x}{x_0} - 1) \tag{5.11}$$

$$\dot{C} = p - \lambda C \tag{5.12}$$

Where x describes the state of the metabolism, λ is the growth rate, p is the production rate, C is the amount of enzyme, τ is a dampening term (X_0 is the equilibrium value), $T_{X \leftarrow Y}$ is the noise transmission constant from X towards Y, c_X and c_X' are constants that set the size of the fluctuations, Γ_X is a white noise source. $R_{X \leftarrow Y}$ indicates a regulatory interaction, an addition to the model, but this notation is just cosmetic, as $T_{\text{effective}} = T + R$. This model assumes all parameters have an average value from which fluctuations deviate, but always return. Hence the dampening terms. With respect to transmission, I furthermore rescale the absolute value of the noise to be comparable to the target noise (hence the x_0^{-1} and c_X terms in combination with the $T_{X \leftarrow Y}$ term).

   This is similar to the model that Philipe Nghe suggested in Kiviet et al. [10], which was inspired by Dunlop et al. [15] (see supplement of that manuscript for a description of the Dunlop model). A difference between my equations, the Nghe equations and the Dunlop equations lies in the dampening terms (those containing τ, β or λ_E). In my model noise is effected through the ODE, and dampening occurs on the parameter of interest. In Dunlop et al., there are two dampening terms, one specifically dampening the noise and a second term dampening the parameters of interest. Nghe also takes the latter approach, see below.

Analytical expressions for the correlations between these equations can be found by linearizing them, writing the correlations in Fourier space, and back-transforming them using residue integration techniques. These notes do not cover this. Out of practical considerations, we numerically solved Eq. 5.9-5.12 by simple Euler propagation implemented in Matlab.

SETTINGS OF THE NUMERICAL IMPLEMENTATION

The script `growthnoisepropagatorv2.m` that numerically solves the equations will be made available in an online Github repository, parameter settings that were used are shown below. For the dilution mode:

$$C_0 = 2000, \lambda_0 = 1, \lambda_0 = 0.0116,$$
$$p_0 = 23.1049, x_0 = 1, \tau_\lambda = 120, \Gamma_\lambda = 8.3255e - 04, \Gamma'_\lambda = 8.3255e - 04,$$
$$\tau_p = 120, \Gamma_p = 0.0048, \Gamma'_p = 0.0048, \tau_x = 120, \Gamma_x = 1.0000e - 03, \Gamma'_x = 1.0000e - 03,$$
$$T_{x \to \lambda} = 0, T_{x \to p} = 0, T_{C \to x} = 1, T_{\lambda \to C} = -1;$$

(λ is given in units of min^{-1} here.) For the catabolic mode:

$$C_0 = 2000, \lambda_0 = 1, \lambda_0 = 0.0116,$$
$$p_0 = 23.1049, x_0 = 1, \tau_\lambda = 60, \Gamma_\lambda = 0, \Gamma'_\lambda = 8.3255e - 06,$$
$$\tau_p = 60, \Gamma_p = 1.5200, \Gamma'_p = 1.5200, \tau_x = 60, \Gamma_x = 0, \Gamma'_x = 0.3162,$$
$$T_{x \to \lambda} = 0.9000, T_{x \to p} = 0, T_{C \to x} = 0.9000, T_{\lambda \to C} = -1;$$

For the common mode:

$$C_0 = 2000, \lambda_0 = 1, \lambda_0 = 0.0116,$$
$$p_0 = 23.1049, x_0 = 1, \tau_\lambda = 6, \Gamma_\lambda = 0, \Gamma'_\lambda = 8.3255e - 06,$$
$$\tau_p = 6, \Gamma_p = 0, \Gamma'_p = 0.0481, \tau_x = 60, \Gamma_x = 0.3162, \Gamma'_x = 0.3162,$$
$$T_{x \to \lambda} = 0.9000, T_{x \to p} = 0.9000, T_{C \to x} = 0, T_{\lambda \to C} = -1;$$

For the combined scenario:

$$C_0 = 2000, \lambda_0 = 1, \lambda_0 = 0.0116,$$
$$p_0 = 23.1049, x_0 = 1, \tau_\lambda = 60, \Gamma_\lambda = 2.4034e - 04, \Gamma'_\lambda = 2.4034e - 04,$$
$$\tau_p = 60, \Gamma_p = 0.5308, \Gamma'_p = 0.5308, \tau_x = 60, \Gamma_x = 0.1459, \Gamma'_x = 0.1459,$$
$$T_{x \to \lambda} = 0.2700, T_{x \to p} = 0.3000, T_{C \to x} = 0.2700, T_{\lambda \to C} = -1,$$

Cross-correlations were calculated based on 100000 one-minute timesteps, noise was introduced with the matlab function `normrnd`. The feedback is added by subtracting 0.27 from the $T_{x \to \lambda}$ parameter, i.e. setting it to zero.

SEPARATE NOISE EQUATIONS

Both Nghe and Dunlop define separate ODEs for the noise terms:

$$\dot{N}_X = \sqrt{I_X} \cdot \Gamma_X - N_X / \tau, \tag{5.13}$$

though their notation might be slightly different (I used Daniel Gillespie's notation [201]; a capital I is used here to follow Gillespie's square root notation, $\sqrt{I_X} = i_x$). With for our case X equaling λ, x or p. Note that $\tau^{-1} = \beta$ (β is used in Nghe and Dunlop).

   Not so relevant for our case, but noteworthy, is that in the Dunlop model, which models a completely different process than the one described here [15], the *solutions* of the ODEs describing the noise are plugged into the ODEs describing the protein dynamics. This leads to an additional memory effect. That is:

$$\dot{X} = N_X + F(X) + X/\tau, \qquad (5.14)$$

with $F(X)$ some arbitrary function of X. Note that the N_X function also contains a τ term (see Eq. 5.13), which is effectively integrated, thus leading to effects of the fluctuations much longer timescales than τ. This effect is (partially) countered by the third term in Eq. 5.14, which also contains the τ term.

KIVIET ET AL. MODEL

As mentioned, the Kiviet et al. model takes a different approach. The formulae that follow correspond to Fig. 5.9.B (Fig. 5.9.A contains the version of the cartoon which was published in Kiviet et al.). The starting point,

$$\dot{C} = p - \lambda C, \qquad (5.15)$$

is the same in my model, but after linearization (defined as $X = X_0 + \delta X$) this leads to only one ODE:

$$\frac{\delta \dot{C}}{C_0 \lambda_0} + \frac{\delta C}{C_0} = \frac{\delta p}{C_0 \lambda_0} + T_{C \leftarrow \lambda} \frac{\delta \lambda}{\lambda_0}. \qquad (5.16)$$

Noise terms are introduced with ODEs that are also shown above in Eq. 5.13. Additionally, two functions are defined for p and λ. These are not ODEs, as the effects on these parameters are thought to happen on fast timescales. The parameters are however linearised around X_0 (and thus written as δX). The equation

$$\frac{\delta \lambda}{\lambda_0} = T_{\lambda \leftarrow C} \frac{\delta C}{C_0} + T_{\lambda \leftarrow G} N_G + N_\lambda \qquad (5.17)$$

simply defines the evolution of $\delta \lambda$. There is also a similar equation for δp:

$$\frac{\delta p}{C_0 \lambda_0} = T_{C \leftarrow C} \frac{\delta C}{C_0} + T_{C \leftarrow G} N_G + N_E, \qquad (5.18)$$

which plays a bit more complicated role. It is defined using terms that pertain to C, like $T_{C \leftarrow G}$, such that it can be directly plugged into Eq. 5.16. Indeed, plugging Eq. 5.18 into Eq. 5.16 leads to equation:

$$\frac{\delta \dot{C}}{C_0 \lambda_0} + \frac{\delta C}{C_0} = \left[T_{C \leftarrow C} \frac{\delta C}{C_0} + T_{C \leftarrow G} N_G + N_E \right] + T_{C \leftarrow \lambda} \frac{\delta \lambda}{\lambda_0} \qquad (5.19)$$

which corresponds to equation 5 in the Kiviet et al. [10] manuscript. I say it is a bit complicated, since Eq. 5.18 has no role in the model (we could also just have defined Eq. 5.19 immediately), except that it shows us which part of the model can be interpreted as being the production rate. This is also the reason why p is depicted as a red dashed box in Fig. 5.9.

Note that dampening terms can be implicitly present in the Nghe model, in the form of a transmission from the parameter to itself. Specifically, $T_{C \leftarrow C}$ can fulfill this role. Second order dampening can also occur, as is pointed out in the Kiviet et al. supplementary information, which states that the time scale of the C fluctuations is set by the term $\lambda_0(1 - T_{C \leftarrow \lambda} T_{\lambda \leftarrow C} - T_{C \leftarrow C})$.

5

III. THE TOWBIN ET AL. MODEL EQUATIONS

THE EQUATIONS

This supplementary note offers a brief overview of the formulae used in Towbin et al. to describe the CRP regulation system, and its relation to metabolite concentrations and growth. This is meant as a reference, to allow comparison between the linearized noise model that was used in this chapter to model the dynamics of metabolism and growth (described in supplementary note II), and a more sophisticated model of the biochemical network that is aimed to model the population average response to changes in the extracellular environment. These equations are directly taken from the supplementary notes of Towbin et al. (Equations 7-9), except that some parameter symbols have been changed to match our notation [153]. The equations are the following:

$$\dot{C_M} = \lambda(\frac{k_f}{k_f + x} - C_M), \tag{5.20}$$

$$\dot{x} = \frac{\alpha}{k_2}\left(\beta f_M(C_M)\frac{k_1}{x + k_1} - \lambda\right), \tag{5.21}$$

$$\lambda = \gamma(1 - C_M)\frac{x}{x + k_2}. \tag{5.22}$$

Here, C_M represents the metabolic enzyme expression, x the pool of carbon metabolites which give feedback to the CRP system, and λ the cellular growth rate[4]. Furthermore, k_1, k_2, and k_f are Michaelis Menten rate constants, where k_1 pertains to carbon import (which is self-inhibited), k_2 to biomass production and k_3 to feedback by metabolites. α, β and γ are also constants, α sets the conversion ratio between metabolite consumption and growth, β the maximum import rate of metabolites, and γ the maximum catalytic rate of ribosomes. $f_M(C_M)$ is a function of C_M, which describes 'the regulation of the limiting enzyme for carbon uptake and catabolism by CRP' [153]. The left and right terms (within brackets) in Eq. 5.20 relate to production and dilution and in Eq. 5.21 the left and right terms in brackets relate to import and consumption. The growth rate is a function of metabolite consumption and ribosome concentration $C_R = 1 - C_M$.

GRAPHIC REPRESENTATIONS OF THE MODEL

The relationships that are set in the differential equations are presented in a diagram in figure 5.10. The contributions of the left and right terms can be plotted separately to understand the dynamic behavior of the system when these terms change independently, see figure 5.11.

---

[4]In this chapter, we use both the symbols λ and μ to indicate growth rate, λ when the units are /min and μ when the units are doublings/hour.
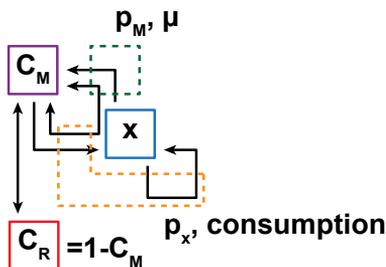
Figure 5.10: Diagram of Towbin et al. ODE model. This diagram shows how the parameters in the Towbin et al. model relate to each other. C_M and x are the metabolic enzyme and metabolite concentration respectively, which are both modeled using differential equations. The ribosome concentration is only implicitly present as $C_R = 1 - C_M$. Metabolic enzyme production p_m, growth rate λ, metabolite production/import rate p_x and metabolite consumption are all terms that are part of these differential equations.
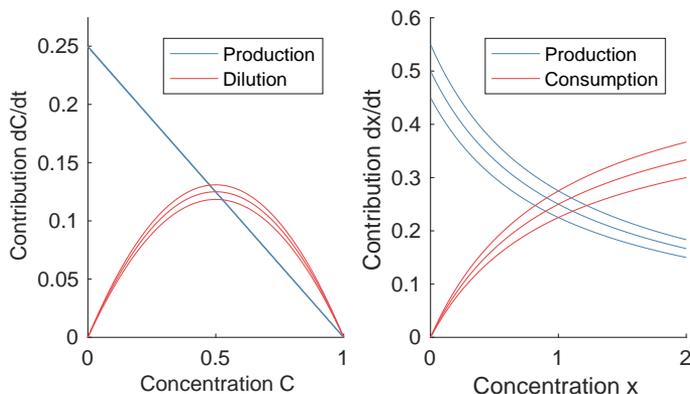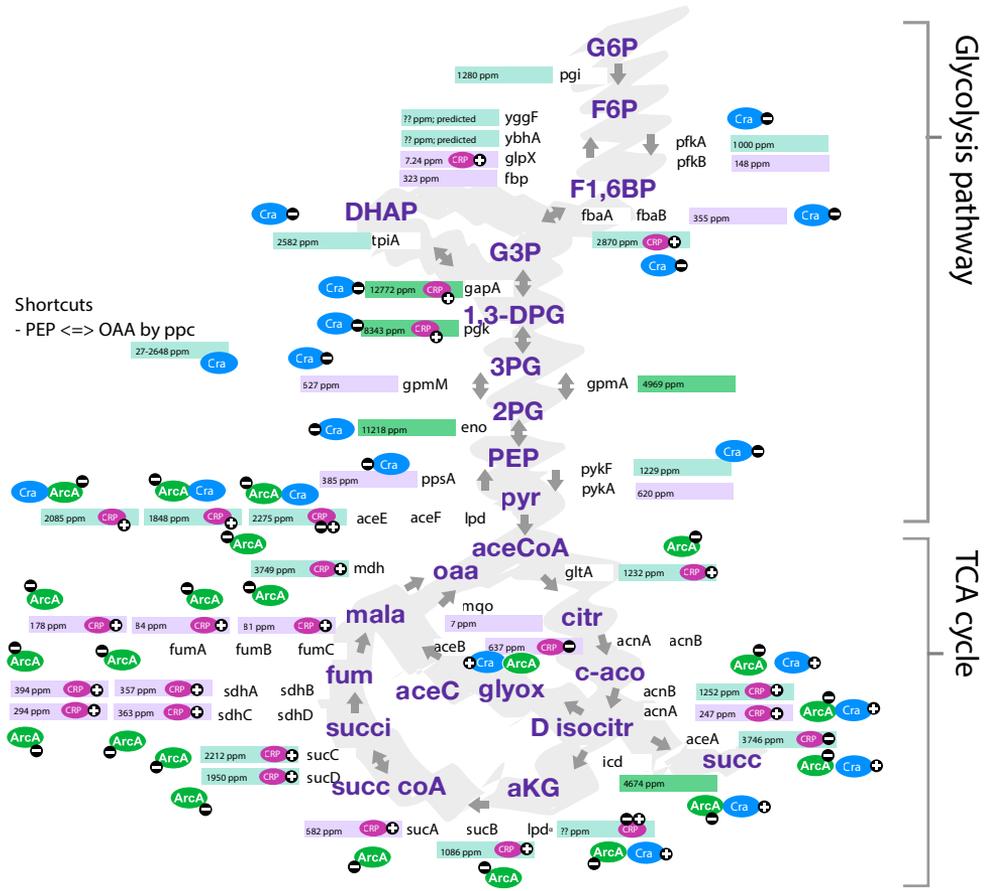

Figure 5.11: Using models to understand the dynamic behavior of stochastic metabolic and growth fluctuations. (Left) Illustrations of the contributions from the positive (production) and negative (dilution) terms to the differential equation that describes the time evolution of the concentration of metabolic enzymes as a function of that concentration C, according to the Towbin et al. model [153]. The different lines correspond to different values of the metabolite concentration x ($x = 0.9, 1.0, 1.1$). (Right) Similar as on the left, except that the terms are production and consumption of metabolites as a function of metabolite concentration x and that the different lines correspond to different values of the metabolic enzyme concentration C ($C = 0.45, 0.50, 0.55$). (Towbin et al. focus on import of metabolites as rate limiting step in the rate at which they are created, but we have used the more general term production here.) These plots are merely to illustrate the systems behavior, so all parameter values were simply set to 1.

SUPPLEMENTARY FIGURES

5

5

Glycolysis pathway

TCA cycle

G6P

1280 ppm | pgi

F6P

?? ppm; predicted | yggF
?? ppm; predicted | ybhA
7.24 ppm CRP + | glpX
323 ppm | fbp

Cra −
1000 ppm
148 ppm
pfkA
pfkB

F1,6BP

fbaA | fbaB
355 ppm | Cra −
2870 ppm CRP +
Cra −

Cra − DHAP

2582 ppm | tpiA

G3P

Cra − 12772 ppm CRP + | gapA
Cra − 8343 ppm CRP + | pgk

1,3-DPG

Cra −

527 ppm | gpmM

3PG

gpmA 4969 ppm

Cra − 11218 ppm | eno

2PG

PEP

Cra −
Cra − 385 ppm | ppsA
pykF | 1229 ppm Cra −
pykA | 620 ppm

pyr

Shortcuts
- PEP <=> OAA by ppc
27-2648 ppm Cra

Cra − ArcA −
2085 ppm CRP +
ArcA − Cra −
1848 ppm CRP +
ArcA − Cra −
2275 ppm CRP +
aceE | aceF | lpd
ArcA −

aceCoA

3749 ppm CRP + | mdh
oaa
gltA
ArcA −
1232 ppm CRP +

ArcA −
178 ppm CRP +
ArcA −
84 ppm CRP +
ArcA −
81 ppm CRP +
mala
mqo
7 ppm
citr

fumA | fumB | fumC
aceB
acnA | acnB

ArcA − | ArcA −
394 ppm CRP + | 357 ppm CRP +
294 ppm CRP + | 363 ppm CRP +
sdhA | sdhB
sdhC | sdhD

fum
aceC | glyox
c-aco
acnB
acnA
637 ppm CRP + Cra + ArcA −

ArcA − Cra +
1252 ppm CRP +
247 ppm CRP + ArcA − Cra +

ArcA − | ArcA −
succ
D isocitr
aceA | 3746 ppm CRP −
ArcA − Cra +

ArcA −
2212 ppm CRP +
1950 ppm CRP +
sucC
sucD
succi
icd
succ
ArcA −

Succ coA
aKG
4674 ppm
ArcA − Cra +

582 ppm CRP + | sucA | sucB | lpd− | ?? ppm CRP +
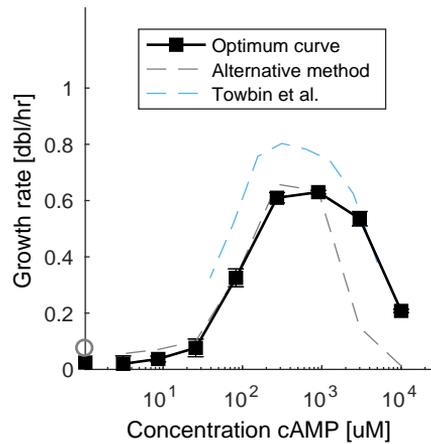1086 ppm CRP +
ArcA −
ArcA − Cra +
ArcA −

Figure 5.13: Optimum curve based on bulk measurements. The black line with solid black squares indicates exponential phase growth rates of ΔcAMP cells at different concentrations of cAMP as measured in a platereader that measured bacterial density over time. The growth rate was determined by an exponential fit, based on a manually selected part of the bacterial density curve that displayed exponential growth. We also used an alternative method to determine growth rates (gray dashed line), were we fitted an exponential to part of the bacterial density curve that fell between two threshold values. We consider this less reliable, as growth is not guaranteed to be exponential. We also show the optimum curve as measured by Towbin et al [153] (dashed blue line), who calculated growth rate based on an exponential fit to a two hour window, which was selected by an algorithm that looked for the longest period of stable growth rate.

Figure 5.12: Regulation of the TCA cycle and ED pathway. Some regulatory proteins control the expression of many tricarboxylic acid (TCA) cycle and/or enzymes glycolysis pathway simultaneously. This diagram shows the regulatory effects of CRP, but also of the Catabolite repressor activator (Cra) and the Anoxic redox control A (ArcA) protein. Cra controls the direction of the metabolite flux through metabolic pathways [171, 202]. Cra is activated by the metabolite fructose-1,6-bisphosphate [203]. ArcA is part of a two component system (ArcAB) which controls gene expression in response to aerobic versus anaerobic conditions [171, 204]. Metabolites are displayed in larger font, enzymes that catalyze reactions in smaller fonts. In the colored boxes, regulation by CRP is displayed by an oval symbol marked "CRP". Similarly, when an enzyme is regulated by Cra or ArcA this is displayed next to the boxes; a plus or minus symbol indicates positive or negative regulation. Additionaly, in the boxes this diagram displays the concentration of particular enzymes in ppm, as annotated in PaxDb (retrieved in 2015) [205]. The boxes are also color-coded according to abundance of the enzymes. The diagram is based on EcoCyc [171]. From top to bottom, clockwise, enzymes abbreviations stand for glucose 6-phosphate, fructose 6-phosphate, fructose 1,6-biphosphate, glyceraldehyde 3-phosphate, 1,3-biphospho-D-glycerate, 3-phospho-glycerate, 2-phospho-glycerate, phosphoenolpyruvate, pyruvate, acetyl coenzyme A, citrate, cis-aconitate, d-threo-isocitrate, (succinate, glyoxalate, acetyl-CoA), 2-oxo-glutarate, succinyl-CoA, succinate, fumarate, malate, oxaloacetate.
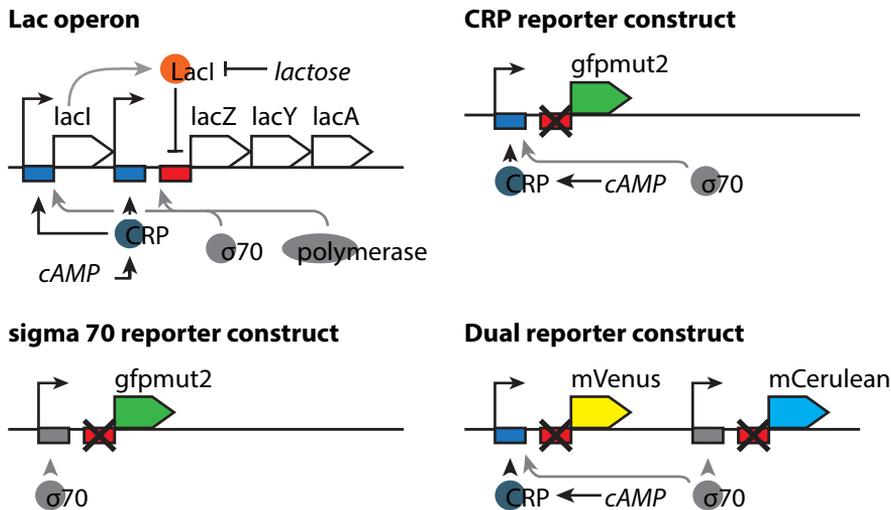
Figure 5.14: The reporters used in this study. (Top left panel) This panel shows the lac operon, on which these promoters are based. The lac operon consists of three genes, *lacZ*, *lacY* and *lacA*, that are required to metabolize lactose. These three genes are regulated by the same promoter, which is activated by CRP. When there is no lactose, transcriptional repressor lacI blocks gene expression from the lac operon. The *lacI* gene is encoded at a location near to the lac operon, and is also displayed (it is also regulated by CRP). When there is lactose, transcriptional repressor lacI is inhibited by lactose and LacZ, LacY and LacA are produced. CRP recruits initiation factor sigma 70, which is followed by polymerase binding and transcription. (It is furthermore also known that pleiotropic transcription factor H-NS can bind to the lac operon and repress gene expression.) (Top right and bottom left panel) Using the lac promoter as a starting point, Towbin et al. created two reporters [153]. Both posses mutations which result in the removal of the LacI binding site (indicated with a black cross). The CRP sensitive promoter is then created by fusing the promoter to a GFP fluorescent protein sequence. We call this the CRP reporter or metabolic reporter. In the promoter of the second reporter, which is intented to measure background fluctuations in gene expression, also mutations have been introduced that remove the CRP binding site. Instead, a consensus binding site for the sigma 70 initiation factor is introduced, effectively making it a constitutive promoter that is otherwise similar to the promoter used for the CRP reporter. It was also fused to GFP, and we call this the sigma 70 reporter or constitutive reporter. (Bottom right panel) To be able to perform single cell measurements that involve both reporters, we replaced the GFP sequences of the metabolic and constitutive reporters with mVenus and mCerulean sequences respectively. The GFP reporters were introduced to cells using plasmids as vector, whilst the mVenus and mCerulean reporters were chromosomally inserted at the *intC* and *galK* locations respectively.
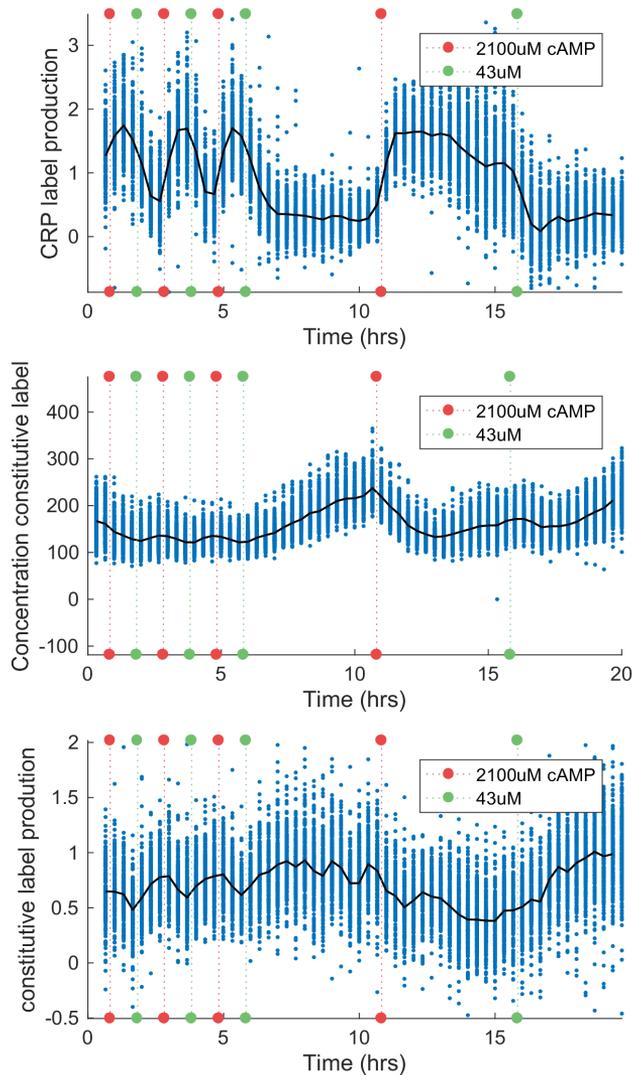
Figure 5.15: Additional time traces from the cAMP pulsing experiment. This figure displays additional data for figure 5.2. It shows time traces for parameters measured in a population of ΔcAMP cells (blue dots are single cell measurements, the black line is the population average) that were grown alternately in minimal medium supplemented with 43 μM cAMP and minimal medium supplemented with 2100 μM cAMP. Red and green dotted lines indicate times where the concentration was switched, as indicated in the legend. For this experiment, also the production rate of the metabolic reporter was determined (top panel). Furthermore, also the production rate and concentration of a constitutive reporter where determined (bottom two panels), see main text for more on these quantities.

5
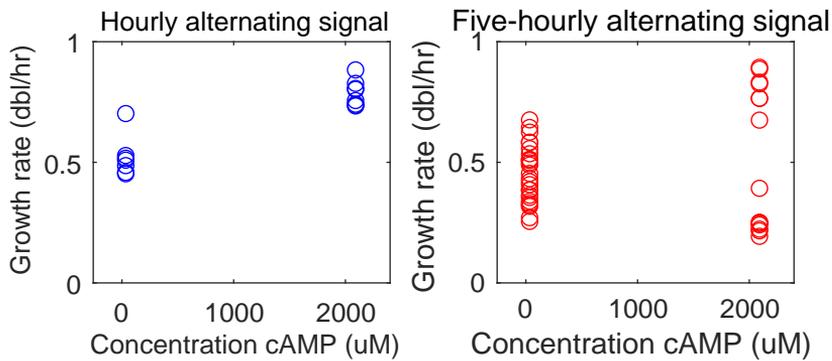
Figure 5.16: Growth rates during the pulsing experiment against cAMP concentration. This plot shows the population average growth rate of ΔcAMP cells during the pulsing experiment (see main text). In the left panel the growth rates are shown from the regime where the cAMP concentration in the medium was switched every hour between low and high concentrations (43 and 2100 μM respectively), and in the right panel data is shown from the regime where the cAMP concentration switched every five hours. Note that these cells are not in steady state. Growth rates are not only determined by the cAMP concentration, but also by the history of the population.
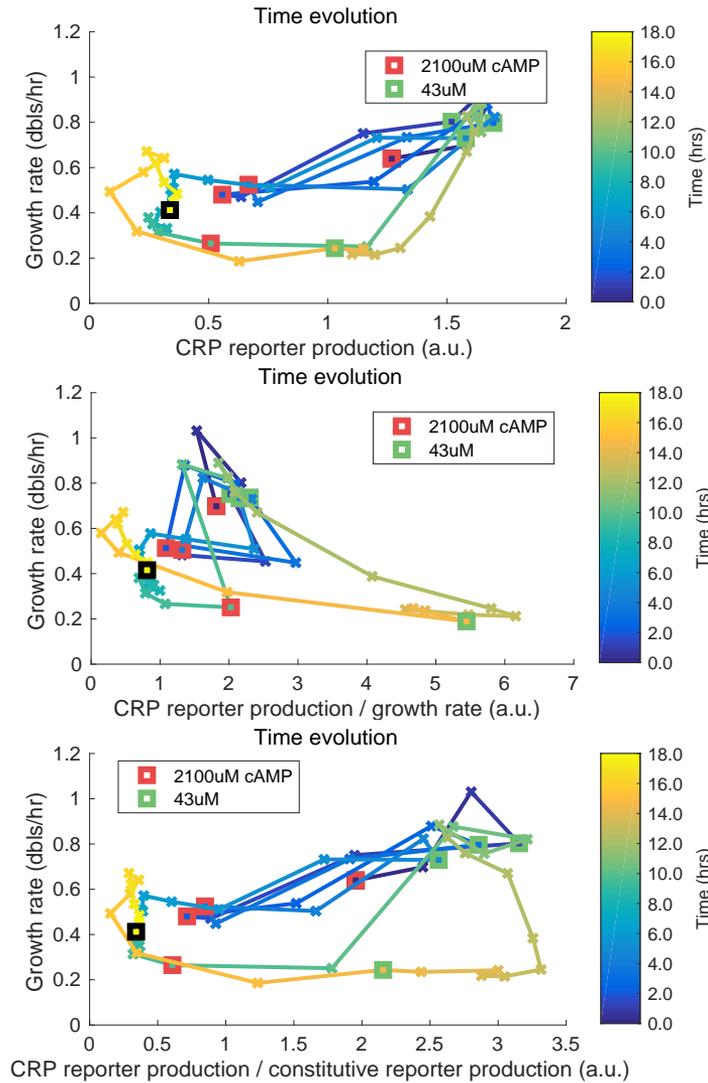
Figure 5.17: Time evolution of production versus growth rate during block pulses of high and low cAMP concentrations. (Top) The production rate of the CRP reporter plotted against growth rate, color coded for time. Production is determined here by the difference in absolute signal between two timepoints, divided by the amount of time inbetween these two points and divided by the area of the cell. Note that the production rate might also depends on the growth rate, as does the final concentration of the protein. This makes this data not trivial to interpret. (Middle) Similar to top panel, but production rate is divided by growth rate, since production rates might be more faithfully represented as fractions of the growth rate. (Bottom) Similar to top panel, but production rate is divided by the production rate of the constitutively expressed label, which serves a reference for the total production rate capacity of the cell.
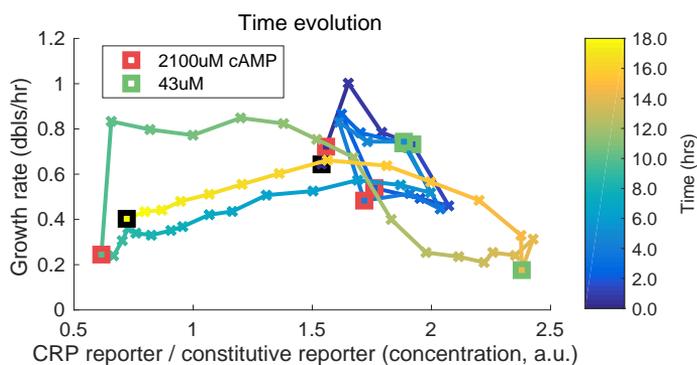
Figure 5.18: Time evolution of metabolic reporter concentration normalized by constitutive reporter concentration. This figure is similar to figure 5.17, but shows the concentration of the metabolic reporter divided by the constitutive reporter on the x-axis, instead of production rates.
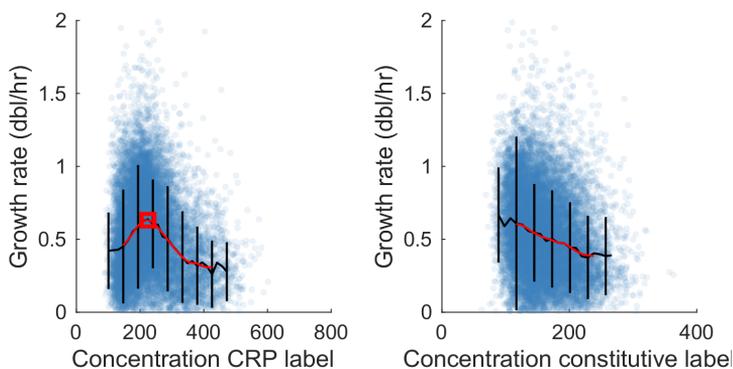
5

Figure 5.19: Scatter plots of growth rate versus metabolic and constitutive reporter from single cell measurements. These plots relate to the pulsing experiment with ΔcAMP cells, figure 5.2 in the main text. Every point in these plots corresponds to a single cell observation on growth rate and fluorescent label concentration that was made during a time series of pulses of high and low concentrations of cAMP. On the left, growth rate plotted against the concentration of the metabolic reporter. On the right, growth rate plotted against the concentration of a constitutive reporter based on the same lac metabolic reporter but with the CRP binding site replaced by a σ70 consensus binding site. The black lines indicate the average signal for corresponding concentrations and standard deviations. The red lines are interpolated 2nd degree polynomial fits of the displayed mean values. The red square shows the maximum value of the interpolated values at a metabolic reporter concentration of 225 a.u. (corresponding to a growth rate of 0.63 doublings/hr).
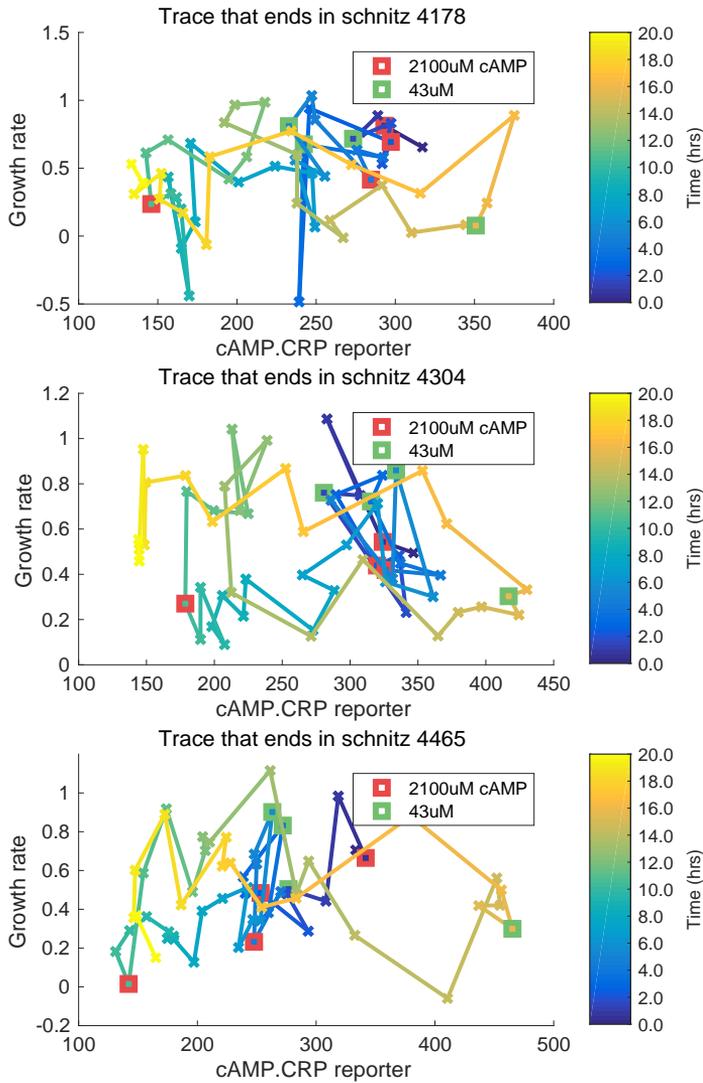
Figure 5.20: Time evolution of growth rate and metabolic reporter concentration for single cells.
Each panel is identical to main figure 5.2.C, except that each panel corresponds to the behavior of
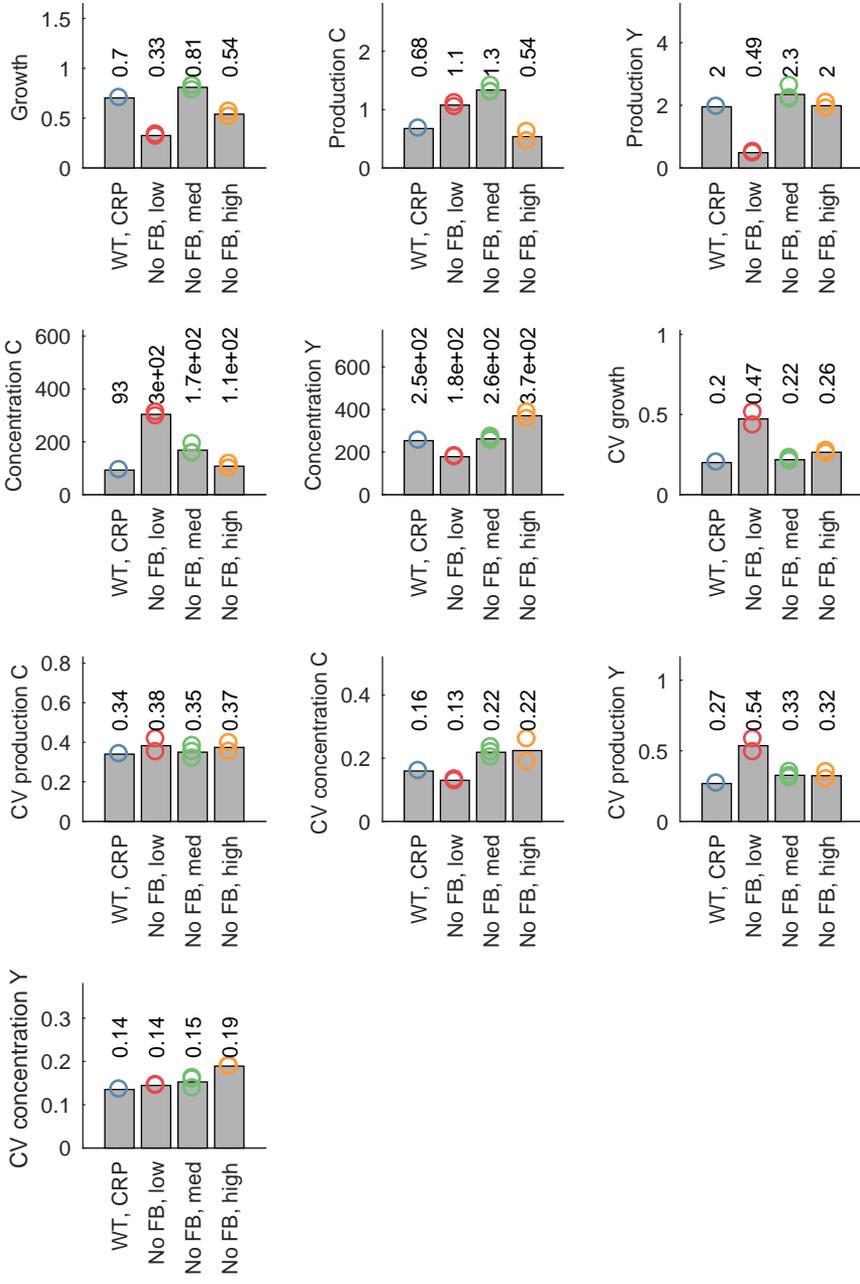a lineage of single cells.

5

Figure 5.21: Comparison of population averages of different parameters in different experimental conditions. This figure shows growth rates (doublings per hour), production rates (a.u./(px min)), concentrations (a.u./px) and coefficients of variations (CV). C is the abbreviation used for the constitutive reporter (which uses a cyan mCerulean label), and Y is the abbreviation used for the metabolic reporter (which uses a yellow mVenus label). WT stands for wild type, indicating wild type cells except the addition of our metabolic and constitutive reporter constructs. No FB stands for no feedback, indicating the ΔcAMP cells that also carry the metabolic and constitutive reporter constructs. Low, med and high correspond to extracellularly provided cAMP concentrations of 80, 800 and 5000 µM cAMP respectively. Bars are averages over multiple experiments, and each open circle corresponds to a value observed in one experiment.
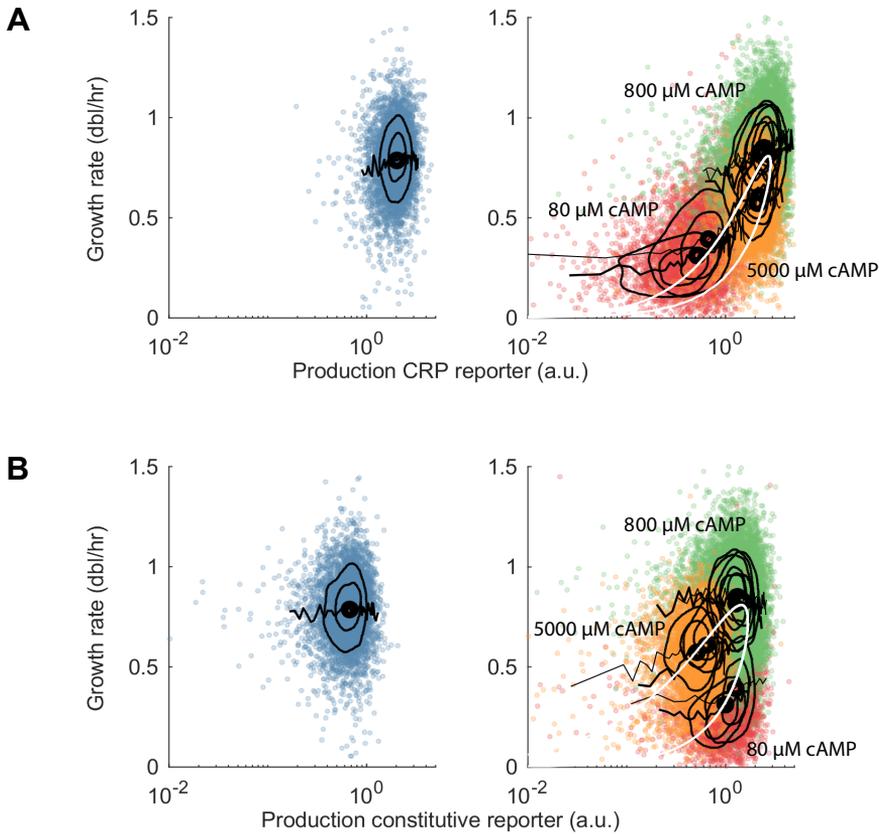
5

Figure 5.22: Scatter plots of production rates against growth rate. (A) Colored dots show single cell growth rate values plotted against respective single cell production rates of metabolic reporter, which is a proxy for the production rate of metabolic enzymes. Dots correspond to data points from WT cells with endogenous negative metabolic feedback, and ΔcAMP cells which have been grown at concentrations of 80, 800 and 5000 μM cAMP (colored blue, and red, green and orange, respectively). The white lines show predicted population average behavior based on a model described in supplementary note I. The black lines show the average growth rate for cells that are binned according to production rates, and the black isolines reflect kernel density estimates of the probability distribution (using the matlab function `kde2d` [50]). The circles with thick black edges show population average values per experiment. (B) As panel A, except that this panel shows the relationship between growth and the production rate of a constitutive reporter.
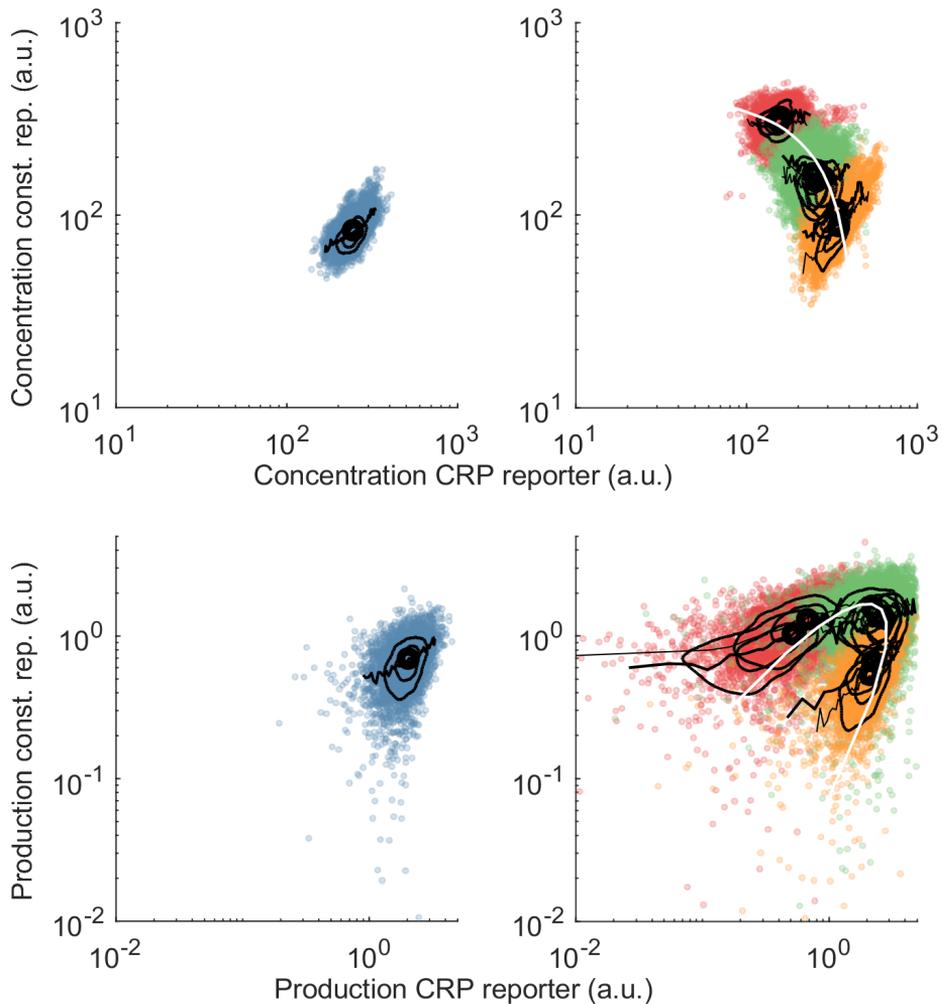
Figure 5.23: Concentration-concentration and production-production relationships between metabolic and constitutive reporters. This figure is similar to Figure 5.22, except that the relationships plotted are different. Blue, red, green and orange dots again correspond respectively to data points from WT cells with endogenous negative metabolic feedback, and ΔcAMP cells which have been grown at concentrations of 80, 800 and 5000 μM cAMP. (Top panel) Scatter plots for concentration of metabolic (CRP) reporter against the concentration of constitutive reporter. The white line reflects the fact that the sum of the concentrations of the two reporters remains constant (not that the axes are log-log scale; this line would be straight on a linear scale). (Bottom panel) Scatter plots for production rate of metabolic reporter against the production rate of constitutive reporter. The white line is based on the optimum curve and a simple model describing the relationship between production rate, concentration and growth, see supplementary note I.
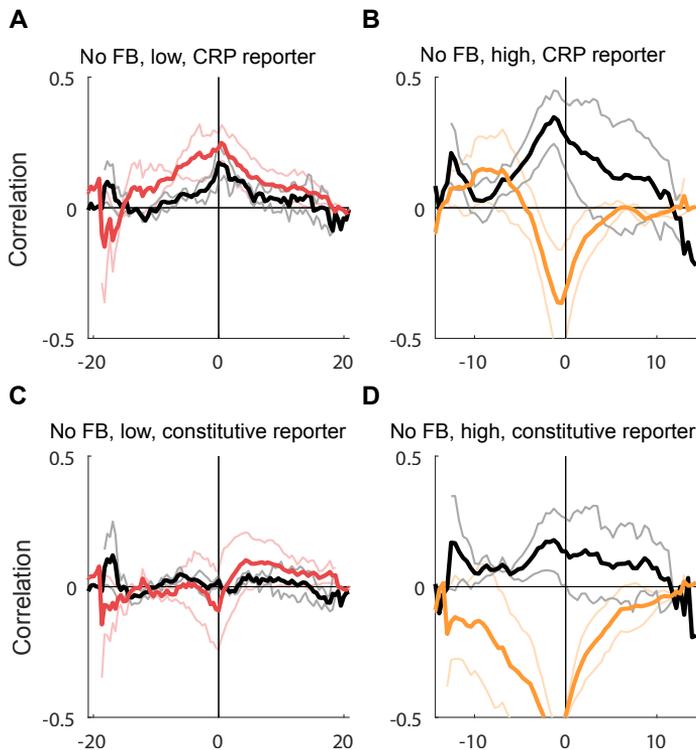
5

Figure 5.24: Cross-correlations for non-optimal cAMP expression levels. Also for the non-optimal concentrations of cAMP, 80 μM ("low") and 5000 μM ("high") respectively, the cross-correlations were determined both for the metabolic and constitutive reporters. Colored lines indicate concentration-growth correlations, whereas black lines indicate production-growth correlations. The shaded lines are experiments from independent colonies, whereas the dark lines indicate the average over those experiments. (A) At lower than optimal cAMP concentrations, we see a similar pattern for the metabolic reporter as at cAMP concentrations of 800 μM (figure 5.4.B), namely positive correlations between growth and expression of metabolic proteins. (B) At higher than optimal cAMP concentrations, we see a slightly different pattern for the metabolic reporter than in the 800 μM case; there is still a positive correlation between protein production and growth rate, but the concentration seems more typical of the dilution mode. (C) The cross-correlations for the constitutive reporter also seem different from the case with 800 μM cAMP (figure 5.4.D), although there is quite some discrepancy between the two experiments shown (shaded lines), and hence more experiments might be needed to elucidate these cross-correlations further. (D) The cross-correlations for the constitutive reporter at 5000 μ cAMP also seem different from the case with 800 μM cAMP (figure 5.4.D), although there is again quite some discrepancy between the two experiments shown (shaded lines).

ADDITIONAL SUPPLEMENTAL FIGURES

This section contains additional supplemental figures that were not referenced in the chapter.
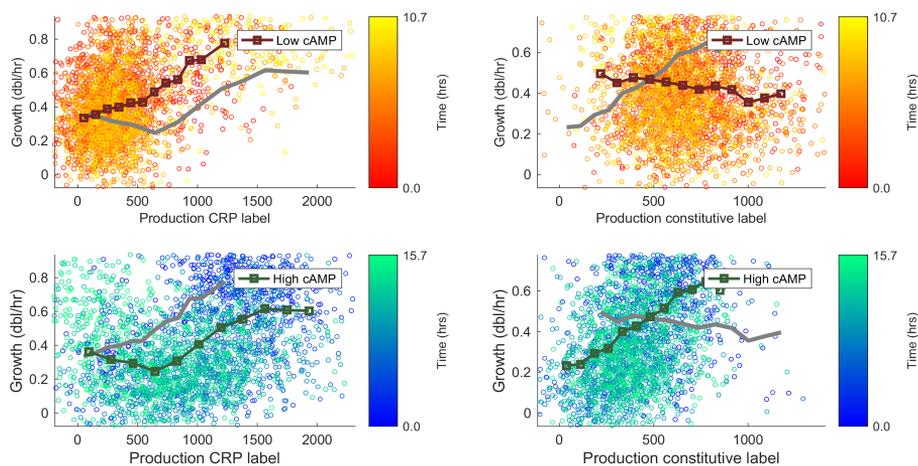
5

Figure 5.25: Scatter plots for growth versus production, with selection on the high/low condition. These plots relate to the pulsing experiment with ΔcAMP cells, figure 5.2 in the main text. Every point in these plots corresponds to a single cell observation on growth rate and fluorescent label concentration that was made during a time series of pulses of high and low concentrations of cAMP. In these scatter plots, the production of the metabolic (left plots) or constitutive (right plots) reporter is plotted against growth rate, where in the top plots only data from the low cAMP condition is shown and in the bottom plot only data from the high cAMP condition is shown. The gray lines are the exception, they show the average values from the high condition in the plot for the low condition, and vice versa.
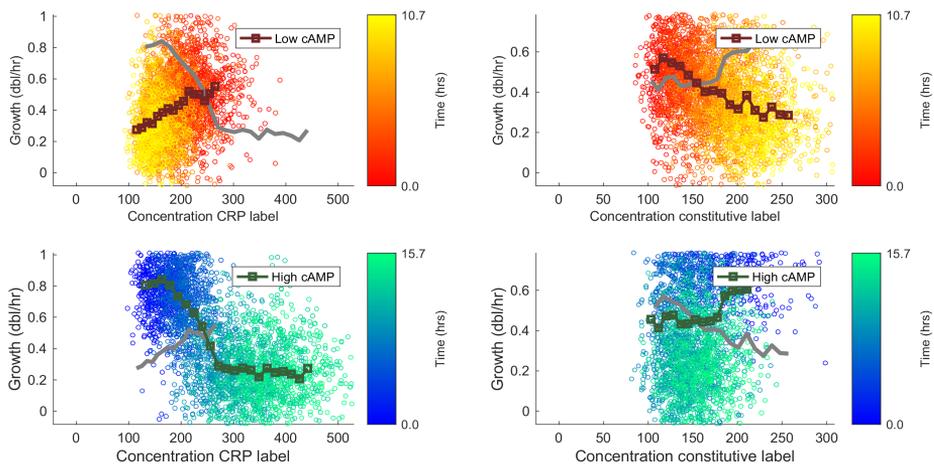
Figure 5.26: Scatter plots for growth versus concentration, with selection on the high/low condition. This figure is similar to supplemental figure 5.25, but deals with concentration instead of production rates. These plots relate to the pulsing experiment with ΔcAMP cells, figure 5.2 in the main text. Every point in these plots corresponds to a single cell observation on growth rate and fluorescent label concentration that was made during a time series of pulses of high and low concentrations of cAMP. In these scatter plots, the concentration of the metabolic (left plots) or constitutive (right plots) reporter is plotted against growth rate, where in the top plots only data from the low cAMP condition is shown and in the bottom plot only data from the high cAMP condition is shown. The gray lines are the exception, they show the average values from the high condition in the plot for the low condition, and vice versa.

5
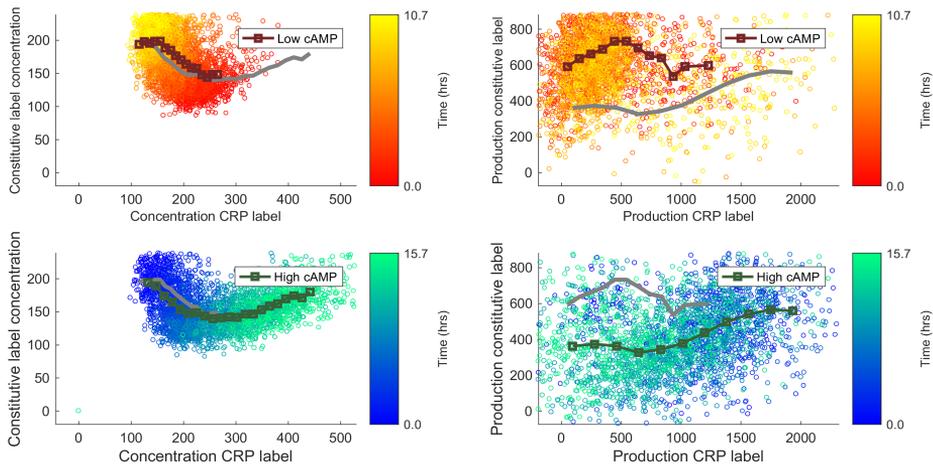
Figure 5.27: Scatter plots for metabolic reporter versus constitutive reporter, with selection on the high/low condition. This figure is similar to figures 5.25 and 5.26. These plots relate to the pulsing experiment with ΔcAMP cells, figure 5.2 in the main text. Every point in these plots corresponds to a single cell observation on growth rate and fluorescent label concentration that was made during a time series of pulses of high and low concentrations of cAMP. In these plots, concentration-concentration (left) or production-production (right) scatter plots are presented for the metabolic versus the constitutive reporters. In in the top plots only data from the low cAMP condition is shown and in the bottom plot only data from the high cAMP condition is shown. The gray lines are the exception, they show the average values from the high condition in the plot for the low condition, and vice versa.
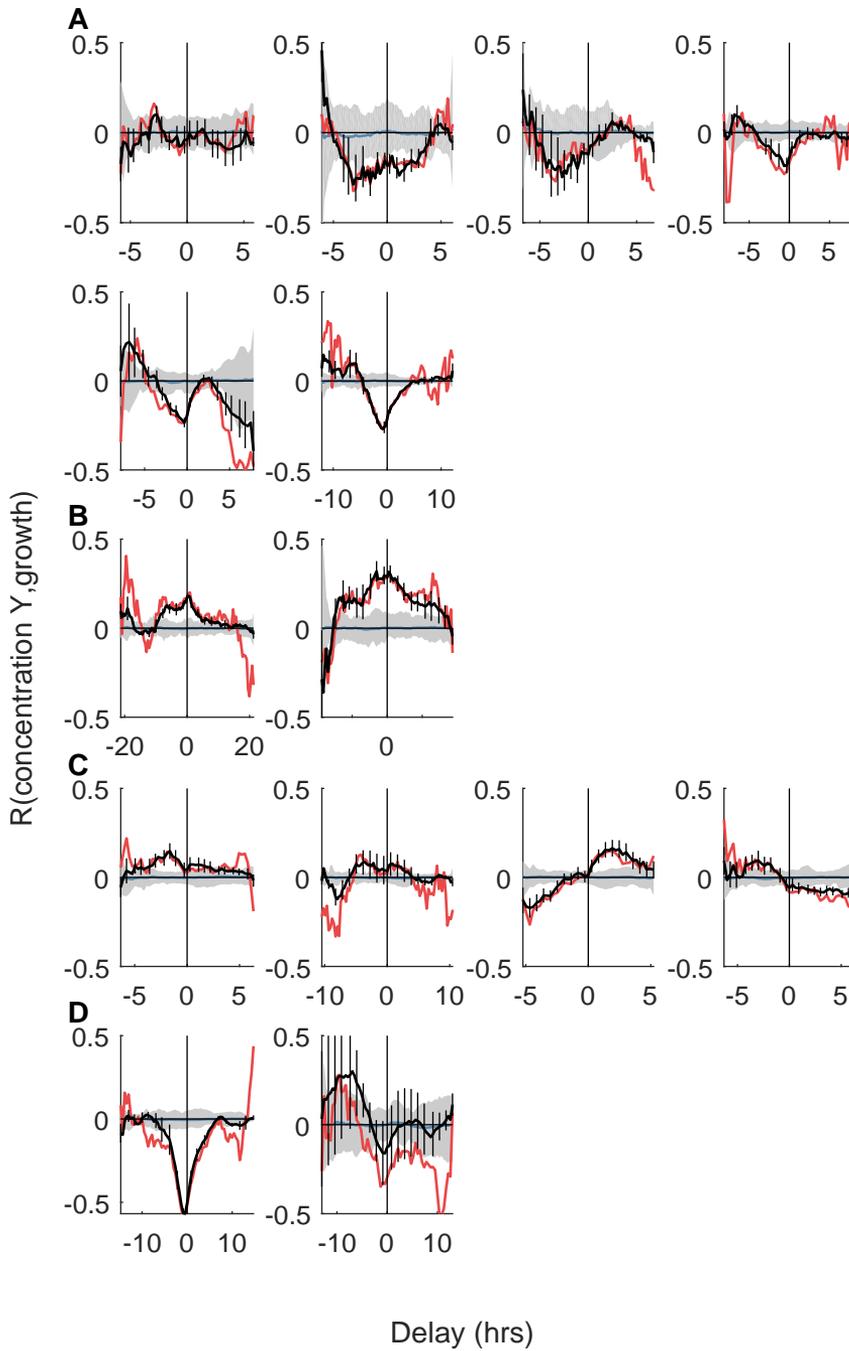
Figure 5.28:   Metabolic concentration-growth cross-correlations $R_{M,\mu}(\tau)$ **per experiment.** These graphs show cross-correlations (CCs) per condition, per experiment. They display not only the overall cross-correlation (black), but also the cross-correlation based on the scatter plots (red) and the control (gray area); see chapter 2 and also figure 2.10, for more information. The error bars in the black curves are SEM, based on dividing data from a single experiment into four groups, and calculating four CCs, for which the SEM is calculated. The black CCs here correspond to the shaded CCs that are shown in other figures where multiple experiments are combined into one plot. (A) $R_{M,\mu}(\tau)$ for the wild type cells. (B) $R_{M,\mu}(\tau)$ for the ΔcAMP cells, with 80 μM cAMP supplemented to the growth medium. (C) $R_{M,\mu}(\tau)$ for the ΔcAMP cells, with 800 μM cAMP supplemented to the growth medium. (D) $R_{M,\mu}(\tau)$ for the ΔcAMP cells, with 5000 μM cAMP supplemented to the growth medium. (In the y-axis label, "Y" stands for the yellow fluorescent reporter concentration that was tracked here.)
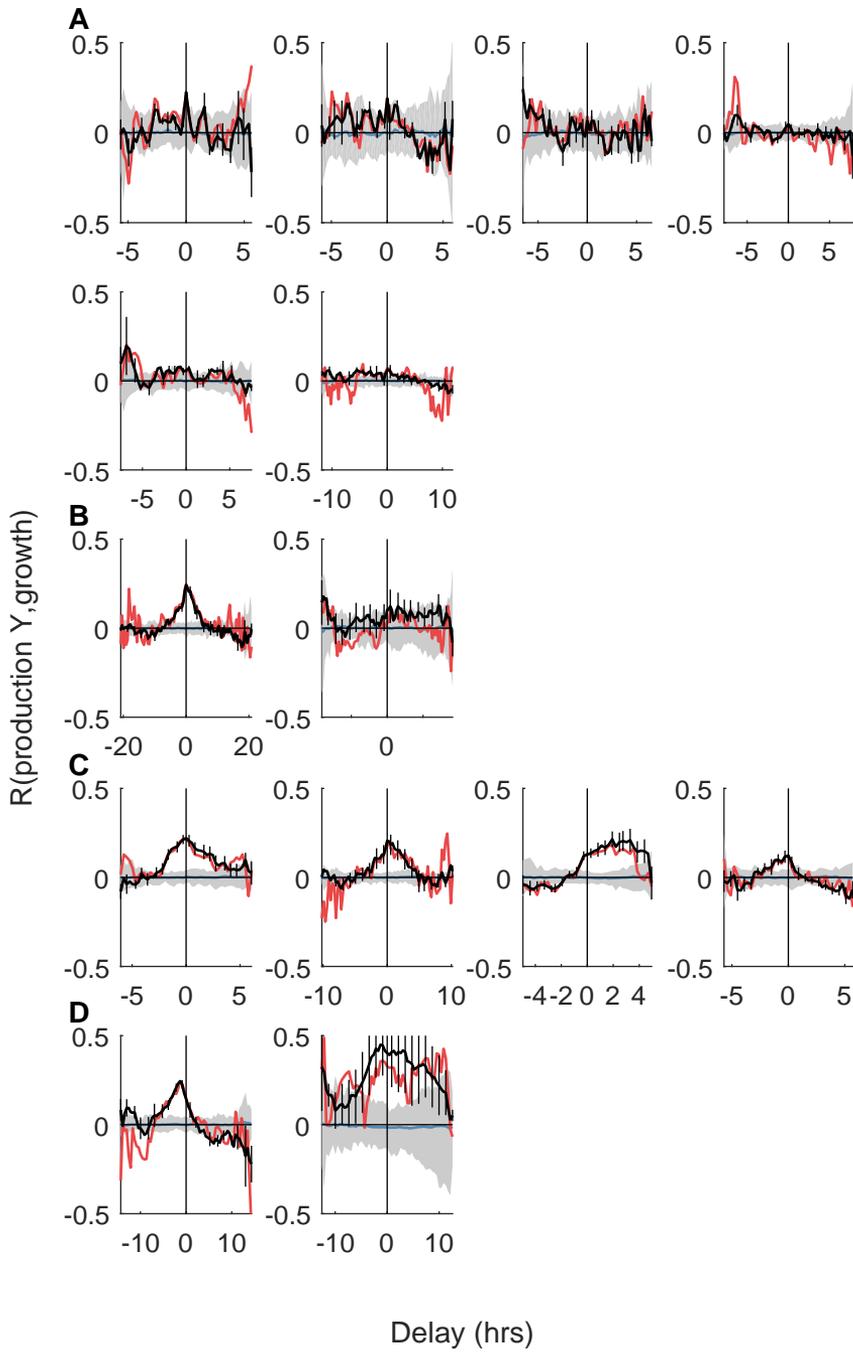
5

5

Figure 5.29: Metabolic production-growth cross-correlations $R_{p_M,\mu}(\tau)$ **per experiment.** This figure is similar to figure 5.28, except that production-growth correlations are shown here, instead of concentration-growth correlations. These graphs show cross-correlations (CCs) per condition, per experiment. They display not only the overall cross-correlation (black), but also the cross-correlation based on the scatter plots (red) and the control (gray area); see chapter 2 and also figure 2.10, for more information. The error bars in the black curves are SEM, based on dividing data from a single experiment into four groups, and calculating four CCs, for which the SEM is calculated. The black CCs here correspond to the shaded CCs that are shown in other figures where multiple experiments are combined into one plot. (A) $R_{p_M,\mu}(\tau)$ for the wild type cells. (B) $R_{p_M,\mu}(\tau)$ for the ΔcAMP cells, with 80 μM cAMP supplemented to the growth medium. (C) $R_{p_M,\mu}(\tau)$ for the ΔcAMP cells, with 800 μM cAMP supplemented to the growth medium. (D) $R_{p_M,\mu}(\tau)$ for the ΔcAMP cells, with 5000 μM cAMP supplemented to the growth medium. (In the y-axis label, "Y" stands for the yellow fluorescent reporter production that was tracked here.)
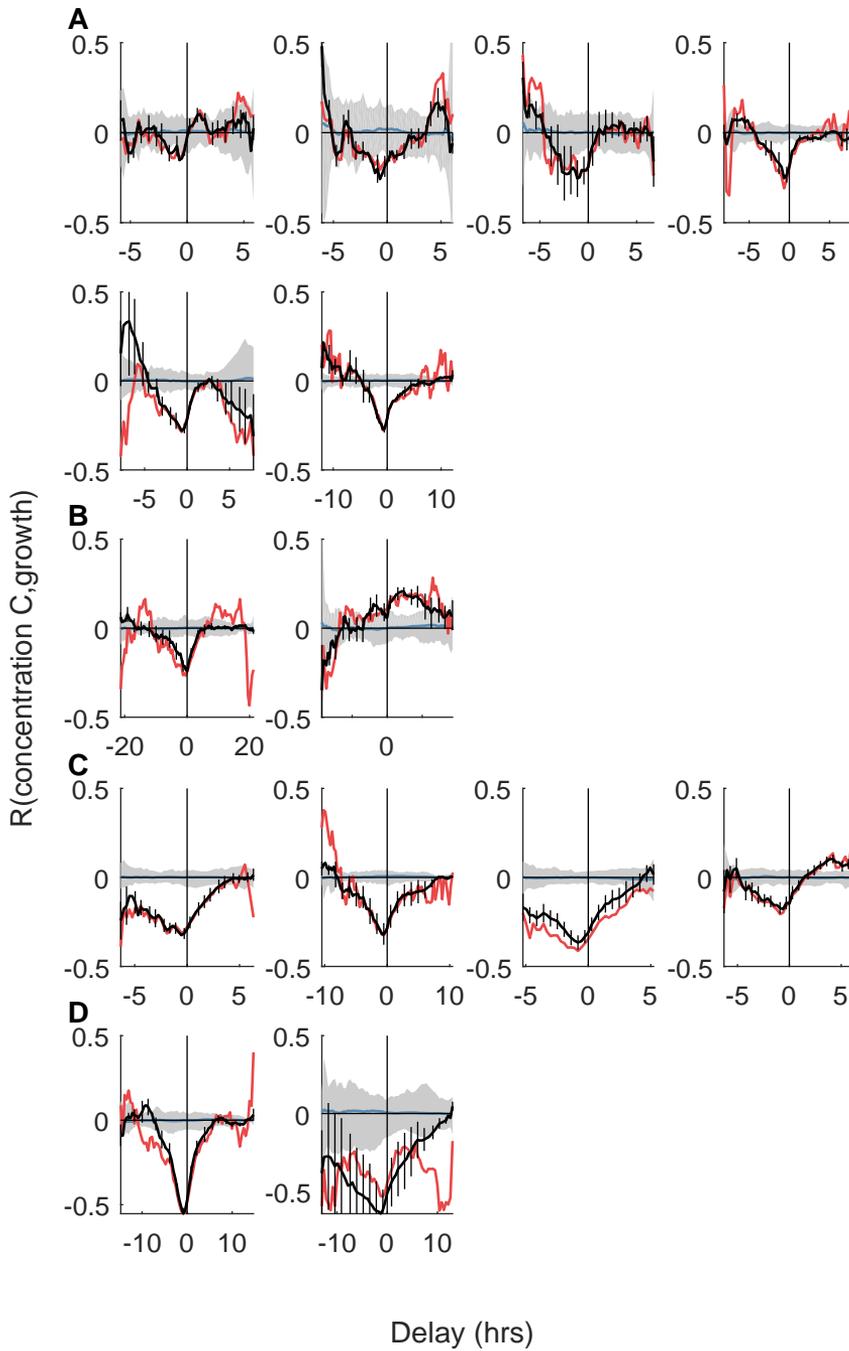
5

R(concentration C,growth)

Delay (hrs)

5

Figure 5.30:   Constitutive concentration-growth cross-correlations $R_{Q,\mu}(\tau)$ per experiment. This graph is similar to supplemental figure 5.28, except that it relates to constitutive reporter measurements. These graphs again show cross-correlations (CCs) per condition, per experiment. They display not only the overall cross-correlation (black), but also the cross-correlation based on the scatter plots (red) and the control (gray area); see chapter 2 and also figure 2.10, for more information. The error bars in the black curves are SEM, based on dividing data from a single experiment into four groups, and calculating four CCs, for which the SEM is calculated. The black CCs here correspond to the shaded CCs that are shown in other figures where multiple experiments are combined into one plot. (A) $R_{Q,\mu}(\tau)$ for the wild type cells. (B) $R_{Q,\mu}(\tau)$ for the ΔcAMP cells, with 80 μM cAMP supplemented to the growth medium. (C) $R_{Q,\mu}(\tau)$ for the ΔcAMP cells, with 800 μM cAMP supplemented to the growth medium. (D) $R_{Q,\mu}(\tau)$ for the ΔcAMP cells, with 5000 μM cAMP supplemented to the growth medium. (In the y-axis label, "C" stands for the cyan fluorescent reporter concentration that was tracked here.)
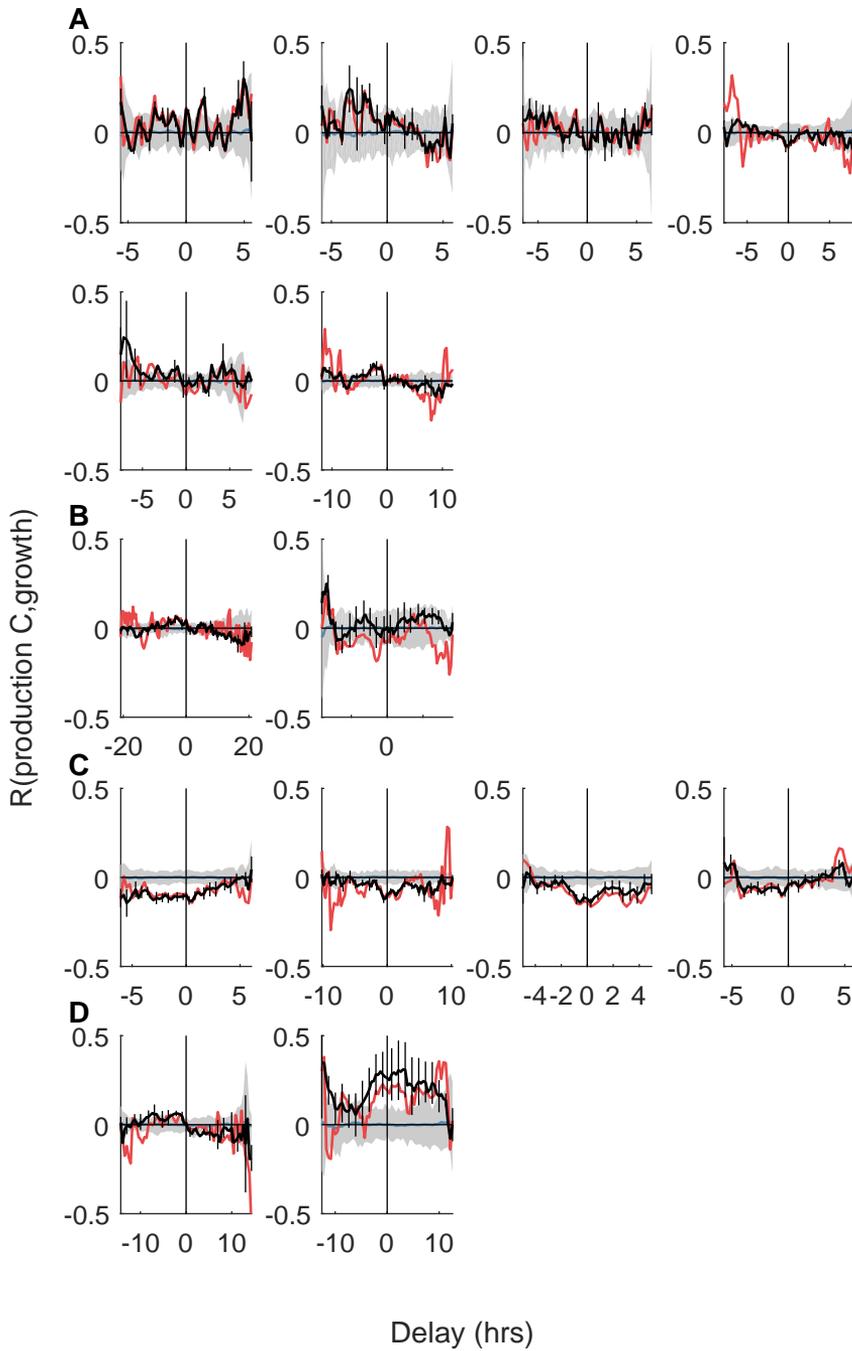
5

Figure 5.31: Constitutive production-growth cross-correlations $R_{p_Q,\mu}(\tau)$ **per experiment.** This figure is similar to figure 5.29, except that it relates to constitutive reporter measurements. These graphs shows cross-correlations (CCs) per condition, per experiment. It displays not only the overall cross-correlation (black), but also the cross-correlation based on the scatter plots (red) and the control (gray area); see chapter 2 and also figure 2.10, for more information. The error bars in the black curves are SEM, based on dividing data from a single experiment into four groups, and calculating four CCs, for which the SEM is calculated. The black CCs here correspond to the shaded CCs that are shown in other figures where multiple experiments are combined into one plot. (A) $R_{p_Q,\mu}(\tau)$ for the wild type cells. (B) $R_{p_Q,\mu}(\tau)$ for the ΔcAMP cells, with 80 μM cAMP supplemented to the growth medium. (C) $R_{p_Q,\mu}(\tau)$ for the ΔcAMP cells, with 800 μM cAMP supplemented to the growth medium. (D) $R_{p_Q,\mu}(\tau)$ for the ΔcAMP cells, with 5000 μM cAMP supplemented to the growth medium. (In the y-axis label, "C" stands for the cyan fluorescent reporter production that was tracked here.)

5

Growth rate (doublings/hr)

Time (minutes)

Figure 5.32: Growth rate for all cell lineages per conditions per microcolony. Each panel corresponds to a microcolony, letters indicate conditions. The gray lines show single lineage traces, the black lines the population average. Colored lines highlight example single lineage traces to illustrate single cell behavior. Dashed and dotted lines indicate $2 \cdot \sigma$ and $5 \cdot \sigma$ boundaries from the overall mean respectively. As before, the displayed conditions are (A) wild type cells, (B) ΔcAMP cells growing on 80 μM cAMP, (C) ΔcAMP cells growing on 800 μM cAMP and (D) ΔcAMP cells growing on 5000 μM cAMP.

5

Fluorophore concentration (a.u./pixel)

Time (minutes)

5

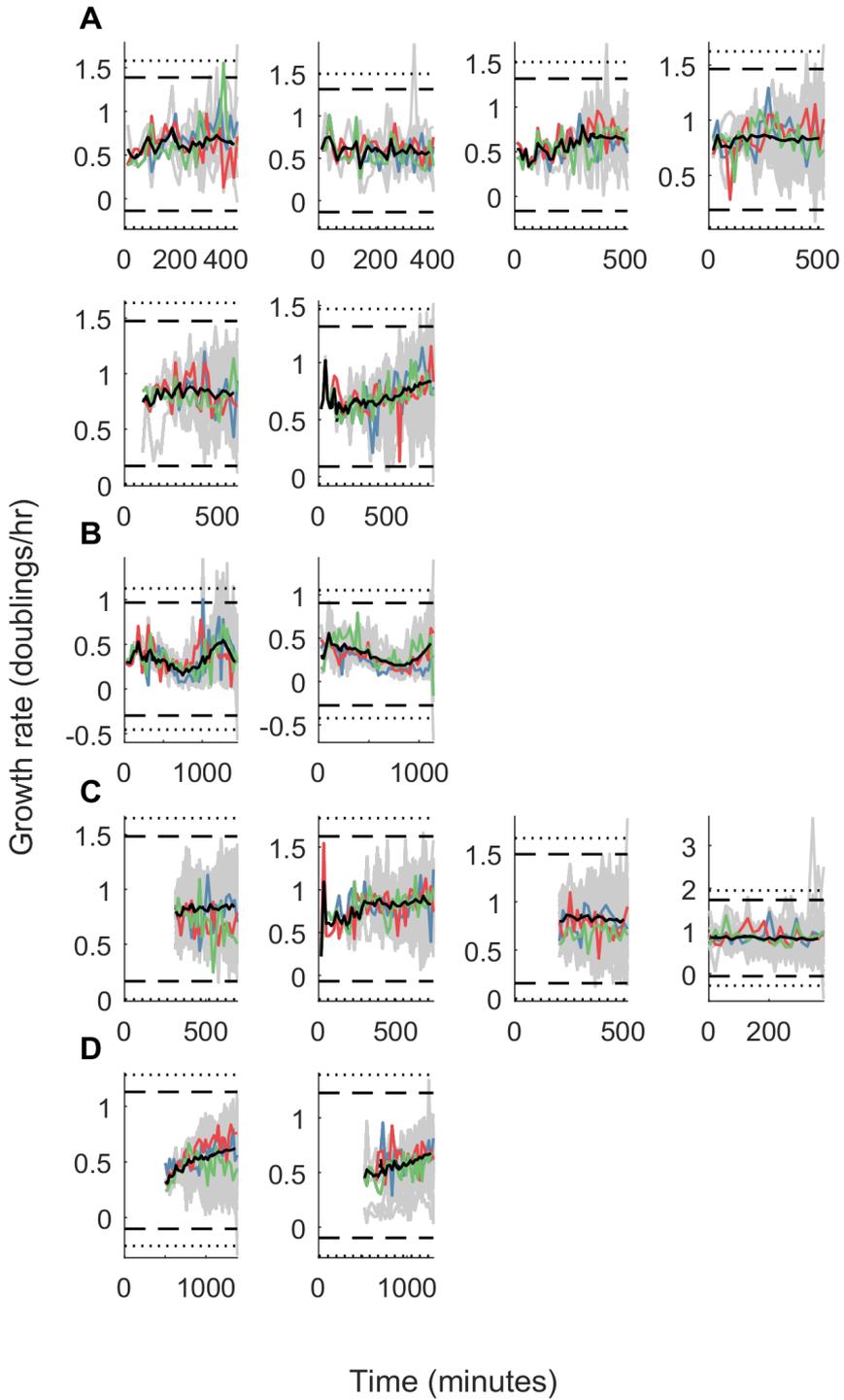**Figure 5.33: Metabolic reporter concentrations for all cell lineages per conditions per micro-colony.** Each panel corresponds to a microcolony, letters indicate conditions. The gray lines show single lineage traces, the black lines the population average. Colored lines highlight example single lineage traces to illustrate single cell behavior. Dashed and dotted lines indicate $2 \cdot \sigma$ and $5 \cdot \sigma$ boundaries from the overall mean respectively. As before, the displayed conditions are (A) wild type cells, (B) ΔcAMP cells growing on 80 μM cAMP, (C) ΔcAMP cells growing on 800 μM cAMP and (D) ΔcAMP cells growing on 5000 μM cAMP.
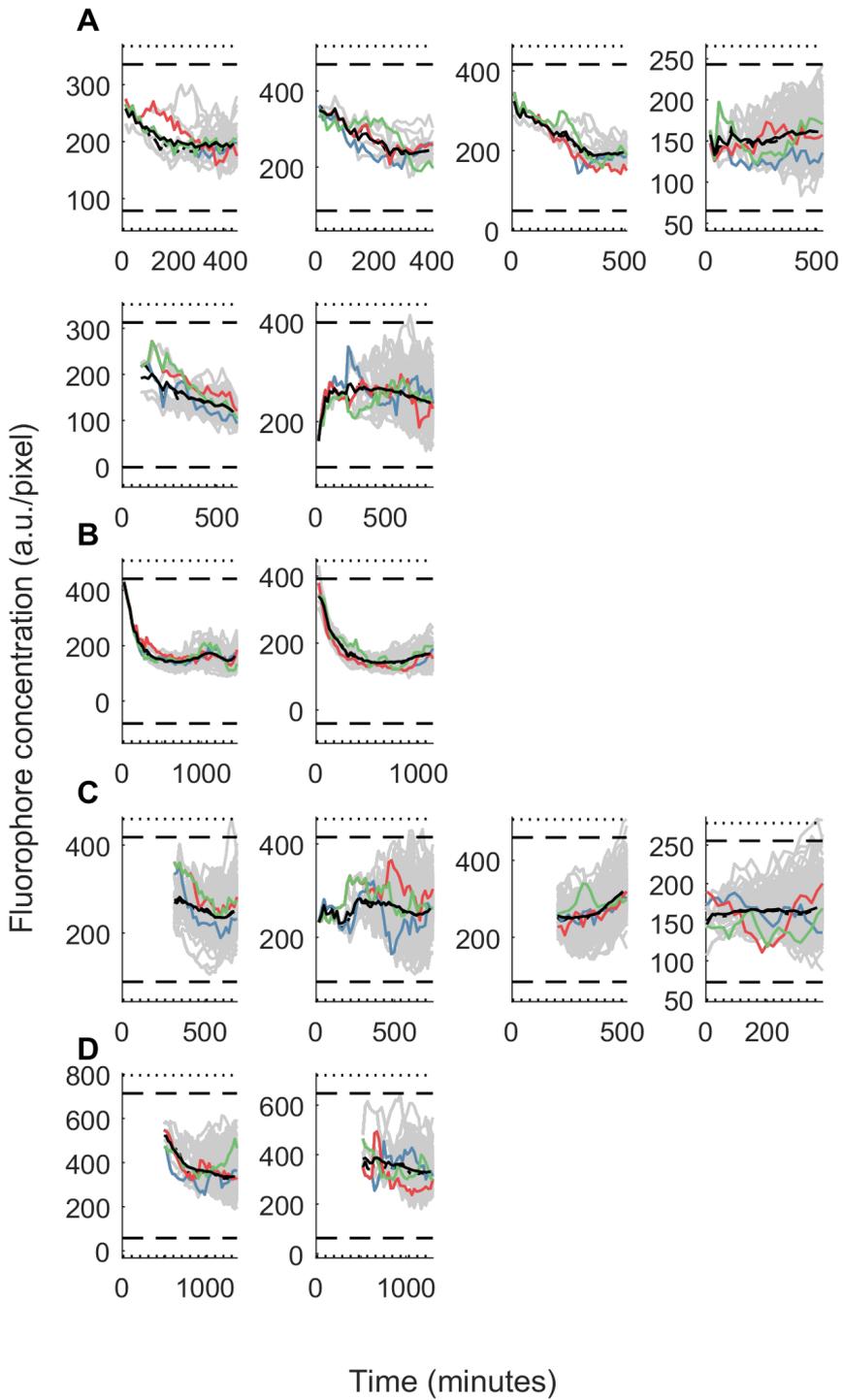
5
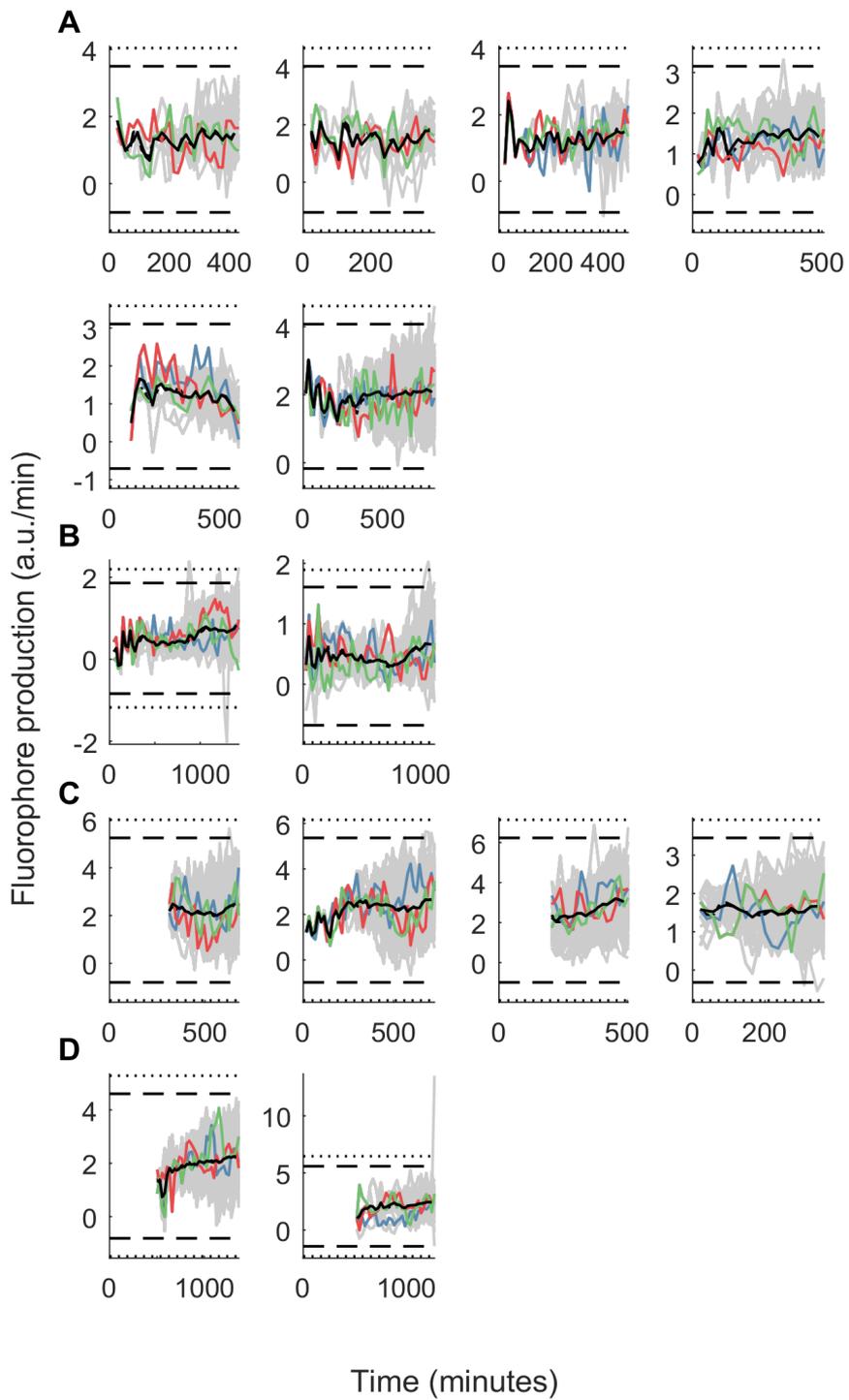
Figure 5.34: Metabolic reporter production rates for all cell lineages per conditions per microcolony. Each panel corresponds to a microcolony, letters indicate conditions. The gray lines show single lineage traces, the black lines the population average. Colored lines highlight example single lineage traces to illustrate single cell behavior. Dashed and dotted lines indicate $2 \cdot \sigma$ and $5 \cdot \sigma$ boundaries from the overall mean respectively. As before, the displayed conditions are (A) wild type cells, (B) ΔcAMP cells growing on 80 µM cAMP, (C) ΔcAMP cells growing on 800 µM cAMP and (D) ΔcAMP cells growing on 5000 µM cAMP.
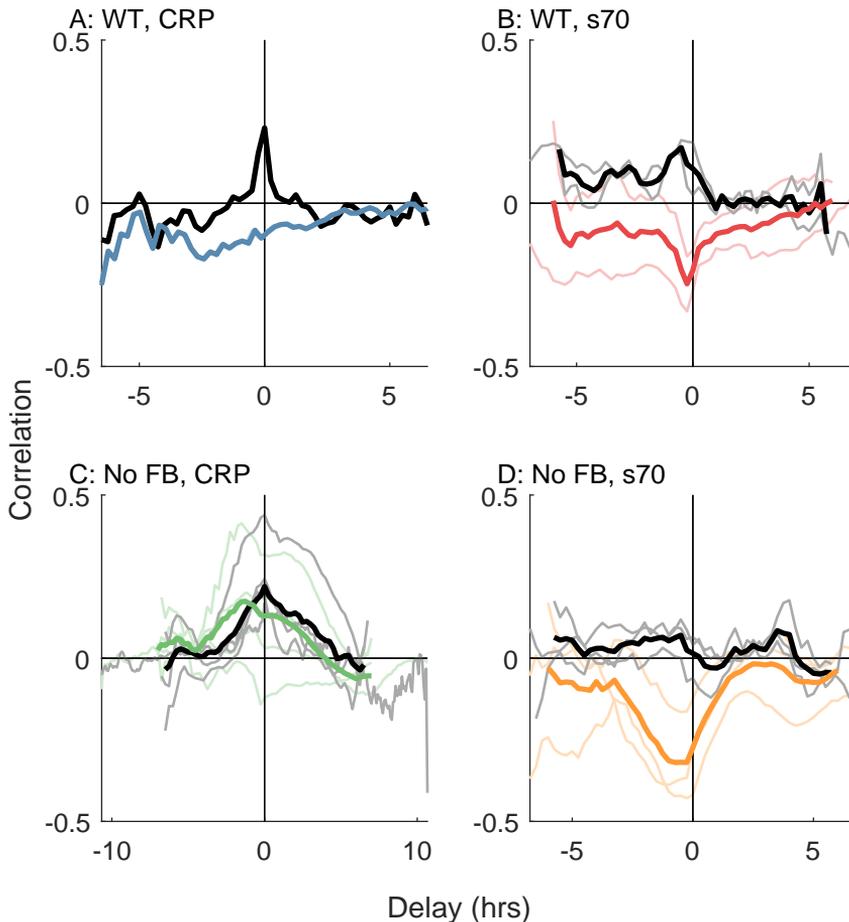
5

Figure 5.35: Exploratory experiments with plasmid reporters show similar results as experiments with our chromosomal reporters. The interaction of metabolic and growth fluctuations was also measured by plasmid constructs in initial experiments, where the metabolic reporter and constitutive reporter were put into different cell lines on similar plasmids with a GFP reporter. The plasmids on which these cross-correlations are based are the same as described in ref [153]. Except for the fact that the reporters are on plasmids instead of chromosomally inserted, and both placed in separate cell lines instead of the same cell line, the experimental conditions are exactly the same as in the experiments presented earlier involving wild type and ΔcAMP cells plus 800 μM cAMP (figure 5.4). Consistently, in this supplemental figure, we observe slightly different but qualitatively similar cross-correlations as in figure 5.4. The constitutive reporter (top right for wild type cells and bottom right for ΔcAMP cells plus 800 μM cAMP) shows cross-correlations that are most similar to a dilution scenario, and as before (figure 5.4), the metabolic reporter in wild type cells shows a similar behavior. On the other hand, without feedback, the metabolic reporter shows very different behavior in the ΔcAMP cells plus 800 μM cAMP. As before, black lines correspond to production rate-growth cross-correlations, and colored lines to concentration-growth cross-correlations. Darker lines are averages of multiple experiments, the shaded lines show the separate experiments (when more than 1 experiment was conducted).

ASC number	Shorthand	Description	Source
ASC840	ΔCRP	*cyaA crp* null mutant, also known as bBT92. In addition to the *crp* gene, also the *cyaA* gene has been knocked out, since some of the constitutive CRP proteins still have a residual response to cAMP.	[153]
ASC843		MG1655 WT strain (bBT12) + GFP w/o promoter PSCS101 (control), plasmid aka pU66 (kanamycin resistance on plasmid).	[153]
ASC847	pHA5	p_CRP-CRP construct in a pBR322 plasmid; stored in DH5α. This is the CRP protein sequence behind its native promoter. (Ampicilin resistance on plasmid.)	[153, 186]
ASC845	pHA7	p_bla-CRP construct in a pBR322 plasmid; stored in DH5α. This is the CRP protein behind a bla promoter. (Ampicilin resistance on plasmid.)	[153, 186]
ASC859	pHA7*1	p_bla-CRP*1 construct in a pBR322 plasmid; stored in DH5α. The star indicates this is a constitutively active version of CRP, the number indicates which version. See Aiba et al. for more information [186]. (Ampicilin resistance on plasmid.)	[153, 186]
n/a	pHA7*2	p_bla-CRP*2 construct in a pBR322 plasmid; plasmid only. Idem. (Ampicilin resistance on plasmid.)	[153, 186]
n/a	pHA7*3	p_bla-CRP*3 construct in a pBR322 plasmid; plasmid only. Idem. (Ampicilin resistance on plasmid.)	[153, 186]

Table 5.2: Additional strains produced for this work, but not used in experiments presented here. ASC stands for AMOLF strain collection. FB indicates this strain was acquired by Ferhat Büke, a member of both the Sander Tans lab and Greg Bokinsky lab. VS indicates this strain is produced at AMOLF by technician Vanda Sunderlikova. Note with ASC844: this plasmid was received from the Alon lab in a DH5α strain, other plasmids were received purified and were transferred to DH5α strains at AMOLF.

ASC number	Shorthand	Description	Source
ASC858	pHA7*4	p_bla-CRP*4 construct in a pBR322 plasmid; stored in DH5α. Idem. (Ampicilin resistance on plasmid.)	[153, 186]
ASC846	pHA7*5	p_bla-CRP*5 construct in a pBR322 plasmid; stored in DH5α. Idem. (Ampicilin resistance on plasmid.)	[153, 186]
ASC844	pHA7*6	P_bla::CRP*6 construct in a pBR322 plasmid; stored in DH5α. Idem. (Ampicilin resistance on plasmid.)	[153, 186]
ASC852		pHA7 strain, with additionally p_CRPr plasmid.	
ASC853		pHA7 strain, with additionally p_s70 plasmid.	
ASC854		pHA7*5 plasmid and p_CRPr in ΔcAMP strain.	
ASC855		pHA7*5 plasmid and p_s70 in ΔcAMP strain.	
n.a.		Wild type strain with pHA7*2 and P_s70. (This strain was attempted to be made, but grew very slowly in TY which resulted in failure to produce it in a first attempt.)	
ASC878		Wild type strain with pHA7*6 and p_s70.	
ASC879		Wild type strain with pHA7*2 and P_RCRP.	
ASC880		Wild type strain with pHA7*6 and P_RCRP.	
ACS893		ΔcAMP strain plus P_CRPr plasmid. (ASC894 is identical to this strain.)	
asc1005		ΔCRP with the following additional chromosomally inserted reporters: delta(galk)::s70-mCerulean-kanR + delta(intc)::rcrp-mVenus-cmR. (Kanamycin and chloramphenicol resistant.)	VS
asc1018		pzs101-yfp; ptet system + repressor, expressing a YFP fluorescent reporter. (Ampicilin resistant.)	FB
asc1021		ASC1005 plus ptet-CRP*2 in asc1018 vector (instead of yfp) (Ampicilin resistant.)	VS
asc1022		ASC1005 plus ptet-CRP*5 in asc1018 vector (instead of yfp) (Ampicilin resistant.)	VS
asc1023		ASC838 strain plus ASC1018 plasmid (control).	VS

Table 5.2 continued.

6

# Ribosomal dynamics, a puzzling affair

6.1. Introduction

6.1.1. The central role of the ribosome

Ribosomes are life. In 1958, Francis Crick, who discovered the structure of DNA with James Watson and Rosalind Franklin, formulated what is called the central dogma of molecular biology. It states that DNA holds the amino acid sequence information required to build a protein, and that this sequence information is first translated to RNA, which is then transcribed into a protein [211]. The dogma relates immediately to some of the features that are said to define life: the capability to store information and the regeneration of components from scratch [212, 213]. Slightly earlier, in 1955, "new cytoplasmic components" were already discovered under an electron microscope [214]. These components turned out to be the components in the cell that perform the last step of the central dogma: the synthesis of proteins using mRNA molecules as templates; also known as transcription. The components became known as ribosomes. Given the idea that early life consisted only of RNA molecules that could catalyze other reactions (the "RNA world" [215]), it is not surprising that the ribosomes consist of both catalytic RNA molecules and (many) proteins, as shown in figure 6.1. Given these facts, we can say ribosomes are (a key element of) life.

6.1.2. Understanding the role of ribosomes in protein fluctuations

In this thesis, we investigate temporal fluctuations in the concentrations of cellular components. As mentioned in earlier chapters (4 and 5), these fluctuations are a result of the intrinsic stochastic nature of chemical processes that occur in the cell. Given the central role of the ribosome in the cell that was just discussed, ribosomes might play a big role in such fluctuations. Indeed, it is often suggested that fluctuations in ribosomal con-
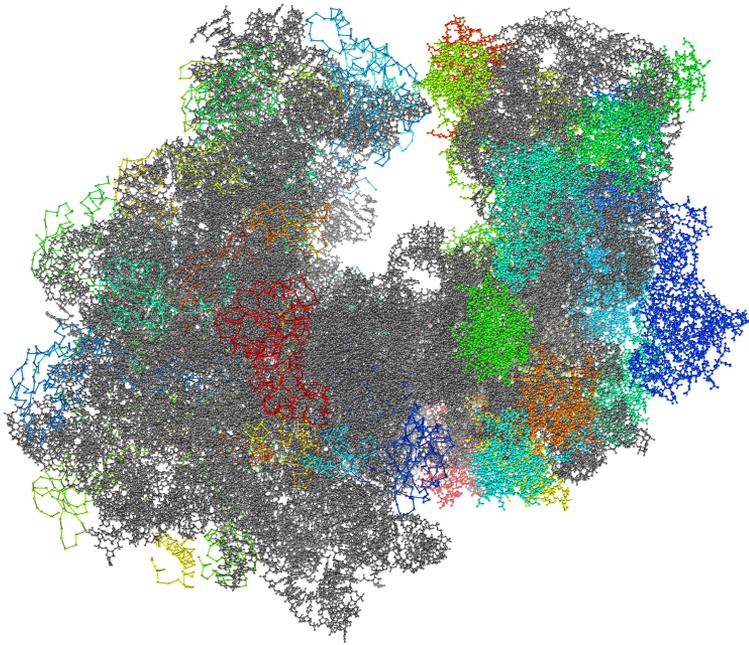
Figure 6.1: Picture of the ribosome. The ribosome consist of a small and a large subunit. These two come together on messenger RNA templates to perform protein synthesis. The complete ribosome is made of 3 ribosomal RNA molecules and more than 50 proteins, shown in grey and color in the picture above, respectively [206]. The 16S rRNA and 23S rRNA molecules function as backbone. Proteins labeled S1 up to S22 bind to the 16S rRNA to form the small subunit, and proteins labeled L1-L36 and the 5S rRNA bind to the 23S rRNA to form the large subunit [171]. The Ecocyc database lists 58 ribsomal proteins, though also different numbers of ribosomal proteins are listed (ref. [206] e.g. talks about 54 ribosomal proteins). Interestingly, the ribosomal RNA constitutes 73-80% of the total RNA found in an *E. coli* cell (mRNA constitutes 3-4.5% and tRNA 15-20%) [207]. Assembly of the ribosome requires many co-factors [206]. This image is based on x-ray crystal-lography (PDB ID: 4v4q, [208]). Pdb files downloaded from `www.rcsb.org` [209] visualized with UCSF Chimera (version 1.11.2, build 41376, [210]).

centration could lead to in cell-wide protein concentration fluctuations [11–14]. Such concerted concentration fluctuations are also referred to as extrinsic noise (as opposed to intrinsic noise, which refers to fluctuations that only occur in one cellular species) [9].

6.1.3. QUANTIFYING RIBOSOMAL DYNAMICS

In a previous thesis from the Tans lab, Noreen Walker quantified the dynamic relationship between single cell ribosomal expression fluctuations and growth in steady state conditions [18]. The experiments involved presented many challenges, of which a number are described in her thesis. Given the previously described central role of the ribosome, it was hypothesized that the stochastically fluctuating concentration of ribosomes might be limiting, meaning that ribosomal fluctuations might result in growth rate fluctuations. In single cell time lapse experiments, no clear indications were found to support this hypothesis. Instead, ribosomal proteins L19 and L31 that were labeled with mCherry (L31-R and L19-R) showed expression-growth cross-correlations (CCs) indicative of dilution-scenario behaviour in minimal medium, or none at all in rich medium (see chapters 4 and 5 for discussions about the use of cross-correlations to interpret dynamics, and also Kiviet et al. [10]). mCerulean labeled L31 ribosomal protein (L31-C) showed expression-growth CCs with small correlations around zero delay in minimal medium. This difference between the L31-C and L31-R experiments was unexpected, as only the label was different. In rich medium, the L31-C cross-correlation was consistent with the L31-R cross-correlation, and showed almost no correlation. Since the ribosome mainly consists of RNA, a GFP reporter under the control of one of the ribosomal RNA promoters was also used (rrna-G). Like the L31-C reporter, the dynamics of this rrna-G reporter showed positive correlations in minimal medium and only very small correlations in rich medium. It remained unclear whether the differences between these observations were due to physical effects or due to day-to-day experimental variation. Also growth-expression scatter plots were used, to investigate possible interesting shapes of growth-expression relationships, but this yielded no noteworthy shapes.

In an attempt to force the cell in a scenario where ribosomes are limiting, experiments were conducted where cells were grown in the presence of sub-inhibitory concentrations of tetracycline, an antibiotic. Though also interesting in itself, the outcome of such an experiment could serve as a reference to interpret other experiments. In rich medium, both for the L31-R and the rrn-G reporter, the antibiotic did not result in more limiting behaviour, i.e. correlations did not become more positive. Given the disparate observations, both between reporters and conditions, and the absence of a point of reference, the nature of the ribosomal dynamics remained elusive; it was concluded that ribosomal fluctuations perhaps do not have a pivotal role in steady state cellular growth dynamics.

6.1.4. NEW WORK

In this chapter, we will describe additional experiments that were aimed at gaining further insights in the ribosomal dynamics. Specifically, we tried to understand whether fluctuations in ribosomal concentration could have cell-wide implications. To answer this question, we tried both new experiments as well increase the throughput of existing

experiments. Since this is an ongoing project, some sections in this chapter will be more succinct.

6.2. RESULTS

6.2.1. ADDITIONAL STRAINS

ASC number	Shorthand	Description	Source
ASC976	Prrna-C, pn25-Y	Δphp::pn25-mVenus-cmR, Δche::Prrsa-mCerulean-kanR. (Kanamycin and chloramphenicol resistance.)	VS
ASC968	L19-C, pn25-Y	L19-gc-mCerulean-kanR (GC linker), Δphp::pn25-mVenus-cmR. (Kanamycin and chloramphenicol resistance.)	VS
ASC1058	L9-R, S2-Y	Also known as JE202. rplI-mCherry-KanR (L9), rpsB-venus-CmR (S2). (Kanamycin and chloramphenicol resistance.) Gift from Johan Elf lab.	[57]
ASC666	L31-R, gltA-G	L31::mCherry-kanR, gltA::gfpA206K-cat	[10]
ASC810	L31-C	L31-gc-mCerulean-kanR (GC linker). (Kanamycin resistant.)	NW, VS

Table 6.1: Strains used in this chapter. Strain ASC1058 was a kind gift from the Johan Elf lab. VS indicates these strains were created by Vanda Sunderlikova, technician in the Tans lab. ASC stands for AMOLF strain collection. Note that Y, C, G and R indicate yellow, cyan, green and red fluorescent reporters, respectively. The first two strains in this table were specifically created for experiments described in this chapter, and the third strain was requested specifically for work described in this chapter. The last two strains in the table were created earlier.

To be able to further explore the implications of ribosomal fluctuations, we produced additional strains. Importantly, we were interested in the effect that fluctuations in ribosomal expression would have on the single cell capacity to produce proteins. We therefore created two dual-label strains that both carried ribosomal reporter constructs, as well as constitutively expressed fluorescent reporters. The constitutive reporter served as a readout for fluctuations in protein production that likely affect all proteins that are produced in the cell. In the first strain, we fused the mCerulean fluorescent reporter sequence to the ribosomal RNA promoter (rrsa), and additionally inserted an mVenus sequence to the constitutive pn25 promoter. Both were chromosomally inserted. In the second strain, we introduced an mCerulean sequence behind the chromosomal L19 ribosomal protein sequence, thus creating a strain that produces L19 ribosomal proteins that are fused to mCerulean fluorescent labels. We also introduced the constitutive pn25-mVenus reporter to this strain. These strains, the Prrna-C, pn25-Y strain and the L19-C, pn25-Y strain are listed in table 6.1. Additionally, supplementary table 6.2 gives an overview of the operons in which the ribosomal proteins are encoded.

Furthermore, as the ribosome consists of many ribosomal proteins, we wanted to be able to confirm dynamics observed for one labeled ribosomal protein, also for other labeled ribosomal proteins. For this purpose, we requested a strain from the Elf lab which had both the L9 and S2 ribosomal proteins labeled by the mCherry and Venus fluorescent proteins, respectively [57]. This strain was kindly supplied by the Elf lab, and is also listed in table 6.1 with the shorthand notation L9-R, S2-Y.

Thus, we expanded our ability to track ribosomal dynamics by expanding our collection of ribosomal reporters, and introduced a way to probe the effect of ribosomal fluctuations on protein expression by the introduction of the pn25 reporter.

6.2.2. ANTIBIOTIC SHIFT EXPERIMENTS TO CREATE LIMITING SITUATIONS

To understand the dynamics of a process, it is often convenient to grasp what happens in extreme cases. We therefore attempted to device an experiment which would lead to a situation where single cells that expressed more ribosomes than the population average would have a growth advantage, thus possibly forcing a "limiting" situation. We did this by growing cells in minimal medium in our microfluidic device 1 for a few hours, and then switching to minimal medium supplemented with antibiotics (see chapters 2 and 3 for a description of the microfluidics device). We hypothesized that right after the medium has switched, cells in the population — which at that point are calibrated for growth in minimal medium — that happen to express more ribosomes have a growth advantage. To analyse whether this was indeed the case, we made growth-concentration scatter plots for multiple points in time during the experiment. We further analysed these measurements by calculating the slope and correlation of these scatter plots for each point in time.

Some data involving an antibiotic switch was already gathered by Tans lab alumni Sarah Boulinea for the L31-R, gltA-G strain (see table 6.1); we also ran our analysis on this dataset. Also, two additional experiments were conducted, one with the L31-C strain where one microcolony was analyzed, and one with the rrsa-C, pn25-Y strain, where three microcolonies were analyzed. Before looking at the scatter plots, we first look at the trends in fluor concentration for all these five datasets, which are shown in figure 6.2. This figure shows that for the Prrna-C strain datasets, the signal goes up after the switch to antibiotics supplemented medium. However, for both for the L31-R and L31-C datasets, the signal appears to be going down after the shift (though pre-shift fluctuations on the population average level of the L31-R data make it unclear what is exactly going on). It is unclear why this happens, since ribosome inhibition should increase ribosomal demand and accordingly, expression [170]. We now turn to the scatter plots. For brevity we only show the representative series of scatter plots that relates to the L31-C dataset, see figure 6.3. This figure shows that there is no clear change in correlation visible by eye after the switch to medium with antibiotics. We further quantified the data by calculating both the slope (least square fit by Matlab's `polyfit` function) and correlation coefficient for each of the scatter plots. The analyses for all five datasets are shown in figures 6.4-6.8. The L31-R dataset (figure 6.4) arguably shows a small increase in correlation approximately 200 minutes after the switch, but the correlation also decreases swiftly after that. The L31-Y dataset (figure 6.5) is difficult to interpret, likely the positive correlations before the switch are caused by chance as there are only a few cells in the
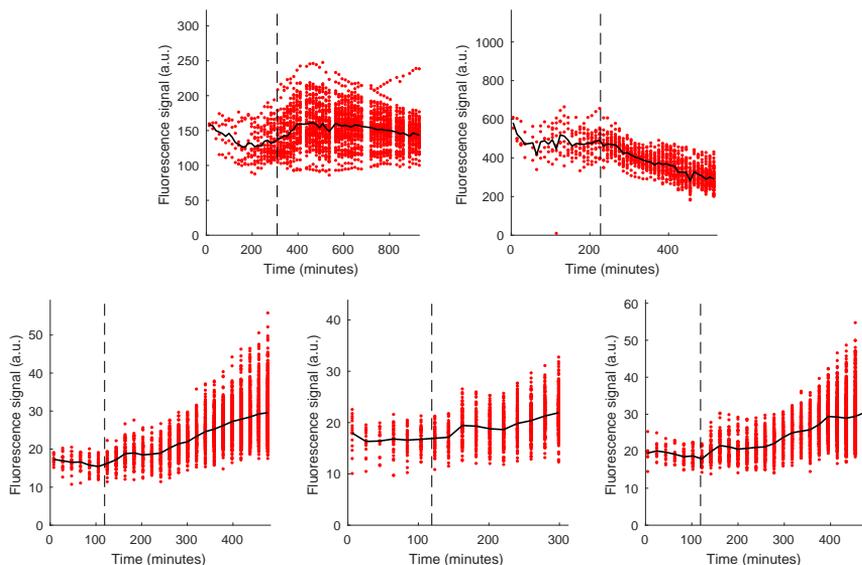
Figure 6.2: Fluorescence intensity for switch from clean medium to medium supplemental with antibiotics. Red dots show single cell obsvervations, the black line indicates the colony average. (Top left) L31-R strain. (Top right) L31-C strain. (Bottom three) All are rrsa-C strains. "C", "Y" and "R" indicate the fluorescent labels by their colors: mCerulean (cyan), mVenus (yellow) and mCherry (red) respectively.

microcolony at those points in time. The first Prrsa-Y dataset (figure 6.6) shows a change in signal right after the switch, but this is not seen in the two other Prrs-Y datasets (6.7-6.8). In conclusion, it is difficult to interpret these datasets, as the different datasets show different trends.

6.2.3. HIGHER THROUGHPUT AND ADDITIONAL CROSS-CORRELATIONS

MANY EXPERIMENTS IN ONE

One reason that it is hard to analyse the data from the antibiotic switch experiments using microfluidic device 1, is that the size of the microcolony starts out small, which leads to a high variability of the correlation during the first part of the experiment. To produce a dataset with a constant large amount of cells, we performed additional antibiotic switch experiments in microfluidic device 2. Conveniently, microfluidic device 2 allows for multi-day experiments, which allowed us to also measure additional steady state expression-growth cross-correlation curves. Importantly, we performed these additional experiments with the new Prrna-C, pn25-Y and L19-C, pn25-Y strains, to investigate the effect of ribosomal fluctuations on protein production rates with the pn25-mVenus reporter. An example of a typical experimental design is given by the measurements we performed on the L19-R pn25-Y strain, which involved the following sequence of supplied medium: *6 hours TY, 6 hours TY + TET, 2 hours TY, 2 hours TY + TET, 2 hours TY, 2*
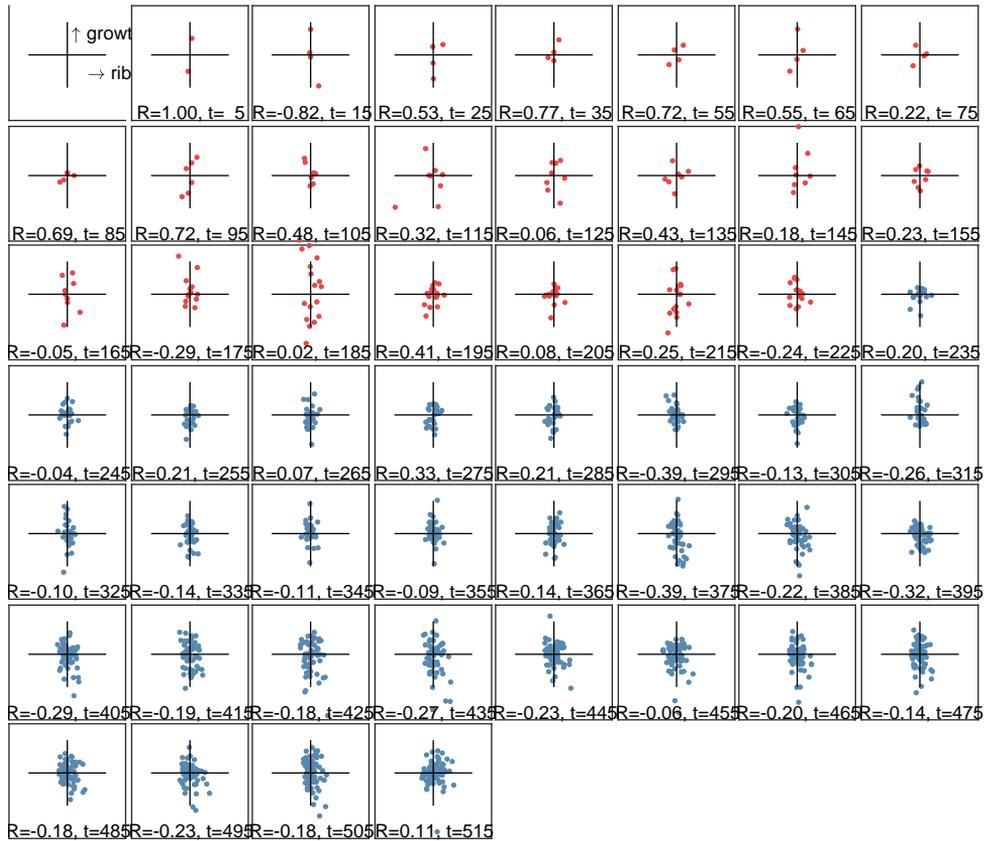
Figure 6.3: Scatter plots for switch from clean medium to medium supplemental with antibiotics. Growth (doublings per hour) is shown on the y axes, and ribosome concentration (a.u.) is shown on the x axis. Data shown here is from the L31-C ribosomal reporter strain. It was first grown in M9 medium (indicated by red scatter plots), and then in M9 medium plus 0.5 μM tetracycline (indicated by blue scatter plots). The labels additionally indicate the correlation R, and the time t (in minutes).

Figure 6.4: Quantification of the correlation between growth and ribosomal expression in the antibiotics switch experiment. Data from the L31-R strain. The number of cells gradually increased over the course of the experiment (A), whilst growth rate decreased after the switch (B). The dynamics of growth and expression were quantified by the slope (C) and correlation coefficient (D) of the growth-expression scatter plots at each point in time. Data displayed in black is based on the whole colony, blue symbols indicate single cell data.
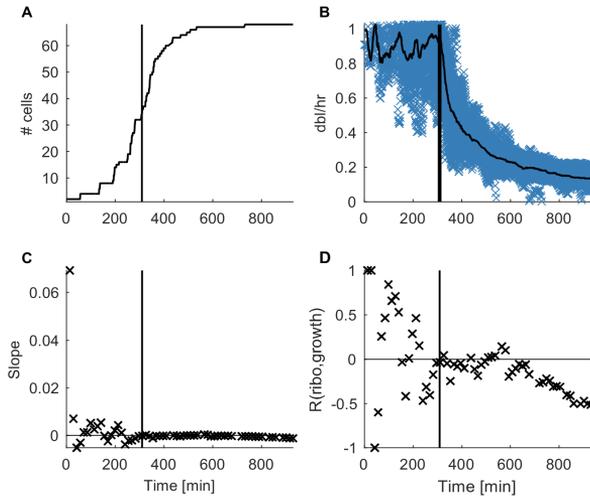
6

Figure 6.5: Quantification of the correlation between growth and ribosomal expression in the antibiotics switch experiment. Data from the L31-C strain. The number of cells gradually increased over the course of the experiment (A), whilst growth rate decreased after the switch (B). The dynamics of growth and expression were quantified by the slope (C) and correlation coefficient (D) of the growth-expression scatter plots at each point in time. Data displayed in black is based on the whole colony, blue symbols indicate single cell data.

Figure 6.6: Quantification of the correlation between growth and ribosomal expression in the antibiotics switch experiment. Data from the Prrsa-C, pn25-Y strain. The number of cells gradually increased over the course of the experiment (A), whilst growth rate decreased after the switch (B). The dynamics of growth and expression were quantified by the slope (C) and correlation coefficient (D) of the growth-expression scatter plots at each point in time. Data displayed in black is based on the whole colony, blue symbols indicate single cell data.
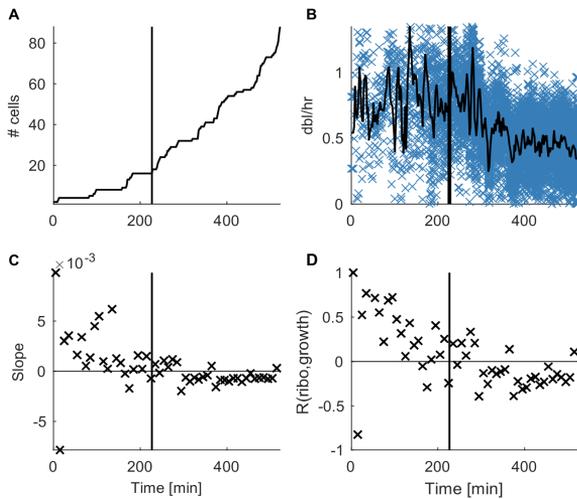
6
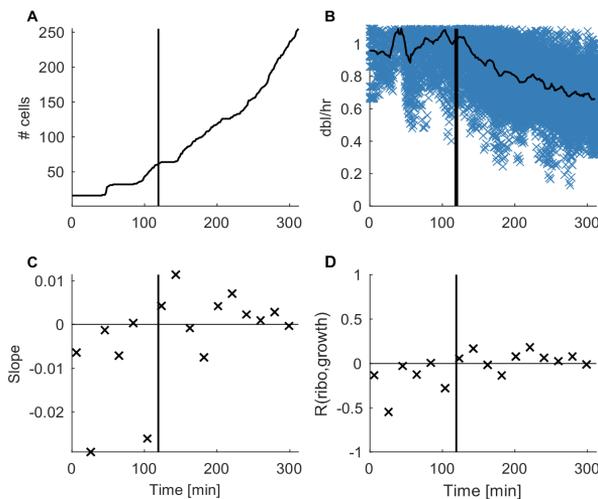
Figure 6.7: Quantification of the correlation between growth and ribosomal expression in the antibiotics switch experiment. Data from the Prrsa-C, pn25-Y strain. The number of cells gradually increased over the course of the experiment (A), whilst growth rate decreased after the switch (B). The dynamics of growth and expression were quantified by the slope (C) and correlation coefficient (D) of the growth-expression scatter plots at each point in time. Data displayed in black is based on the whole colony, blue symbols indicate single cell data.
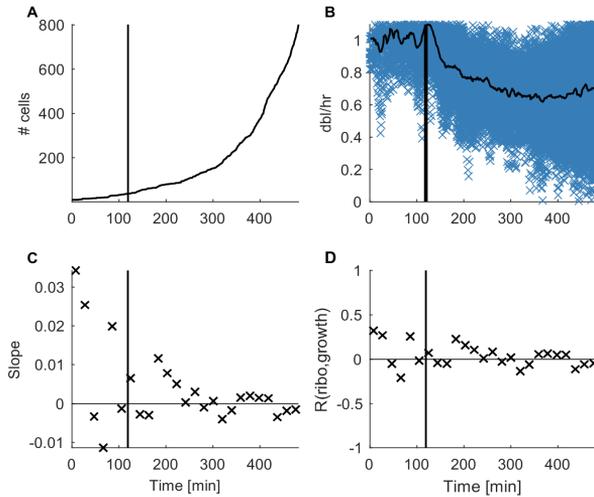
Figure 6.8: Quantification of the correlation between growth and ribosomal expression in the antibiotics switch experiment. Data from the Prrsa-C, pn25-Y strain. The number of cells gradually increased over the course of the experiment (A), whilst growth rate decreased after the switch (B). The dynamics of growth and expression were quantified by the slope (C) and correlation coefficient (D) of the growth-expression scatter plots at each point in time. Data displayed in black is based on the whole colony, blue symbols indicate single cell data.
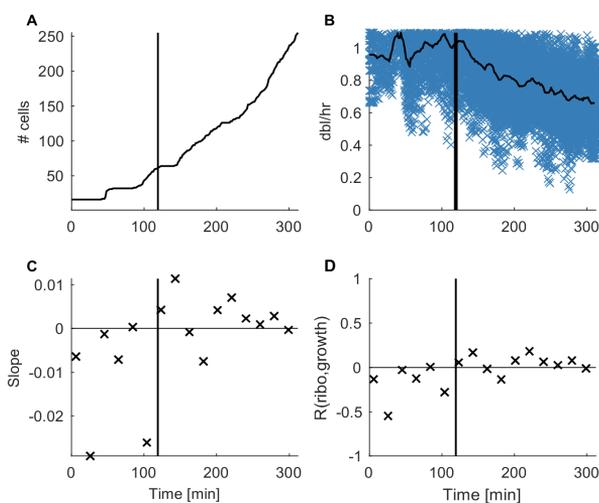
6

hours TY + TET, 2 hours TY, 6 hours M9, 6 hours M9 + TET, 2 hours M9, 2 hours M9 + TET, 2 hours M9, 2 hours M9 + TET. (M9 indicates M9 minimal medium here, supplemented with lactose, uracil and tween20; see chapter 3 for further description of used media).

Thus, to recapitulate, each experiment with microfluidic device 2 can potentially give (1) data on the ribosome-growth dynamics during antibiotic shifts, (2) data on ribosome-protein expression dynamics during antibiotic shifts, (3) additional CCs regarding ribosome-growth dynamics to characterize steady state relationships, and (4) CCs of previously uncharacterised ribosome-protein expression dynamics. Additionally, we could also analyse the response growth medium shifts and the response to medium shifts that occur at high frequency, but we will not focus on this here.

Results of the experiments

The large amount of data generated by microfluidic device 2 is both an advantage and a disadvantage. We will not go into technical details here, but both the amount of data and also the nature of the growth in the wells presented the computer analysis with new challenges. We therefore only analysed parts of the data that was produced by these experiments.We chose to first analyse the data where we assumed cells had reached steady state; and where we can perform cross-correlation analyses (i.e. analyses mentioned in points 3 and 4 in the previous paragraph). Additional data recorded at the times of the switches exists but has not been fully analysed yet. We also note that additional manual corrections could perhaps further refine the analyses that we show here. For example, some of the tracked lineages show unrealistic fluctuations which might be due to artefacts in the computer analysis and might be removed or corrected (see supplemental

figures 6.26 and 6.27), and also the analysis now only takes into account a selection of the total amount of imaged cells, this selection could be extended to obtain more data.

RESULTS FROM STRAINS WITH RIBOSOME AND CONSTITUTIVE REPORTERS

Figures 6.10-6.15 show cross-correlation functions obtained from the Prrna-C, pn25-Y and L19-C, pn25-Y strains. We also generated negative control cross-correlations by correlating the time trace from one lineage branch with the time trace from another randomly selected lineage branch; random pairs of branches should not correlate for steady state measurements with a sufficient sample size. This procedure is also described in section 2.4.5 in the methods chapter, under subheading "Controls". The mean and maximum correlations of 50 of such randomly generated negative control curves are also plotted in figures 6.10-6.15. Most of the Prrna-C, pn25-Y and L19-C, pn25-Y cross-correlations show large error bars, and often the values are in the same range as negative control values (figures 6.10-6.15). This indicates that more data is required to draw definitive conclusions. Furthermore, figure 6.9 shows that the concentration of ribosomal RNA reporters was quite low in comparison to other reporters. This might be due a weak ribosomal binding site, and further emphasizes caution is required when interpreting this data. Nevertheless, figures 6.10 and 6.11 show the CCs calculated for the pn25 constitutive expression-growth relationship for growth in minimal medium, TY medium, and TY medium supplemented with sub-inhibitory concentration of the antibiotic tetracycline (0.5 μM). Most of the pn25 concentration-growth curves show negative correlation values for negative delays (figure 6.10). The production rate-growth curves do not show clear trends (figure 6.11). We saw earlier (see chapter 5) that constitutive reporters often show dilution mode dynamics. The CCs we observe here could be consistent with that transmission mode. Furthermore, figures 6.12 and 6.13 show data from the same strains, but show the expression-growth CCs for the ribosomal reporters (rRNA or L19). The concentration-growth and production-growth CCs for the labeled L19 protein (panel A in figures 6.12 and 6.13 respectively) do not show a clear trend. The concentration-growth and production-growth CCs for the behaviour of the *rrna* reporter (shown in panels 6.12B-D and 6.13.B-D respectively), show behaviour that also might be consistent with dilution mode transmission of noise; negative concentration-growth at negative delays correlations are seen both for M9 and TY medium, and also for the condition where antibiotic was added to the medium. It is striking that there are no noticeable differences in the rRNA expression growth dynamics between different conditions. One explanation might be that the promoter–reporter construct does not fully capture the dynamics of ribosomal RNA expression, as the construct does not directly label ribosomal RNA (which is something that can only be done for proteins).

The dual reporter strains with both ribosomal and pn25 constitutive reporters were constructed to allow us to not only study ribosome-growth dynamics, but also allow us to study the impact of ribosomal fluctuations and protein expression. To understand whether ribosomes indeed have an effect on protein production, we look at the CCs between ribosomal concentration and constitutive reporter concentration, and CCs between ribosomal production rate and constitutive production rate. The L19-C, pn25-Y concentration-concentration CC shows a very unclear pattern (figure 6.14.A), but the Prrna-C, pn25-Y concentration-concentration CCs show a positive peak around $\tau = 0$

delay in various conditions, indicating concerted fluctuations (figure 6.14.B-D). Positive correlations are expected, since concerted fluctuations were observed earlier for a pair of two constitutive reporters [9], for groups of proteins [141] and in general is expected to some extent for any two proteins because of the existence of extrinsic noise [13]. However, here, we look at two proteins of which one reports for ribosomal concentration, which might have an impact on protein production itself. This might change the dynamics. Additional features in the CC on top of the positive peak at 0 delay might tell us something about the ribosome-protein expression dynamics. For example, if we were to observe more pronounced correlations at positive delays this could indicate transmission of fluctuations from ribosomes to protein expression. However, The Prrna-C, pn25-Y concentration correlation for M9 medium (figure 6.14.B) seems to show the opposite: there is more positive correlation at negative delays. This implies that ribosome concentration fluctuations correlates with past protein concentration fluctuations. The two reporters might have different maturation times, which could lead to artificial correlations negative delay. Maturation times are however in the order of tens of minutes [18, 216], and a discrepancy between the two is thus not expected to cause such a big effect as observed here. A very speculative hypothesis is that fluctuations in protein abundance might positively regulate ribosomal production, which might help the cell anticipate ribosomal demand.

For TY medium (figure 6.14.C), the CC does not show a stronger correlation at negative delays. One could speculate this is because ribosomes become more limiting here, thus shifting the balance from negative delays in M9 medium to positive delays in TY medium. The CC for TY medium with antibiotics (figure 6.14.D) shows multiple peaks, but the negative control CCs also shows high correlation values. This indicates the second peak in the cross correlation curve at negative delays might be an artefact of insufficient data. More data is required to draw any conclusions for this condition.

The production-production CCs for the same conditions (figure 6.15) all appear to show peaks at 0 delay, that are rather symmetrical around the y-axis, and appear to not show additional features. This is consistent with aforementioned concerted protein expression fluctuations.

Results from strains with two ribosomal reporters

As mentioned, to get a wider view on ribosomal dynamics, we employed a strain which has labels on additional ribosomal proteins. This L9-R, S2-Y strain has a red fluorescent reporter on the L9 ribosomal protein, and a yellow reporter on the S2 ribosomal protein, see table 6.1. We performed measurements on this strain to obtain additional data on the single cell dynamics between ribosomal expression and growth. Furthermore, the presence of two ribosomal reporters in one strain allowed us to investigate the interaction between them. To investigate the effect of ribosomal stress on ribosomal dynamics, we analysed strain L9-R, S2-Y growing in steady state both in M9 minimal medium and M9 minimal medium with sub-inhibitory doses of the translational inhibitory antibiotic tetracycline (0.5 μM). To check the validity of the reporters, we first looked at the population level response of the reporter concentration to the presence of antibiotics. Figure 6.16 shows that due to the presence of antibiotics, the population average growth rate decreased by 12%, whilst the concentration of ribosomal reporter more than doubles.
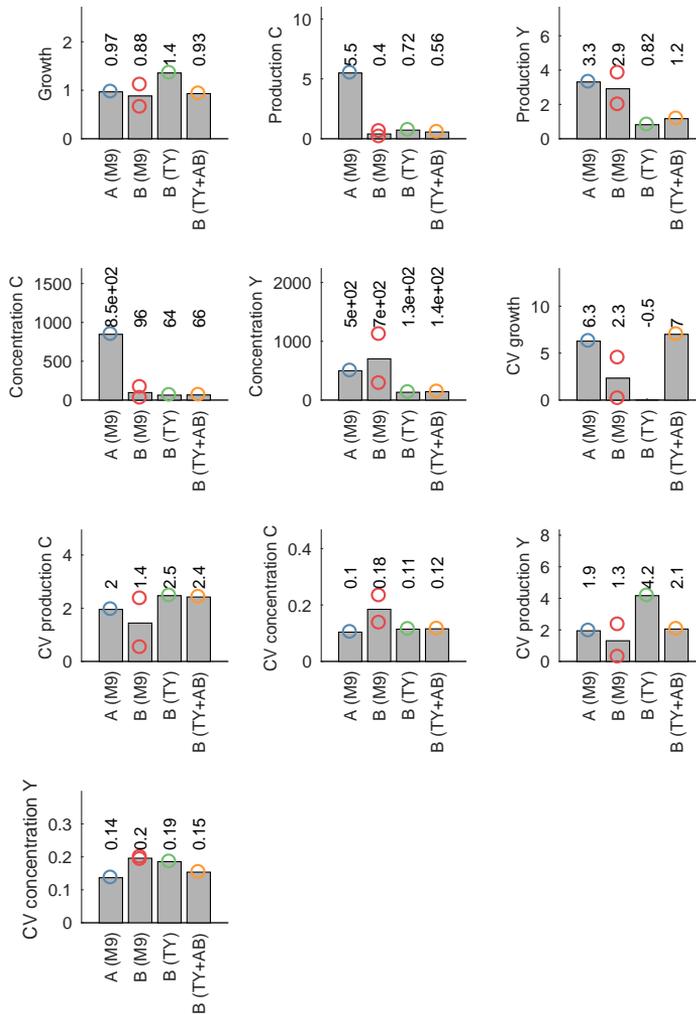
Figure 6.9: Population mean values of different parameters measured in different strains and conditions. In the labels, (A) indicates the L19-C, pn25-Y strain and (B) the Prrna-C, pn25-Y strain. Growth media are indicated between brackets. M9 indicates M9 minimal medium, TY indicates TY medium and TY+AB indicates TY medium supplemented with antibiotics. "C", "Y" and "R" indicate the fluorescent labels by their colors: mCerulean (cyan), mVenus (yellow) and mCherry (red) respectively.

Figure 6.10: Cross-correlations between concentration of constitutive reporter and growth. (A) L19-C, pn25-Y strain. Grown in M9 minimal medium. (B) Prrna-C, pn25-Y strain. Grown in M9 minimal medium. (C) Prrna-C, pn25-Y strain. Grown in TY medium. (D) Prrna-C, pn25-Y strain. Grown in TY medium supplemented with antibiotics. "C", "Y" and "R" indicate the fluorescent labels by their colors: mCerulean (cyan), mVenus (yellow) and mCherry (red) respectively. Both the black and red curves show the cross-correlation, based on two different calculation methods (see section 2.4.5). Grey areas and blue lines indicate respectively the range and mean of negative control curves (see text for discussion). Except for the left panel in panel B, which is based on a gel pad experiment, all data is based on experiments with microfluidic device 2 (see also methods section).

Figure 6.11: Cross-correlations between production rate of constitutive reporter and growth.
(A) L19-C, pn25-Y strain. Grown in M9 minimal medium. (B) Prrna-C, pn25-Y strain. Grown in M9 minimal medium. (C) Prrna-C, pn25-Y strain. Grown in TY medium. (D) Prrna-C, pn25-Y strain. Grown in TY medium supplemented with antibiotics. "C", "Y" and "R" indicate the fluorescent labels by their colors: mCerulean (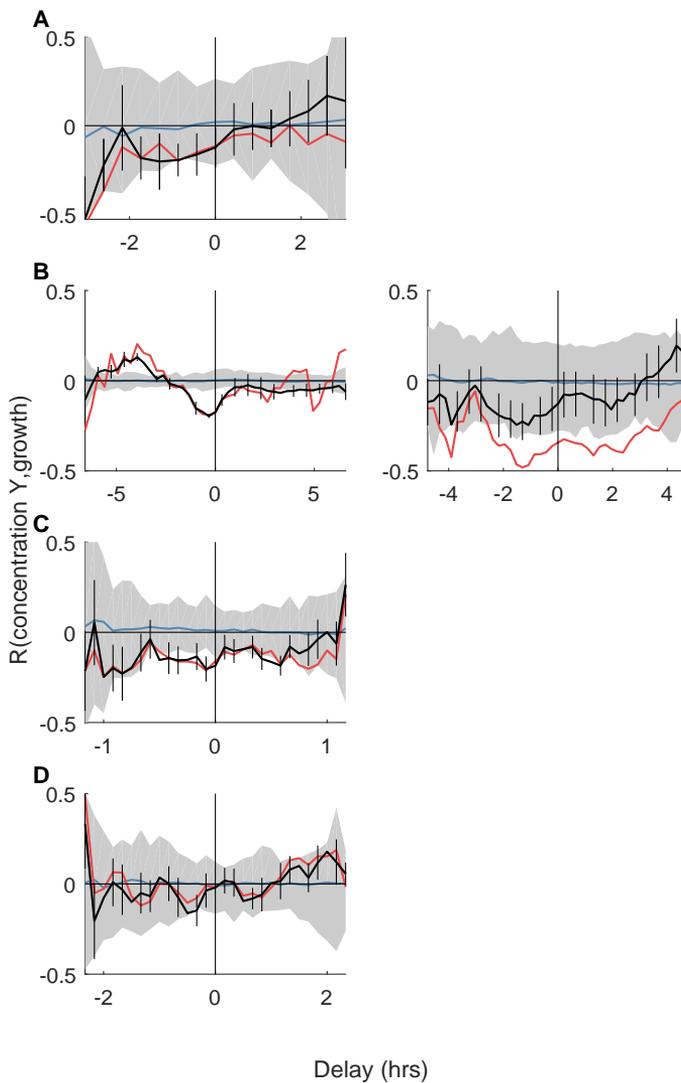cyan), mVenus (yellow) and mCherry (red) respectively. Both the black and red curves show the cross-correlation, based on two different calculation methods (see section 2.4.5). Grey areas and blue lines indicate respectively the range and mean of negative control curves (see text for discussion). Except for the left panel in panel B, which is based on a gel pad experiment, all data is based on experiments with microfluidic device 2 (see also methods section).

Figure 6.12: Cross-correlations between concentration of ribosomal reporter and growth. (A) L19-C, pn25-Y strain. Grown in M9 minimal medium. (B) Prrna-C, pn25-Y strain. Grown in M9 minimal medium. (C) Prrna-C, pn25-Y strain. Grown in TY medium. (D) Prrna-C, pn25-Y strain. Grown in TY medium supplemented with antibiotics. "C", "Y" and "R" indicate the fluorescent labels by their colors: mCerulean (cyan), mVenus (yellow) and mCherry (red) respectively. Both the black and red curves show the cross-correlation, based on two different calculation methods (see section 2.4.5). Grey areas and blue lines indicate respectively the range and mean of negative control curves (see text for discussion). Except for the left panel in panel B, which is based on a gel pad experiment, all data is based on experiments with microfluidic device 2 (see also methods section).
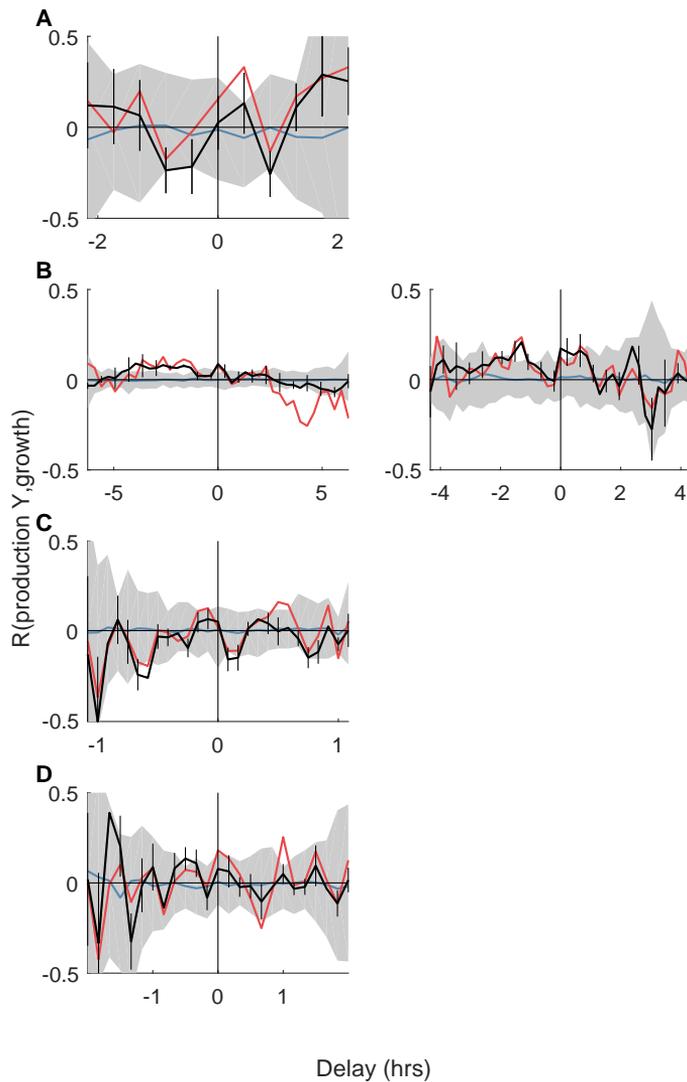
Figure 6.13: Cross-correlations between production rate of ribosomal reporter and growth. (A) L19-C, pn25-Y strain. Grown in M9 minimal medium. (B) Prrna-C, pn25-Y strain. Grown in M9 minimal medium. (C) Prrna-C, pn25-Y strain. Grown in TY medium. (D) Prrna-C, pn25-Y strain. Grown in TY medium supplemented with antibiotics. "C", "Y" and "R" indicate the fluorescent labels by their colors: mCerulean (cyan), mVenus (yellow) and mCherry (red) respectively. Both the black and red curves show the cross-correlation, based on two different calculation methods (see section 2.4.5). Grey areas and blue lines indicate respectively the range and mean of negative control curves (see text for discussion). Except for the left panel in panel B, which is based on a gel pad experiment, all data is based on experiments with microfluidic device 2 (see also methods section).
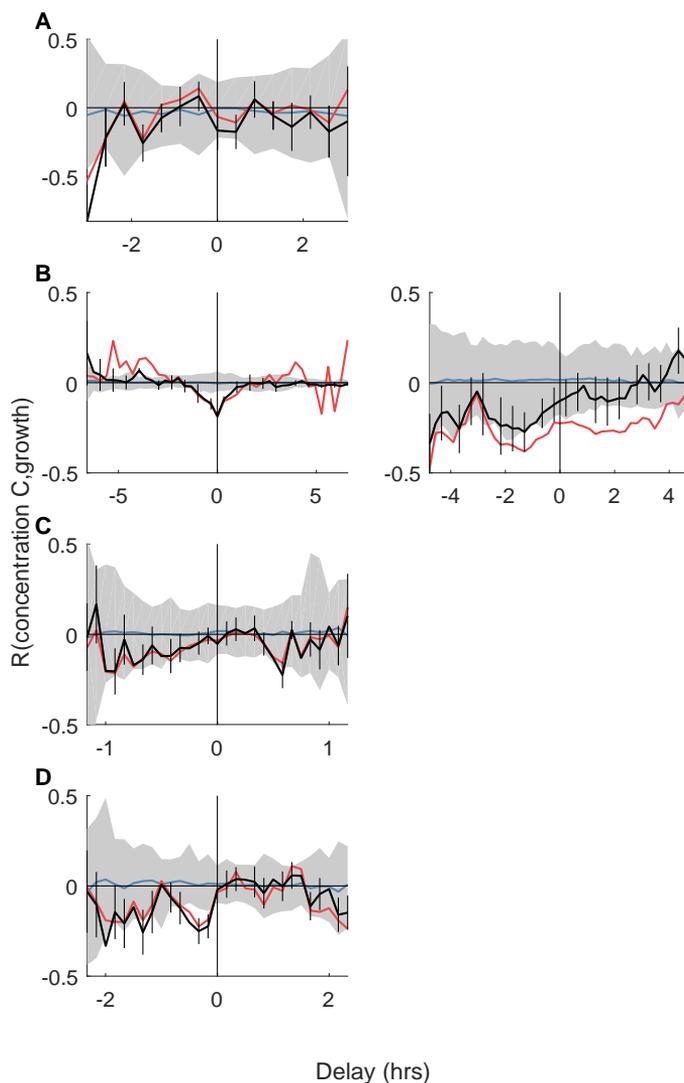
6

Figure 6.14: Cross-correlations between concentration of ribosomal reporter and concentration of pn25 reporter. (A) L19-C, pn25-Y strain. Grown in M9 minimal medium. (B) Prrna-C, pn25-Y strain. Grown in M9 minimal medium. (C) Prrna-C, pn25-Y strain. Grown in TY medium. (D) Prrna-C, pn25-Y strain. Grown in TY medium supplemented with antibiotics. "C", "Y" and "R" indicate the fluorescent labels by their colors: mCerulean (cyan), mVenus (yellow) and mCherry (red) respectively. Both the black and red curves show the cross-correlation, based on two different calculation methods (see section 2.4.5). Grey areas and blue lines indicate respectively the range and mean of negative control curves (see text for discussion). Except for the left panel in panel B, which is based on a gel pad experiment, all data is based on experiments with microfluidic device 2 (see also methods section).
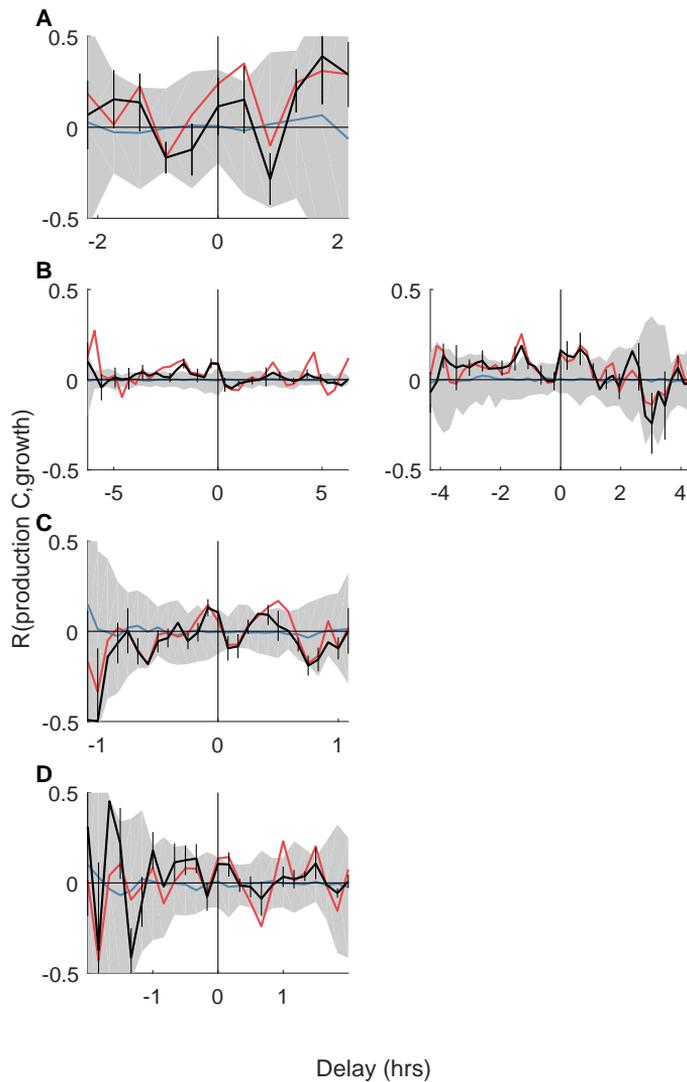
Figure 6.15: Cross-correlations between production rate of ribosomal reporter and production rate of pn25 reporter. (A) L19-C, pn25-Y strain. Grown in M9 minimal medium. (B) Prrna-C, pn25-Y strain. Grown in M9 minimal medium. (C) Prrna-C, pn25-Y strain. Grown in TY medium. (D) Prrna-C, pn25-Y strain. Grown in TY medium supplemented with antibiotics. "C", "Y" and "R" indicate the fluorescent labels by their colors: mCerulean (cyan), mVenus (yellow) and mCherry (red) respectively. Both the black and red curves show the cross-correlation, based on two different calculation methods (see section 2.4.5). Grey areas and blue lines indicate respectively the range and mean of negative control curves (see text for discussion). Except for the left panel in panel B, which is based on a gel pad experiment, all data is based on experiments with microfluidic device 2 (see also methods section).
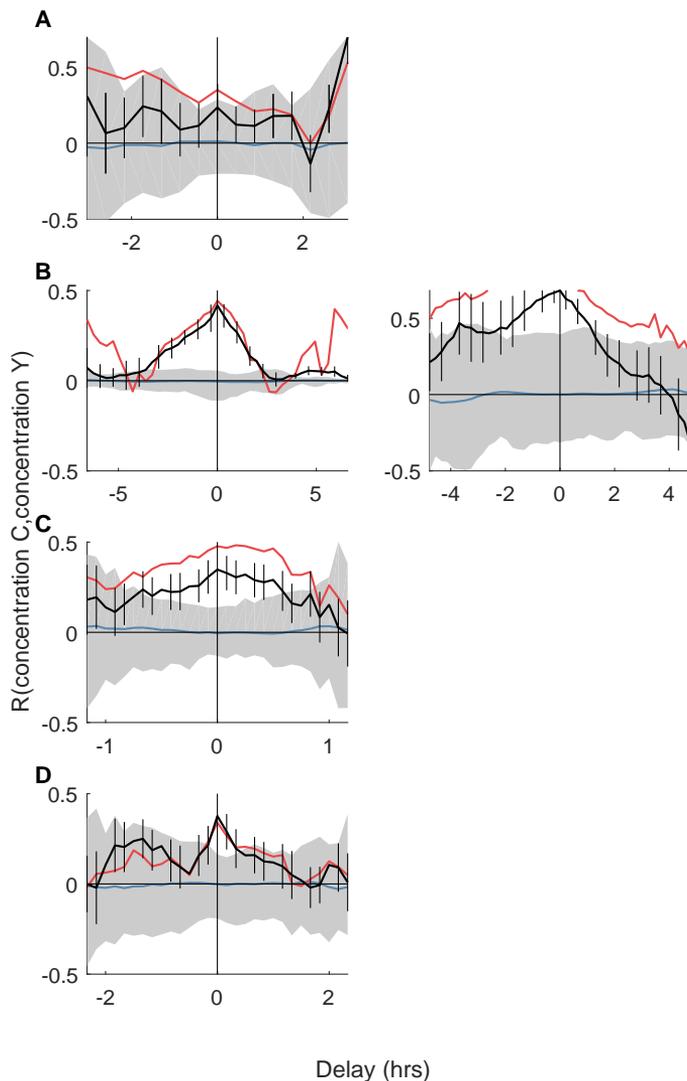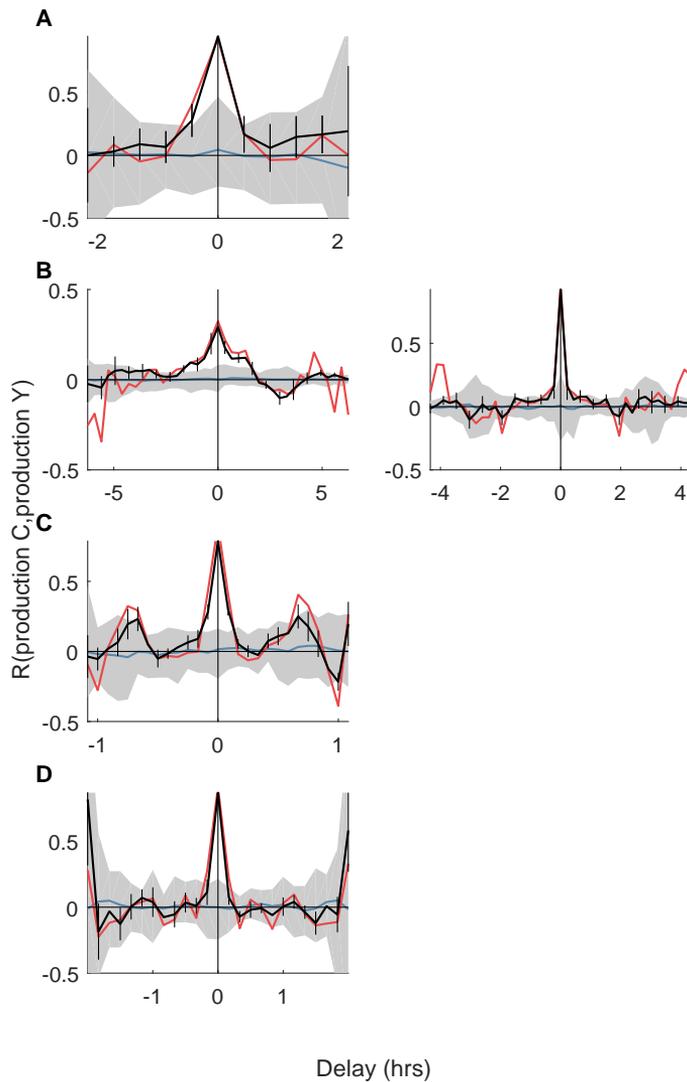
Such large increases of ribosomal concentration are consistent with earlier observations of growth in the presence of antibiotics, although in earlier observations this large increase was accompanied by a larger decrease in growth rate [140]. Nevertheless, earlier ribosomal labels even showed a decrease in concentration with increasing growth rate (see section 6.2.2), so these labels show behavior that is much more consistent with expectations. Figures 6.17 and 6.18 show CCs from the S2 r-protein expression-growth dynamics, and figures 6.19 and 6.20 show CCs from the L9 r-protein expression-growth dynamics. Furthermore, figures 6.21 and 6.22 show expression–expression CCs between the ribosomal labels. Just as we did for the strains discussed earlier, we also generated negative control CCs for these datasets using randomly combined time traces. These negative controls again showed high correlations (grey areas in figures 6.17-6.22) instead of very low values as expected for steady state experiments with sufficient amounts of data. This indicates that more experiments are required to make claims about this data.

Nevertheless, we can try to interpret the current data. The S2 reporter concentration-growth cross correlation for cells growing in M9 minimal medium (left panel in figure 6.17.A) shows negative correlations. Negative correlations at negative delays are consistent with the dilution mode of noise transmission [10], but this curve also shows negative correlations at positive delays, which is not consistent with that mode of noise transmission. The CC of second experiment with this strain in the same conditions (right panel in figure 6.17.A) shows less pronounced negative correlations, but we also note that cross correlation analysis of other parameters from this experiment showed indications of poor data quality (as discussed later, see also figure 6.21.A). Interestingly, the S2 reporter concentration-growth CC for cells growing in the same M9 medium supplemented with antibiotics, does not show these negative correlations, see figure 6.17.B. This could be indicative of a situation where ribosomes are more limiting due to the antibiotic stress, and fluctuations transmit from ribosomal fluctuations to the growth rate of the cell. Production rate-growth CCs of the S2 reporter, shown in figure 6.18, are harder to interpret, both for M9 medium and M9 medium with antibiotics. They are however not inconsistent with the earlier observations and interpretation.

Next, using the same experimental data, we calculated the CCs for the other ribosomal reporter in this strain to see whether it shows similar behaviour. The first experiment (in M9 medium) also showed negative concentration-growth CCs of the L9 reporter (figure 6.19.A, left panel), consistent with the CCs of the S2 reporter. Yet, the CC from the second experiment in M9 medium shows a different pattern; though we note that cross-correlation analysis of other parameters from this experiment showed indications of poor data quality (as discussed later, see also figure 6.21.A). The dynamics in M9 medium plus antibiotics again seemed different compared to the experiment without antibiotics (figure 6.19.B), as negative correlations now manifest as positive delays (instead of at negative delays). The negative correlations are however different from the positive correlations we saw for the S2 reporter in the same experiment. This indicates that either repetitions of these experiments might yield different results, or one of the labels does not appropriately reflect ribosomal concentration, or different ribosomal proteins might have their own dynamics. Production-growth CCs of the L9 reporter (figure 6.20) are again hard to interpret, although the condition with antibiotic stress shows a distinct peak around zero delay.

Aside from expression-growth dynamics, we can additionally look at the interaction between expression of the two ribosomal proteins in the experiments with this strain. In general, we expect CCs for this interaction to show positive correlations that are symmetrical around zero delay, since we expect both reporters to represent the ribosomal concentration. The first experiment (in M9 medium) showed an S2-L9 concentration-concentration CC that is — against expectations — not symmetric (left panel in figure 6.21.A). The second experiment in the same condition (M9 medium) showed an S2-L9 concentration-concentration CC that has a shape that we cannot interpret in our current framework (figure 6.21.A, right panel). The correlations in this CC also do not go to zero for large delays, which is indicative for poor data quality. If data quality is indeed poor for this experiment, this might also have implications for other CCs based on this experiment (all the right panels in figures 6.17.A, 6.18.A, 6.19.A and 6.20.A.) Furthermore, the S2-L9 concentration-concentration CC for the experiment with this strain in M9 minimal medium supplemented with antibiotics is different, but also not symmetric (figure 6.21.B). This asymmetry is consistent with the earlier differences in expression growth CCs between the two labels. Again, it might be that repetitions of these experiments might yield different results, or one of the labels does not appropriately reflect ribosomal concentration, or different ribosomal proteins might have their own dynamics.

Furthermore, the CCs between the production rates of the S2 and L9 labels (figure 6.21) do seem symmetric in both the M9 minimal medium and M9 minimal medium with antibiotics, which is consistent with aforementioned expectations. We do not observe further noteworthy features in the production–production CCs.

6.3. DISCUSSION AND CONCLUSION

6.3.1. ANTIBIOTIC SHIFT EXPERIMENTS ON DEVICE 1

In conclusion, we have used different single cell approaches in an attempt to further understand the ribosome. We first performed an experiment using our microfluidic device 1 to subject cells to a switch from growth in minimal medium to growth in minimal medium with antibiotics. Based on literature, we expected that cells needed a higher concentration of ribosomes when they were exposed to translational antibiotics (see e.g. ref. [170]), and thus that single cells that expressed more ribosomes relative to the population average would benefit shortly after the switch. If we would be able to observe such single cell growth advantages, this might imply ribosomes are limiting in this scenario. Their behaviour could then serve as positive control and point of reference for other experiments. To our surprise, the L31 reporter did not show the expected increased concentration during exposure to antibiotics after the switch. This was also observed before [18] for this particular ribosomal protein, and it raises the question how representative the L31 reporter is for the ribosomal concentration. A second reporter we used, a ribosomal RNA promoter fused to a cyan fluorescent reporter protein, did show the expected increase of ribosomal concentration after the switch to medium supplemented with antibiotics. The results from these experiments regarding single cell growth advantage were however inconclusive, they neither showed a clear absence nor a clear presence of an increased correlation after the switch.
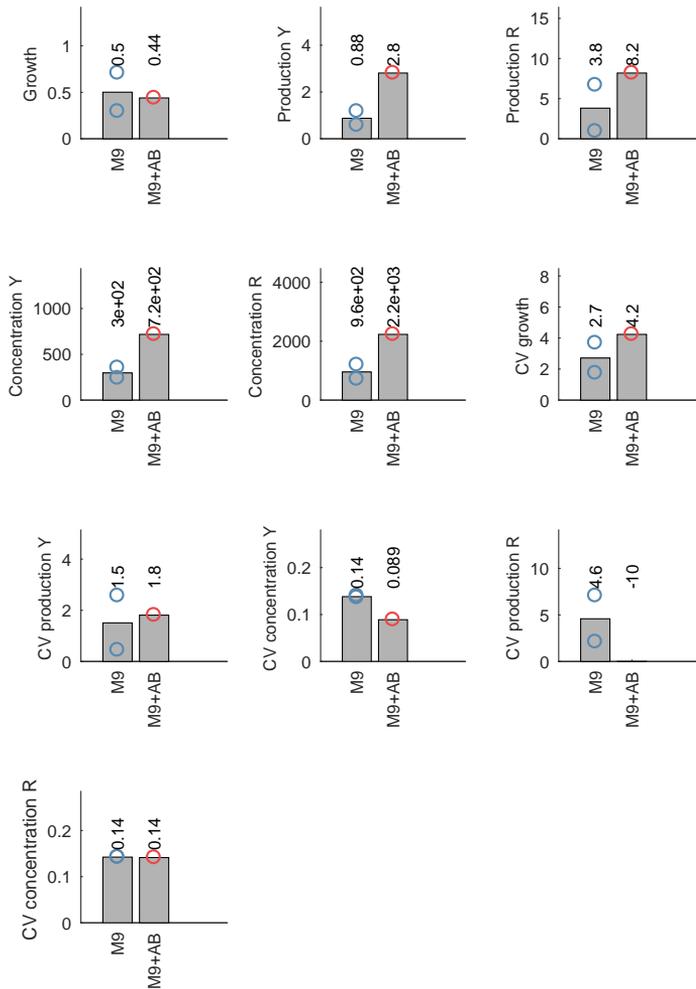
Figure 6.16: Population mean values of different parameters measured in L9-R, S2-Y strain experiments. Shown are data from growth in M9 minimal medium ("M9") versus growth in M9 minimal medium supplemented with antibiotics ("M9+AB"). "C", "Y" and "R" indicate the fluorescent labels by their colors: mCerulean (cyan), mVenus (yellow) and mCherry (red) respectively.

Figure 6.17: Cross-correlations between the concentration of the S2 ribosomal protein and growth. (A) L9-R, S2-Y strain. Grown in in M9 minimal medium. (B) L9-R, S2-Y strain. Grown in in M9 minimal medium supplemented with antibiotics. "C", "Y" and "R" indicate the fluorescent labels by their colors: mCerulean (cyan), mVenus (yellow) and mCherry (red) respectively. Both the black and red curves show the cross-correlation, based on two different calculation methods (see section 2.4.5). Grey areas and blue lines indicate respectively the range and mean of negative control curves (see text for discussion). All data is based on experiments with microfluidic device 2 (see also methods section).

Figure 6.18: Cross-correlations between the production rate of the S2 ribosomal protein and growth. (A) L9-R, S2-Y strain. Grown in in M9 minimal medium. (B) L9-R, S2-Y strain. Grown in in M9 minimal medium supplemented with antibiotics. "C", "Y" and "R" indicate the fluorescent labels by their colors: mCerulean (cyan), mVenus (yellow) and mCherry (red) respectively. Both the black and red curves show the cross-correlation, based on two different calculation methods (see section 2.4.5). Grey areas and blue lines indicate respectively the range and mean of negative control curves (see text for discussion). All data is based on experiments with microfluidic device 2 (see also methods section).
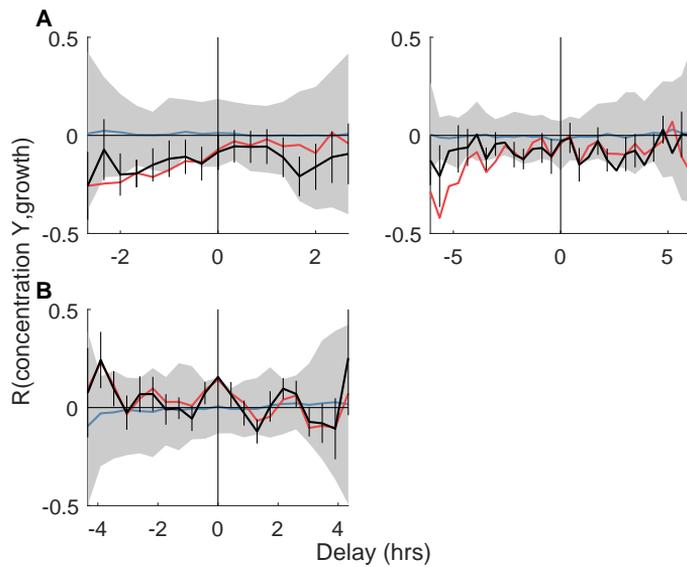
Figure 6.19: Cross-correlations between the concentration of the L9 ribosomal protein and growth. (A) L9-R, S2-Y strain. Grown in in M9 minimal medium. (B) L9-R, S2-Y strain. Grown in in M9 minimal medium supplemented with antibiotics. "C", "Y" and "R" indicate the fluorescent labels by their colors: mCerulean (cyan), mVenus (yellow) and mCherry (red) respectively. Both the black and red curves show the cross-correlation, based on two different calculation methods (see section 2.4.5). Grey areas and blue lines indicate respectively the range and mean of negative control curves (see text for discussion). All data is based on experiments with microfluidic device 2 (see also methods section).

6
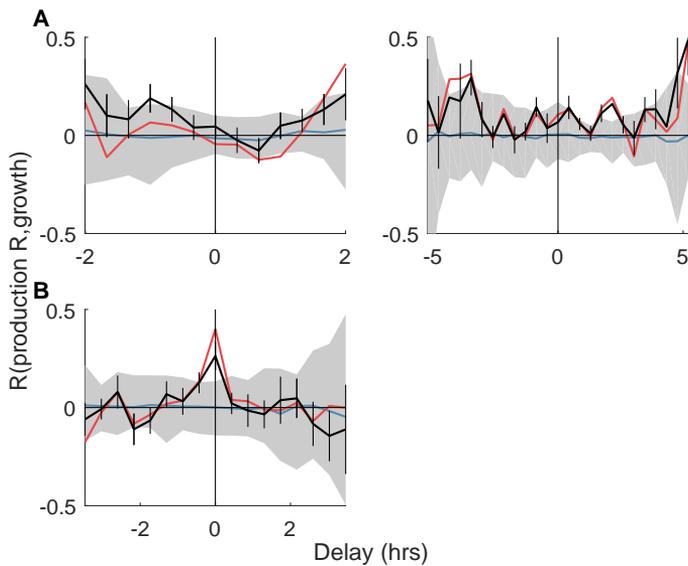
Figure 6.20: Cross-correlations between the production rate of the L9 ribosomal protein and growth. (A) L9-R, S2-Y strain. Grown in in M9 minimal medium. (B) L9-R, S2-Y strain. Grown in in M9 minimal medium supplemented with antibiotics. "C", "Y" and "R" indicate the fluorescent labels by their colors: mCerulean (cyan), mVenus (yellow) and mCherry (red) respectively. Both the black and red curves show the cross-correlation, based on two different calculation methods (see section 2.4.5). Grey areas and blue lines indicate respectively the range and mean of negative control curves (see text for discussion). All data is based on experiments with microfluidic device 2 (see also methods section).
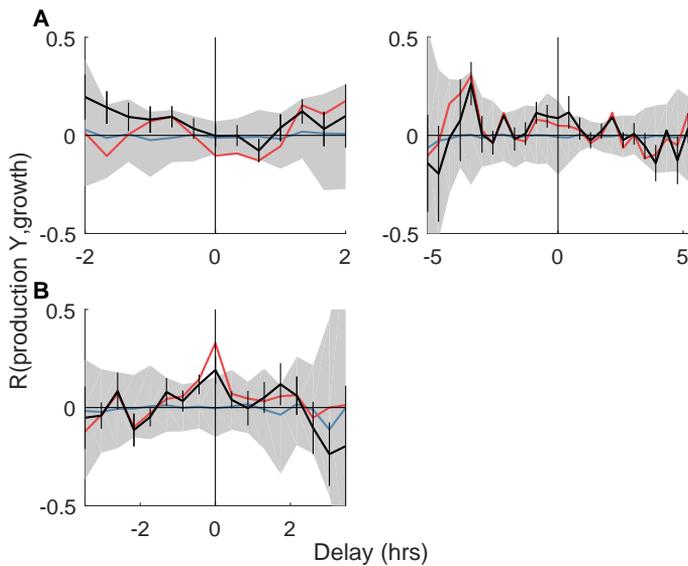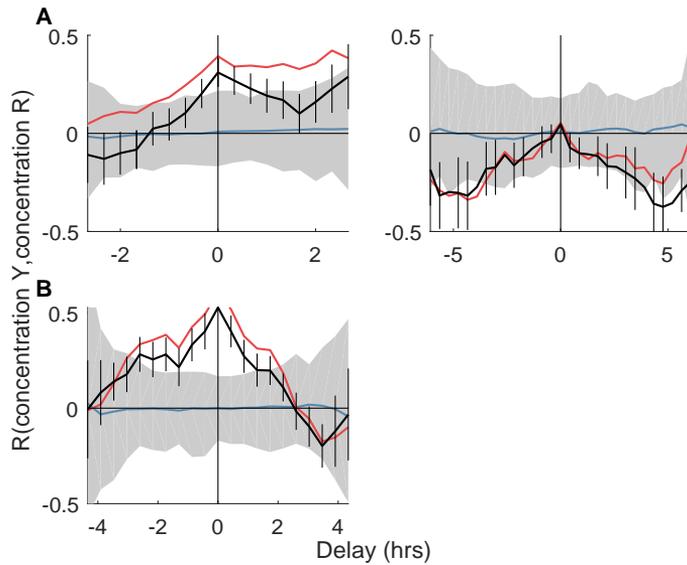
Figure 6.21: Cross-correlations between concentrations of two ribosomal proteins. (A) L9-R, S2-Y strain. Grown in in M9 minimal medium. (B) L9-R, S2-Y strain. Grown in in M9 minimal medium supplemented with antibiotics. "C", "Y" and "R" indicate the fluorescent labels by their colors: mCerulean (cyan), mVenus (yellow) and mCherry (red) respectively. Both the black and red curves show the cross-correlation, based on two different calculation methods (see section 2.4.5). Grey areas and blue lines indicate respectively the range and mean of negative control curves (see text for discussion). All data is based on experiments with microfluidic device 2 (see also methods section).
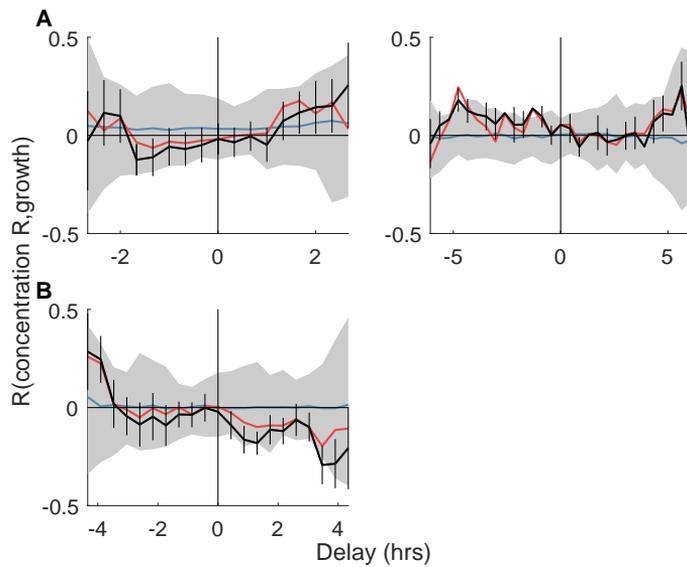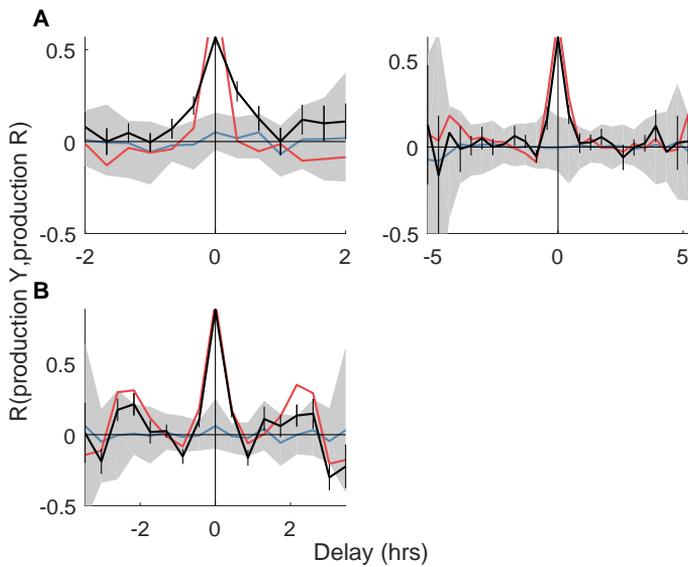
Figure 6.22: Cross-correlations between production rates of two ribosomal proteins. (A) L9-R, S2-Y strain. Grown in in M9 minimal medium. (B) L9-R, S2-Y strain. Grown in in M9 minimal medium supplemented with antibiotics. "C", "Y" and "R" indicate the fluorescent labels by their colors: mCerulean (cyan), mVenus (yellow) and mCherry (red) respectively. Both the black and red curves show the cross-correlation, based on two different calculation methods (see section 2.4.5). Grey areas and blue lines indicate respectively the range and mean of negative control curves (see text for discussion). All data is based on experiments with microfluidic device 2 (see also methods section).

6.3.2. SINGLE CELL RIBOSOME EXPRESSION AND GROWTH

To follow up on these experiments, we performed experiments using microfluidic device 2, which allows for measurements that are longer and thus allows for more switches, and additionally enables tracking a higher and constant number of cells. This set of experiments could have provided additional information on the ribosomal behavior during switches from medium that requires less ribosomal expression to medium that requires more ribosomal expression.Additionally, the long time lapse capabilities of the device would allow us to record data in steady state. On top of this, we used newly constructed strains in these experiments that allowed us to measure the effect of ribosomal fluctuations on single cell protein expression, as the strains also contained a constitutive pn25 reporter construct. However, the analysis of data in this device presented additional challenges for the analysis. We therefore only analysed a subselection of the data and focused on analyses of steady state behaviour.

On top of this, the data that we did analyse showed a very high background signal, which indicated that more data needs to be analysed or additional experiments are required to draw definitive conclusions. Nevertheless, we made several observations regarding the interaction between single cell ribosomal fluctuations and single cell growth rate fluctuations. Firstly, in experiments involving the ribosomal RNA reporter, we saw that cross-correlation analysis in most conditions could be consistent with dilution mode transmission of noise. This would be consistent with the idea that ribosomal RNA does not affect growth rate, but growth rate fluctuations instead result in ribosomal fluctuations. Secondly, in experiments in M9 minimal medium involving labeled ribosomal proteins (L9 and S2), we also saw cross-correlations that might be consistent with dilution mode dynamics. However, when this strain grew on M9 minimal medium supplemented with antibiotics, cross-correlations were less consistent with dilutions mode dynamics. Hence, contrary to what the r-RNA data indicated, this data suggested that ribosomal fluctuations could become more consequential for the cell when it experiences translational stress.

Similar experiments were described earlier by Noreen Walker [18], we briefly summarized her experimental observations in section 6.1.3 of this chapter. Observations by Walker on a strain with a ribosomal RNA reporter in medium M9 medium supplemented with acetate (our lactose-supplemented M9 sustains a higher growth rate) or defined rich medium did not show dilution mode behaviour, which could mean ribosomal RNA shows different dynamics in different media. Some of the experiments involving L19 and L31 reporters that she performed showed dilution mode behavior in M9 minimal medium, similar to our observations on the labeled S2 and L9 ribosomal proteins in minimal medium, but some of her experiments did not. (This depended on the fluorescent label that was used, mCherry or mCerulean respectively.) Walker also performed CC analyses of steady state growth in rich medium supplemented with antibiotics, which resulted only in subtle changes in the dynamics, both for a labelled L31 ribosomal protein and for a ribosomal RNA promoter reporter construct. It is unclear how these results obtained in different media connect to our results.

6.3.3. SINGLE CELL RIBOSOME EXPRESSION AND PROTEIN EXPRESSION

In this chapter, we also performed experiments on strains that carried both a ribosomal expression reporter and a constitutive reporter. The latter reporter was intended to probe the effect of ribosomal expression fluctuations on the cell's ability to produce proteins. We observed positive correlations in the CCs between the concentrations of these two reporters. Such correlations are however expected between the expression of any two proteins, and thus do not necessarily imply that there is an effect of ribosomal expression fluctuations on protein production in the cell. We did observe additional features in the CCs. These depended on the medium in which the cells were growing. In most media, these features were not consistent with an effect of ribosomal fluctuations on protein production, except in TY. In TY medium we observed additional correlations between constitutive protein expression and past ribosome expression, which is consistent with an effect of ribosomal fluctuations on protein production rates. TY medium sustains very fast bacterial growth rate, which could explain an increased sensitivity to ribosomal fluctuations. As mentioned however, more experiments are needed to draw definitive conclusions.

To shed further light on these matters, one could additionally look at the cross-correlation between the concentration of the ribosomal reporter and the production of the constitutive reporter. This would require a rather straightforward extension of our analysis scripts.

6.3.4. THE RIBOSOME IS A COMPLEX STRUCTURE

We have discussed many experiments, which employed different ribosomal labels and probed dynamics in different conditions. Different conditions sometimes showed different results. Often however, results also depended on which ribosomal protein was tracked. One explanation for that observation is that ribosomal subunits have their own dynamics. Though the ribosome is often viewed as a complex that has a fixed composition, recent discoveries in eukaryotes show that this might actually not be the case [217, 218]. Purification and mass-spectrometry of ribosomes in yeast and mouse embryonic stem cells showed that pools of ribosomes with different ribosomal protein composition exist, which could also be linked to functional properties [218]. This raised the suggestion that the cell can modulate expression of different ribosomal proteins separately, to achieve specific regulatory goals. It is not impossible this is also the case in *E. coli*, and this would explain why experiments tracking different ribosomal proteins show different dynamics. This potential added layer of regulation makes it more difficult to understand the effect of ribosomal concentration fluctuations. Coincidentally, for eukaryotes, it has been observed that especially ribosomal surface proteins vary. Surface proteins are also convenient targets for labelling. In fact, the L9, S2 and L19 proteins that we labelled are also on the surface of the ribosome, see figure 6.23. Also the L31 protein that we labelled is only loosely attached to the ribosome [18]. Consistent with the view that the ribosomal proteins are not expressed equally for functional reasons, not all operons from which the ribosomal proteins are expressed are regulated by exactly the same regulators [171]. Additionally, some ribosomal subunits are observed to be essential, whilst others are not (i.e. some respective null mutants are viable, whilst others are not). This further illustrates that different ribosomal proteins might have different functional expression
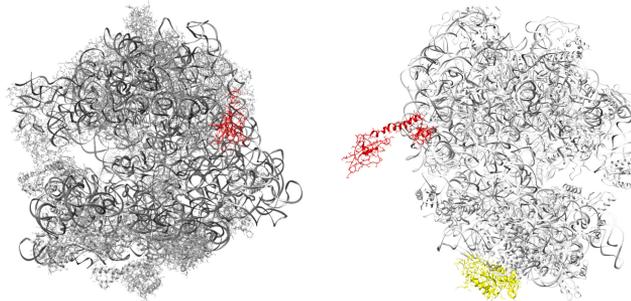
Figure 6.23: Some of the ribosomal proteins that were labeled for experiments in this chapter.
The structure of the ribosome is displayed twice in grey, with on the left the L19 ribosomal protein
highlighted in red, and on the right the L9 and S2 ribosomal subunits highlighted in red and yellow,
respectively. This image is based on x-ray crystallography (PDB ID: 4v4q, [208]). Pdb files were
downloaded from `www.rcsb.org` [209] visualized with UCSF Chimera (version 1.11.2, build 41376,
[210]). These two visualizations were made by Giulia Bergamaschi.

profiles. From the proteins that we labelled, L9 and L31 are non-essential, and S2 and
L19 are essential. This distinction might also be reflected in different relationships be-
tween expression and growth rate for the different ribosomal proteins, though our data
might be too limited to draw conclusions about this. Concluding, understanding ribo-
somal dynamics based on labelling ribosomal proteins might prove a challenging task.

6.4. OUTLOOK

6.4.1. MAKING RIBOSOMES LIMITING

As mentioned in the results section, it can sometimes be very useful to create an ex-
treme situation to better understand the dynamics of a system. For ribosomal dynam-
ics, an extreme situation might be one where ribosomes are limiting, i.e. a situation in
which single cell growth rates correlate strongly with ribosomal expression.This topic
is also discussed in the previously mentioned thesis by Walker [18]. We think it is not
straightforward to create such a situation. For example, whilst it might be a sensible idea
that adding sub-inhibitory concentrations of antibiotics to the cellular growth medium
might cause translational stress and thus more limiting behaviour of ribosomes, the ad-
dition of antibiotics to growth medium was observed to have limited consequences in
experiments. One explanation for this might be that a population of cells reacts to an
antibiotic disturbance by altering its ribosomal expression level in such a way that lo-
cal deviations from the newly established level have similar effects as deviations from
the original expression level. This idea is further illustrated by the cartoon in figure 6.24.
Different stressors instead of antibiotics might be used to put the cell under translational
duress. To create an alternative situation where ribosomal capacity is challenged, we
constructed strain ASC1088 (see table 6.3) which carries a plasmid that expresses the
mCherry protein at a high expression level. Similar to the antibiotic treatment, protein
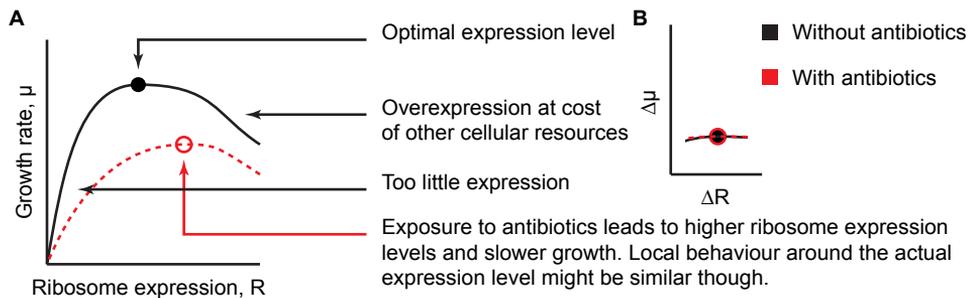
Figure 6.24: Cartoon of how ribosomal expression might affect growth rate. Ribosomal expression in single cell deviates from the population average. Both the average and single cell expression levels might not lie exactly at the optimum. Low expression might lead to slower growth because ribosomes become limiting. High expression might decrease growth rate because resources to produce these ribosomes are drawn from other essential cellular processes. Panel (a) shows how the expression-growth relationships in single cells might relate between cells not exposed to antibiotics (black line) and cells exposed to antibiotics (red line). Panel (b) shows that these relationships might be similar if we consider only the deviations from the population averages (which are here assumed to coincide with optimal expression levels). R and μ respectively indicate ribosomal expression and growth rate, ΔR and $\Delta \mu$ indicate mean-subtracted respective values.

over–expression might also put a large burden on ribosomal capacity, thus increasing the effect ribosomal fluctuations have on growth rate (see Walker's thesis for a longer discussion on this topic [18]). However, partially based on earlier observations of the limited effect of unneeded protein expression [219], we suspected that experiments with over-expression might yield similar results as the antibiotic experiments, we abandoned experiments with this strain.

Future experiments might be devised in which ribosomes are forced to become limiting. As also mentioned by Walker [18], previously constructed *rrs* null mutants [220–224] might provide a venue towards this aim. Given the role of the ribosomal RNA as backbone for the ribosome, shortages in ribosomal RNA content might have profound effects on the cellular growth rate. Ribosomal RNA is expressed from multiple partially redundant operons labelled *rrna* up to *rrnh*, as shown in supplementary figure 6.25. Bollenbach et al. [223] created several null mutants which had more and more of these operons removed, as also shown in supplementary figure 6.25. In these null mutants, ribosomal RNA fluctuations might have a large effect on growth rate. Ribosomal proteins on the other hand might still be expressed at normal levels, and thus would be present at plentiful levels. Hence, the concentration of ribosomal RNA might be limiting, whilst the concentration of ribosomal proteins is not. The expression of ribosomal RNA is however hard to track directly. One way for future experiments to get around this might be the creation of a plasmid that expresses ribosomal RNA and a reporter simultaneously. A fluorescent reporter sequence (and ribosome binding site) could be fused at the end of an *rrn* operon sequence, which could be expressed from a plasmid in one of the deletion strains. This would result in both rRNA and coding RNA being expressed simultaneously, and thus allow for a way to track whether fluctuations in ribosomal content can affect

single cell growth rate. Some research on how such a construct could look was already performed and can be found in supplementary note I (section 6.7) to this chapter.

Furthermore, single cell methods are developing rapidly, and future advances might help us to further understand the complex dynamics between the ribosome, protein production and cellular growth rate.

6.5. METHODS

Where minimal medium is mentioned, this refers to M9 minimal medium (also described in chapter 3). Minimal medium in this chapter was always supplemented with uracil (0.2 mM), lactose (0.1% g/mL). TY medium is described in chapter 3. During time lapse experiments under the microscope, tween20 (0.001% g/mL) was added to the medium additionally. When antibiotics are mentioned, this refers to tetracycline, which was used at a concentration of 0.5 µM. All experiments were conducted using microfluidic device 2 as described in the methods section of chapter 3, except for one experiment (this is indicated in the figure caption), which was performed using a gel pad (see also chapter 3). See also chapter 2 for further information.

6.6. ACKNOWLEDGEMENTS

6

6.7. Supplementary note I

6.7.1. Making rRNA concentration limiting and traceable

In this supplementary section, we aim to show how one could create an artificially limiting situation involving ribosomal genes and also track ribosomal expression to show this. Or in other words, how to create a strain in which traceable ribosomal RNA fluctuations have an effect on the growth rate. As briefly described in the main text, the Δ*rrn* strains created by Bollenbach et al. [223] could experience ribosomal RNA limitations. One could introduce a plasmid containing a ribosomal RNA operon sequence, followed by a ribosomal binding site sequence and fluorescent reporter sequence, to investigate correlations between additional ribosomal RNA expression (from the plasmid) and single cellular growth rate. To achieve this, one could either use the pK4–16 plasmid constructed by Quan et al. [225] — which contains the *rrnb* operon — as a starting point, or clone the *rrnb* operon into another vector (for which one could use primers sequences for *rrnb* also published by Quan et al [225]). The new vector should have a low copy number and its antibiotic resistance module should not be the same as the modules used in strains the vector might be introduced to (ampicilin might be a convenient choice, but not streptavadin or chloramphenicol). To track rRNA expression one could add a fluorescent reporter (e.g. mCerulean2.0) to the *rrnb* operon, which could be inserted after the coding sequences and accompanying RNA cleavage sites, but before the operon's terminator sites. The structure of the operon and the cleavage and terminator sites are described in Kaczanowska et al. [226] (see also Shajani et al. [227]), and terminator sites are also annotated in the Ecocyc database[1] [171]. In addition to a fluorescent protein sequence itself, the reporter sequence would also have to contain a ribosome binding site, which could be synthetically made using the Salislab ribosome binding site (RBS) calculator[2] [228, 229]. Furthermore, to prevent interference with ribosomal RNA maturation, a linker could be introduced between the ribosomal RNA sequences and the reporter sequence. The resulting plasmid could be introduced in strain Δ2 (also known as CGSC 12340 and SQ40), which lacks the *rrne* and *rrng* operons that are both similar to *rrnb*, see table 6.25. The copy number of a low copy number plasmid would likely be variable in single cells, which aids towards the goal of investigating the effect of different single cell expression levels of the ribosomal RNA.

6.7.2. Additional considerations

An additional consideration might be the strength of the *rrn* promoters. These are different per operon, the *rrne* is by far the strongest [230]. The *rrnb* promoter strength is two to several times less depending on the condition. This is compatible with our aim to keep ribosomal RNA expression limited.

　　A challenge to this strategy might be that tRNA turns out to be limiting, as suggested

---

[1]In the *E. coli* genome, the region with the terminators lies between nucleotide 4171756 and 4171760, and an annotated sequence can be found via `https://ecocyc.org/ECOLI/select-gen-el`.

[2]Using `https://salislab.net/software/` and the mCerulean2.0 sequence, an RBS sequence might look like:
`CTTTCAATCAATTCGTAATTTTAGATAAATAATAAGGAGGTCTCTAA.`

	Ribosomal parts		Regulated by	
rrnA	rrsA-ileT-alaT-rrlA-rrfA: 16S, tRNA-Ile, tRNA-Ala, 23S, 5S	‡	ppGpp, DksA, Lrp, Fis Sigma70, sigma32	✕✕✕✕
rrnB	rrsB-gltT-rrlB-rrfB 16S, tRNA-glu, 23S, 5S	*	ppGpp, DksA, Lrp, Fis, HNS Sigma70, sigma32	✕✕ ✕
rrnC	rrsC-gltU-rrlC-rrfC 16S, tRNA-glu, 23S, 5S	*	ppGpp, DksA, Lrp, Fis, [H-NS?] Sigma70, sigma32	
rrnD	rrsD-ileU-alaU-rrlD-rrfD-thrV-rrfF 16S, tRNA-ile, tRNA-ala, 23S, 5S, tRNA-thr, 5S		ppGpp, DksA, [Lrp?], Fis, [H-NS?] Sigma70, sigma32	✕✕✕
rrnE	rrsE-gltV-rrlE-rrfE 16S, tRNA-glu, 23S, 5S	*	ppGpp, DksA, [Lrp?], Fis, (PhoP) Sigma70, sigma32	✕✕ ✕✕
rrnG	rrsG-gltW-rrlG-rrfG 16S, tRNA-glu, 23S, 5S	*	ppGpp, DksA, [Lrp?], Fis, [H-NS?] Sigma70, sigma32	✕✕✕✕✕
rrnH	rrsH-ileV-alaV-rrlH-rrfH 16S, tRNA-ile, tRNA-ala, 23S, 5S	‡	ppGpp, DksA, [Lrp?], Fis, [H-NS?] Sigma70, sigma32	✕✕

Figure 6.25: Table showing the *rrn* operons in *E. coli* with cartoon overlay indicating different *rrn* null mutant strains created by Bollenbach et al.. Ribosomal RNA in *E. coli* is expressed from multiple partially overlapping operons. Operons contain different ribosomal parts (and tRNA molecules). *rrs* encodes 16S rRNA, *rrl* encodes 23S rRNA, and *rrf* encodes 5S rRNA. Other abbreviations indicate genes that encode tRNA molecules. This is also indicated in the second line of the "ribosomal parts" column. Symbols in the third column indicate equivalence between the operons. The last column indicates regulators of the operons, symbols in brackets indicate evidence is not definitive. (This information was obtained from the online database Ecocyc [171]). The overlayed crosses and circles indicate the null mutants that were constructed by Bollenbach et al. [223]. Each column of crosses corresponds to one strain, a cross indicates that the respective strain lacks the operon indicated. A circle indicates additional tRNA was expressed from a plasmid. These strains were also referred to as Δ1-6, where the number corresponds to the number of operons that were removed.

by Quan et al. [222]. tRNA is also expressed from the *rrn* operons. To make sure we are not mistaking tRNA limitations for ribosomal limitations, additional artificial tRNA expression (for example from plasmid ptRNA67 described in Zaporojets et al. [222, 231]) could be a tool to make sure this is not the case.

6.8. Supplemental figures and tables

6

r-protein	Gene name	Operon	Fluorescent label
L31	rpmE	none	Cerulean, mCherry
L19	rplS	rpsP-rimM-trmD-rplS	Cerulean, mCherry
L9	rplI	rpsF-priB-rpsR-rplI	mCherry
S2	rpsB	ttf-rpsB-tsf	Venus

Table 6.2: Labeld ribosomal proteins in the Tans lab.

Figure 6.26: Growth of single cells of the different strain populations in the different conditions.
(A) L19-C, pn25-Y strain. Grown in M9 minimal medium. (B) Prrna-C, pn25-Y strain. Grown in M9 minimal medium. (C) Prrna-C, pn25-Y strain. Grown in TY medium. (D) Prrna-C, pn25-Y strain. Grown in TY medium supplemented with antibiotics.

ASC number	Shorthand	Description	Source
ASC656	L31-R	L31-mCherry-kanR (no linker). (Kanamycin resistant.)	NW, VS
ASC657	L19-R	L19-mCherry-kanR (no linker). (Kanamycin resistant.)	NW, VS
ASC680	L31-R, Prrn-G	L31-mCherry-kanR (no linker), Δ(cheZ)::Prrn-GFP-catR. (Kanamycin and chloramphenicol resistant.)	NW, VS
ASC779	Prrn-G	Δ(cheZ)::Prrn-GFP, rrsa promoter. (No resistance.)	NW, VS
	L19-C	L19-mCerulean L19-gc-mCerulean-kanR (GC linker) (Kanamycin resistance.)	VS
	Prrn-C	Δ(cheZ)::Prrn-mCerulean-kanR (Kanamycin resistance)	VS
ASC1088	mCherry+	ASC968 + high copy plasmid w. pn25-mCherry (Resistances: kanamycin, chloramphenicol, ampicilin.)	VS

Table 6.3: Miscellaneous strains. NW indicates these strains were used in Noreen Walker's thesis [18]. VS indicates these strains are created by Vanda Sunderlikova.
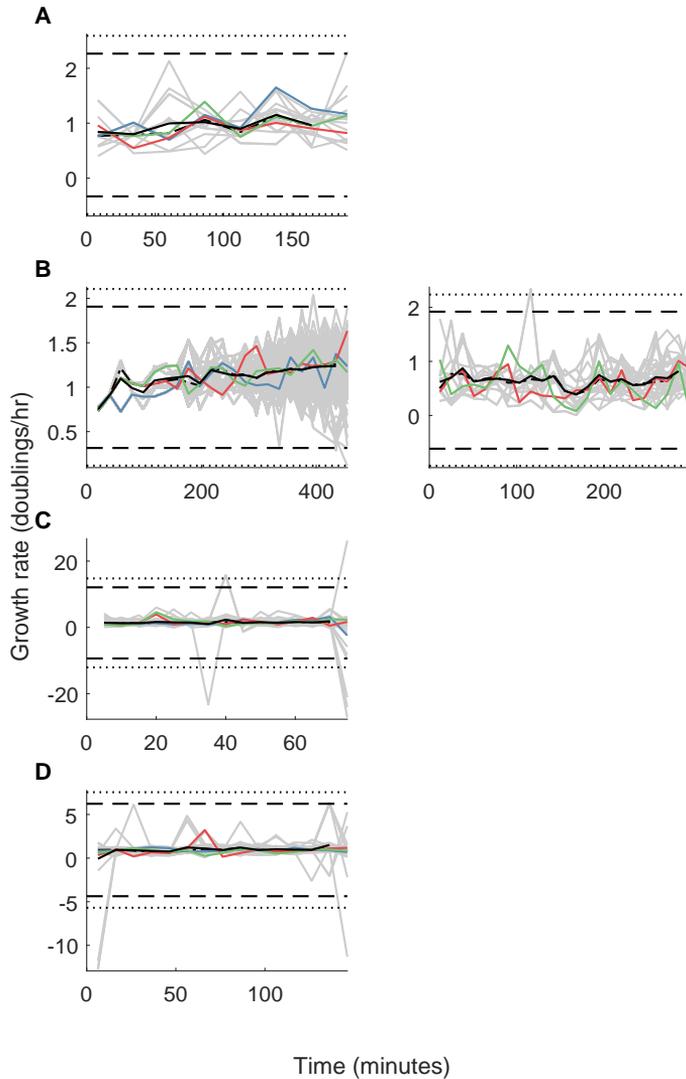
6
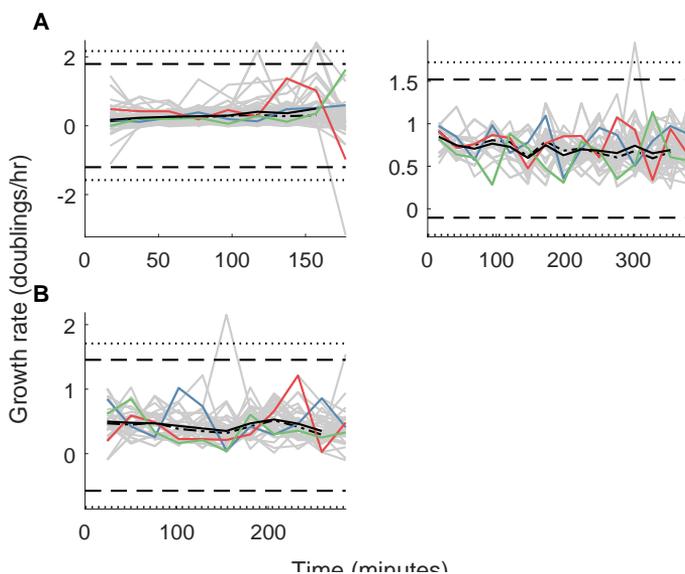
Figure 6.27: Growth of single cells of the different strain populations in the different conditions.
(A) L9-R, S2-Y strain. Grown in in M9 minimal medium. (B) L9-R, S2-Y strain. Grown in in M9 minimal medium supplemented with antibiotics.

7

IMPLICATIONS FOR SOCIETY

This chapter discusses scientific and technical implications of the research in this thesis for society.

7.1. AN UNDERSTANDING OF THE FUNDAMENTALS

All organisms, from bacteria to bird, from jellyfish to jungle tree, consist of cells. Some simple organisms, like bacteria, consist of only one cell. Some organisms consist of many more cells. Humans for example consist of around ten trillion cells (a one with thirteen zeroes) [232, BNID 102390]. A cell consists of a membrane that seperates the outside environment from its inside environment, which is filled with proteins, other biomolecules like signalling molecules and building blocks, and DNA which stores genetic building plans. Interactions between proteins and other cellular molecules, combined with sensory input from the environment outside the cell, determine how the cell behaves, what molecules are produced in the cell and whether it will grow. Like in an electronic circuit board, the properties of the components and their interactions can be identified and quantified, such that the workings of the cell can be understood and manipulated. Currently, we do not understand all the processes that go on in the biological cell. Quantitative biology approaches like this work allow us to further chart the processes that go on in the cell, where in this case we have used the bacterium *Escherichia coli* as a model organism. The aim is to obtain a better understanding and ability to manipulate the biological processes in the biological cell. Since cells are building blocks of all organisms, including humans and pathogenic microorganisms, this eventually might bring us closer to understanding and solving human disease. Additionally, microorganisms play an important role in industry, for example in food production and sewage treatment. What exactly the direct implications of understanding fundamental cellular processes are is hard to predict, but such knowledge has the potential to have a wide impact on society.

7.2. KILLING BACTERIA

Nevertheless, we will try to reflect more specifically on the impact of different aspects of this work. In chapter 3, we described how bacteria manage resuming their divisions after they paused the division process due to situations they experienced as stressful. This is thought to be a bacterial survival mechanism. Knowledge about bacterial survival mechanisms might be useful in situations where from a human perspective it is desired to kill or inhibit bacterial growth. Though bacteria filament in response to many adverse conditions, practically relevant conditions include antibiotic exposure and high or low temperatures. These conditions are relevant in clinical and food preservation settings, respectively. Information about the filamentation process might improve predictions about bacterial survival in these settings, and eventually might help adjusting clinical treatments or preservation processes to increase their effectiveness. Furthermore, targeting the components of the filamentation mechanism specifically might decrease bacterial proliferation in such settings.

7.3. METHODS WITH A WIDER RELEVANCE

Besides division of filamentous bacteria, there are two more topics in this thesis. These are: (1) how stochasticity, regulation and individuality of bacteria relate to each other, investigated in the CRP metabolic regulatory system, discussed in chapter 5, and (2) how stochasticity in ribosomal expression relates to bacterial individuality and growth, discussed in chapter 6. Firstly, it is noteworthy to mention that the methods used in the studies of these two topics (see chapter 2) could have a societal relevance. Automated image analysis of large quantities of individual cells as used in this thesis has wide applications, and experience from this type of analysis might also prove useful in a more clinical settings. Conceivable examples include analysis of blood cells or other tissue samples for the purpose of diagnostics. Furthermore, growing biological cells in microfluidic chambers in order to obtain single cell data is also a technique that can be applied more broadly. Single cell data can be and is already being acquired from human cell lines, which might improve our understanding of human biology and disease. Improvements in single cell microfluidic methods might contribute to the advancement of this research.

7.4. THE IMPORTANCE OF HETEROGENEITY

One can speculate about the more direct, albeit future, impacts on society of the work described in chapters 5 and 6. In chapter 5 we saw that regulation systems might be important in transmitting stochastic fluctuations, which contributes to heterogeneity in the population. In chapter 6 we investigated the role of ribosomes in cellular individuality, which relates immediately to population heterogeneity. Understanding the mechanisms behind heterogeneity gives us fundamental insights in the workings of the biological cell, but heterogeneity might also have a more direct relevance. For example, microorganisms are also used in industrial applications. Practical examples include the production of yoghurt and beer. Perhaps less well known is the effort to genetically modify organisms to produce fuels [233, 234]. In such applications, microorganisms are

usually employed to perform a single task. Like the production of a specific beer, or a specific molecule that can be used as fuel. There likely is an optimal state for an individual microorganism to perform such a task, but the phenomenon of heterogeneity prevents cells in the population to uniformly be in that state. Thus, whilst heterogeneity might be beneficial from an evolutionary perspective, genetic engineering to alter this phenomenon could potentially improve the yield and/or allow better fine-tuning of industrial processes that involve microorganisms.

Furthermore, as also discussed in chapter 4, heterogeneity might contribute to enhance the survival chances of a population, because different individuals can be prepared for different future scenarios (bet hedging). Knowledge about the origins of heterogeneity might thus be relevant in a clinical setting, where the desire often exists to prevent the growth of microorganisms like bacteria.

7

IMPLICATIES VOOR DE SAMENLEVING (NL)

In dit hoofdstuk worden wetenschappelijke en technische implicaties van dit onderzoek voor de samenleving besproken.

HET BEGRIJPEN VAN DE FUNDAMENTEN

Alle organismen, van bacterie tot boomvalk, van kwal tot Koreaanse spar, zijn opgebouwd uit cellen. Sommige simpele organismen, zoals bacteriën, bestaan uit slechts één cel. Sommige organismen bestaan uit heel veel cellen. Mensen bestaan bijvoorbeeld uit ongeveer tien biljoen cellen (een één met dertien nullen) [232, BNID 102390]. Een cel heeft een membraan dat een scheidingsbarriere vormt tussen dat wat zich in de cel bevindt en dat wat zich daarbuiten bevindt. In de cel bevinden zich eiwitten, andere biomoleculen zoals signaaleiwitten en bouwstenen, en DNA waarop genetische informatie opgeslagen is. Interacties tussen eiwitten en andere cellulaire moleculen, in combinatie met sensorische inputs vanuit wat buiten de cel gebeurt, bepalen hoe de cel zich gedraagt, welke moleculen er geproduceerd worden in de cel, en of de cel gaat groeien. Zoals in een elektronisch circuit kunnen de eigenschappen van componenten en hun interacties geïdentificeerd en gekarakteriseerd worden, zodat de werking van de cel begrepen en gemanipuleerd kan worden. Momenteel begrijpen we niet alle processen die in de biologische cel plaatsvinden. De aanpak van de kwantitatieve biologie, zoals in deze thesis, is erop gericht cellulaire processen verder in kaart te brengen. Hierbij wordt de bacterie *Escherichia coli* als modelorganisme gebruikt. Het doel is om een beter begrip te krijgen van biologische processen in de cel, en die processen ook te kunnen manipuleren. Aangezien alle organismen opgebouwd zijn uit cellen, inclusief mensen en ziekteverwekkende micro-organismen, kan dit uiteindelijk bijdragen aan het begrijpen en genezen van menselijke ziektes. Daarnaast spelen micro-organismen een belangrijke rol in de industrie, bijvoorbeeld in voedselproductie en rioolwaterzuivering. Wat precies de directe implicaties zullen zijn van een beter begrip van fundamentele processen die plaatsvinden in de cel is moeilijk te voorspellen, maar in potentie kan zulke kennis een grote impact hebben.

BACTERIËN DODEN

Desalniettemin zullen we meer specifiek proberen te reflecteren op de impact van het werk in deze thesis. In hoofdstuk 3 beschrijven we hoe bacteriën verder kunnen delen nadat het delingsproces gepauzeerd is door situaties die de bacterie als stressvol ervaart. Er wordt gedacht dat dit proces een overlevingsmechanisme is. Kennis over hoe overlevingsmechanismen werken zijn nuttig in situaties waarin het vanuit een menselijk

oogpunt gewenst is bacteriën te doden of hun groei te remmen. Bacteriën filamente-
ren (stoppen met delen terwijl ze blijven groeien) in reactie op vele vijandige condities.
Praktisch relevante condities waarin dit gebeurt zijn bijvoorbeeld blootstelling aan an-
tibiotica en hoge en lage temperaturen. Deze condities zijn respectievelijk relevant in
klinische context en in de context van conservering van voeding. Kennis over het fila-
mentatieproces kan bijvoorbeeld voorspellingen met betrekking tot overlevingskansen
van bacteriën verbeteren, en zou uiteindelijk kunnen bijdragen aan het aanpassen van
klinische behandelingen of conservering van voeding effectiever kunnen maken. Daar-
naast kunnen in deze context componenten van het filamentatiemechanisme wellicht
specifiek worden uitgeschakeld om bacteriegroei te remmen.

METHODEN MET EEN BREDERE RELEVANTIE

Naast de deling van filamenteuze bacteriën zijn er nog twee onderwerpen in deze thesis.
Te weten: (1) hoe stochasticiteit, regulatie en individualiteit van bacteriën in verhouding
staan met elkaar, een vraag die we onderzoeken in het CRP metabole regulatiesysteem,
zoals beschreven in hoofdstuk 5, en (2) hoe stochasticiteit in ribosomale expressie zich
verhoudt tot bacteriële individualiteit en groei, zoals beschreven in hoofdstuk 6. Ten eer-
ste is het noemenswaardig dat de methoden die in deze twee studies gebruikt zijn (zie
hoofdstuk 2) van nut kunnen zijn voor de samenleving. Geautomatiseerde analyse van
afbeeldingen met grote hoeveelheden individuele cellen heeft bredere toepassingen, en
opgedane ervaringen met dit type analyse kan nuttig zijn in een klinische setting. Dit zou
bijvoorbeeld nuttig kunnen zijn bij de analyse van bloedcellen of andere weefselanalyse
voor diagnostische toepassingen. Daarnaast kan ook de techniek waarmee individuele
cellen in microkamers gekweekt worden (*single cell microfluidics*) breder toegepast wor-
den. Momenteel zijn er al onderzoeken waarbij data wordt verkregen van individuele
humane cellijnen, wat kan bijdragen aan een beter begrip van humane biologie en ziek-
tes. Vorderingen in *single cell microfluidics* methoden kunnen wellicht bijdragen aan de
verdere ontwikkeling van dit onderzoek.

HET BELANG VAN HETEROGENITEIT

We kunnen slechts speculeren over meer directe, toekomstige impact op de samenle-
ving van hoofdstukken 5 en 6. In hoofdstuk 5 zagen we dat regulatiesystemen wellicht
een rol spelen in het doorgeven van fluctuaties, wat bijdraagt aan de heterogeniteit van
populaties. In hoofdstuk 6 hebben we onderzocht hoe ribosomen bijdragen aan cel-
lulaire individualiteit, deze individualiteit draagt ook bij aan de heterogeniteit van po-
pulaties. Het begrijpen van heterogeniteit in populaties verschaft ons fundamentele
inzichten in hoe de cel werkt, maar heterogeniteit heeft ook een directere relevantie.
Micro-organismen worden ook gebruikt in de industrie. Bijvoorbeeld bij de productie
van yoghurt en bier. Wellicht minder bekend is het idee om micro-organismen gene-
tisch te modificeren om zo brandstof te produceren [233, 234]. Bij zulke toepassingen
worden de micro-organismen doorgaans gebruikt om één specifieke taak uit te voeren.
Bijvoorbeeld de vergisting van een specifiek bier, of de productie van een molecuul dat
als brandstof gebruikt kan worden. Waarschijnlijk is er een optimale staat (configuratie)

waarin een individueel micro-organisme zo'n taak kan uitvoeren, maar heterogeniteit in de populatie zorgt ervoor dat niet de gehele populatie diezelfde staat kan aannemen. Dus hoewel heterogeniteit wellicht voordelig is uit evolutionair oogpunt, zou genetische modificatie van deze eigenschap wellicht de opbrengst en/of regulatie van industriële processen waar micro-organismen bij betrokken zijn kunnen verbeteren.

Voorts, zoals ook in hoofdstuk 4 beschreven, kan heterogeniteit bijdragen aan de overlevingskans van populaties, omdat verschillende individuen voorbereid kunnen zijn op verschillende toekomstige scenario's (*bet hedging*). Kennis over de oorsprong van die heterogeniteit kan dus relevant zijn in een klinische context, waar het vaak juist gewenst is de groei van micro-organismen zoals bacteriën af te remmen.

7

BIBLIOGRAPHY

[1] A. J. Lotka, *Analytical Note on Certain Rhythmic Relations in Organic Systems,* Proceedings of the National Academy of Sciences **6**, 410 (1920).

[2] V. Volterra, *Variations and Fluctuations of the Number of Individuals in Animal Species living together,* ICES Journal of Marine Science **3**, 3 (1928).

[3] A. M. Turing, *The Chemical Basis of Morphogenesis,* Philosophical Transactions of the Royal Society B: Biological Sciences **237**, 37 (1952).

[4] I. Lazebnik, *Can a biologist fix a radio, or what I learned while studying apoptosis.* Uspekhi Gerontologii **12**, 166 (2003).

[5] H. Kitano, *Systems Biology: A Brief Overview,* Science (2002), 10.1126/science.1069492.

[6] R. Wollman, *Robustness, accuracy, and cell state heterogeneity in biological systems,* Current Opinion in Systems Biology (2018), 10.1016/j.coisb.2017.11.009.

[7] C. D. Kelly and O. Rahn, *The Growth Rate of Individual Bacterial Cells.* Journal of Bacteriology **23**, 147 (1932).

[8] J. L. Spudich and D. E. K. Jr., *Non-genetic individuality: chance in the single cell,* Nature **262** (1976).

[9] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, *Stochastic gene expression in a single cell.* Science (New York, N.Y.) **297**, 1183 (2002).

[10] D. J. Kiviet, P. Nghe, N. Walker, S. Boulineau, V. Sunderlikova, and S. J. Tans, *Stochasticity of metabolism and growth at the single-cell level,* Nature (2014), 10.1038/nature13582.

[11] C. J. Davidson and M. G. Surette, *Individuality in bacteria.* Annual review of genetics **42**, 253 (2008).

[12] A. Raj and A. van Oudenaarden, *Nature, nurture, or chance: stochastic gene expression and its consequences.* Cell **135**, 216 (2008).

[13] G. Chalancon, C. N. J. Ravarani, S. Balaji, A. Martinez-Arias, L. Aravind, R. Jothi, and M. M. Babu, *Interplay between gene expression noise and regulatory network architecture,* Trends in Genetics **28**, 221 (2012).

[14] F. J. Bruggeman and B. Teusink, *Living with noise: on the propagation of noise from molecules to phenotype and fitness,* Current Opinion in Systems Biology , 1 (2018).

[15] M. J. Dunlop, R. S. Cox, J. H. Levine, R. M. Murray, and M. B. Elowitz, *Regulatory activity revealed by dynamic correlations in gene expression noise.* Nature genetics **40**, 1493 (2008).

[16] J. W. Young, J. C. W. Locke, A. Altinok, N. Rosenfeld, T. Bacarian, P. S. Swain, E. Mjolsness, and M. B. Elowitz, *Measuring single-cell gene expression dynamics in bacteria using fluorescence time-lapse microscopy.* Nature protocols **7**, 80 (2012).

[17] D. J. Kiviet, *The lac Operon: fluctuations, growth and evolution,* Ph.D. thesis, Delft University of Technology (2010).

[18] N. E. Walker, *A single-cell study on stochasticity in growth and gene expression,* Ph.D. thesis, Delft University of Technology (2016).

[19] S. Gude, *Dispersal and growth of bacteria,* Ph.D. thesis, Delft University of Technology (2016).

[20] G. Dong, Q. Yang, Q. Wang, Y. I. Kim, T. L. Wood, K. W. Osteryoung, A. van Oudenaarden, and S. S. Golden, *Elevated ATPase Activity of KaiC Applies a Circadian Checkpoint on Cell Division in Synechococcus elongatus,* Cell **140**, 529 (2010).

[21] P. Wang, L. Robert, J. Pelletier, W. L. Dang, F. Taddei, A. Wright, and S. Jun, *Robust growth of escherichia coli,* Current Biology **20**, 1099 (2010).

[22] J. C. W. Locke and M. B. Elowitz, *Using movies to analyse gene circuit dynamics in single cells.* Nature reviews. Microbiology **7**, 383 (2009).

[23] P. Hammar, M. Walldén, D. Fange, F. Persson, O. Baltekin, G. Ullman, P. Leroy, and J. Elf, *Direct measurement of transcription factor dissociation excludes a simple operator occupancy model for gene regulation.* Nature genetics **46**, 405 (2014).

[24] A. S. Hansen, N. Hao, and E. K. O. Shea, *High-throughput microfluidics to control and measure signaling dynamics in single yeast cells,* Nature Protocols **10**, 1181 (2015).

[25] M. R. Bennett and J. Hasty, *Microfluidic devices for measuring gene network dynamics in single cells,* Nat. Rev. Genet. **10**, 628 (2010).

[26] K. Nanatani, T. Shijuku, Y. Takano, L. Zulkifli, T. Yamazaki, A. Tominaga, S. Souma, K. Onai, M. Morishita, M. Ishiura, M. Hagemann, I. Suzuki, H. Maruyama, F. Arai, and N. Uozumia, *Comparative Analysis of kdp and ktr Mutants Reveals Distinct Roles of the Potassium Transporters in the Model Cyanobacterium Synechocystis sp. Strain PCC 6803,* Journal of Bacteriology **197**, 676 (2015).

[27] M. Hashimoto, T. Nozoe, H. Nakaoka, R. Okura, S. Akiyoshi, K. Kaneko, E. Kussell, and Y. Wakamoto, *Noise-driven growth rate gain in clonal cellular populations,* Proceedings of the National Academy of Sciences of the United States of America **113**, 3251 (2016).

[28] G. Lambert and E. Kussel, *Memory and Fitness Optimization of Bacteria under Fluctuating Environments,* PLoS Genetics **10** (2014), 10.1371/journal.pgen.1004556.

[29] G. Ullman, M. Wallden, E. G. Marklund, A. Mahmutovic, I. Razinkov, and J. Elf, *High-throughput gene expression analysis at the level of single proteins using a microfluidic turbidostat and automated cell tracking.* Philosophical transactions of the Royal Society of London. Series B, Biological sciences **368**, 20120025 (2013).

[30] J. W. Young, J. C. W. Locke, and M. B. Elowitz, *Rate of environmental change determines stress response specificity.* Proceedings of the National Academy of Sciences of the United States of America **110**, 4140 (2013), arXiv:arXiv:1408.1149 .

[31] M. Ferry, I. Razinkov, and J. Hasty, *Synthetic Biology, Part A,* 1st ed., Vol. 497 (Elsevier Inc., 2011) pp. 295–372.

[32] S. Boulineau, F. Tostevin, D. J. Kiviet, P. R. ten Wolde, P. Nghe, and S. J. Tans, *Single-cell dynamics reveals sustained growth during diauxic shifts.* PloS one **8**, e61686 (2013).

[33] S. Taheri-Araghi, S. Bradde, J. T. Sauls, N. S. Hill, P. A. Levin, J. Paulsson, M. Vergassola, and S. Jun, *Cell-Size Control and Homeostasis in Bacteria,* Current Biology **25**, 1 (2014).

[34] A. Estévez-Torres, A. Yamada, and L. Wang, *An inexpensive and durable epoxy mould for PDMS,* (2009).

[35] K. Haubert, T. Drier, D. Beebe, D. Duffy, J. McDonald, O. Schueler, G. Whitesides, B. Jo, L. V. Lerbeerghe, K. Motsegood, D. Beebe, B. Ginn, O. Steinbock, D. Eddington, J. Puccinelli, and D. Beebe, *PDMS bonding by means of a portable, low-cost corona system,* Lab on a Chip **6**, 1548 (2006).

[36] S. Stylianidou, C. Brennan, S. B. Nissen, N. J. Kuwada, and P. A. Wiggins, *SuperSegger: robust image segmentation, analysis and lineage tracking of bacterial cells,* Molecular Microbiology **102**, 690 (2016).

[37] A. Paintdakhi, B. Parry, M. Campos, I. Irnov, J. Elf, I. Surovtsev, and C. Jacobs-Wagner, *Oufti: An integrated software package for high-accuracy, high-throughput quantitative microscopy analysis,* Molecular Microbiology **99**, 767 (2016), arXiv:arXiv:1011.1669v3 .

[38] M. Kaiser, F. Jug, O. Silander, S. Deshpande, T. Julou, G. Myers, and E. V. Nimwegen, *Tracking single-cell gene regulation in dynamically controlled environments using an integrated microfluidic and computational setup* . bioRxiv , 1 (2016), arXiv:/dx.doi.org/10.1101/076224 [http:] .

[39] C. C. Sachs, A. Grünberger, S. Helfrich, C. Probst, W. Wiechert, D. Kohlheyer, and K. Nöh, *Image-based single cell profiling: High-throughput processing of mother machine experiments,* PLoS ONE **11**, 1 (2016).

[40] J.-B. Nobs and S. J. Maerkl, *Long-term single cell analysis of S. pombe on a microfluidic microchemostat array.* PloS one **9**, e93466 (2014).

[41] O. Sliusarenko, J. Heinritz, T. Emonet, and C. Jacobs-Wagner, *High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics,* Molecular Microbiology **80**, 612 (2011), arXiv:arXiv:1011.1669v3 .

[42] S. K. Sadanandan, O. Baltekin, K. E. G. Magnusson, A. Boucharin, P. Ranefall, J. Jalden, J. Elf, and C. Wahlby, *Segmentation and Track-Analysis in Time-Lapse Imaging of Bacteria,* IEEE Journal of Selected Topics in Signal Processing **10**, 174 (2016).

[43] S. Chowdhury, M. Kandhavelu, O. Yli-Harja, and A. S. Ribeiro, *Cell segmentation by multi-resolution analysis and maximum likelihood estimation (MAMLE).* BMC Bioinformatics **14**, S8 (2013).

[44] D. A. Van Valen, T. Kudo, K. M. Lane, D. N. Macklin, N. T. Quach, M. M. DeFelice, I. Maayan, Y. Tanouchi, E. A. Ashley, and M. W. Covert, *Deep Learning Automates the Quantitative Analysis of Individual Cells in Live-Cell Imaging Experiments,* PLoS Computational Biology **12**, 1 (2016).

[45] E. W. Weisstein, *Mathworld,* .

[46] B. Munsky, G. Neuert, and A. van Oudenaarden, *Using gene expression noise to understand gene regulation.* Science (New York, N.Y.) **336**, 183 (2012).

[47] C. Shannon, *A Mathematical Theory of Communication,* Bell Syst Tech **27** (1948), 10.1016/S0006-3495(96)79210-X, arXiv:9411012 [chao-dyn] .

[48] C. M. Bishop, *Pattern recognition and machine learning* (Springer Science, New York, 2006).

[49] *Mutual Information* (Wikipedia, The Free Encyclopedia).

[50] Z. I. Botev, J. F. Grotowski, and D. P. Kroese, *Kernel density estimation via diffusion,* Annals of Statistics **38**, 2916 (2010), arXiv:arXiv:1011.2602v1 .

[51] M. Wehrens, D. Ershov, R. Rozendaal, N. Walker, D. Schultz, R. Kishony, P. A. Levin, and S. J. Tans, *Size Laws and Division Ring Dynamics in Filamentous Escherichia coli cells,* Current Biology , 1 (2018).

[52] W. F. Marshall, K. D. Young, M. Swaffer, E. Wood, P. Nurse, A. Kimura, J. Frankel, J. Wallingford, V. Walbot, X. Qu, and A. H. K. Roeder, *What determines cell size?* BMC biology **10** (2012), 10.1186/1741-7007-10-101.

[53] P. Jorgensen and M. Tyers, *How Cells Coordinate Growth and Division,* Current Biology **14**, R1014 (2004).

[54] L. Robert, *Size sensors in bacteria, cell cycle control, and size control,* Frontiers in Microbiology **6** (2015), 10.3389/fmicb.2015.00515.

[55] A.-C. Chien, N. S. Hill, and P. A. Levin, *Cell Size Control in Bacteria,* Current Biology **22**, R340 (2012).

[56] J. J. Turner, J. C. Ewald, and J. M. Skotheim, *Cell size control in yeast,* Current biology **22**, R350 (2012).

[57] M. Wallden, D. Fange, E. Gregorsson Lundius, Ö. Baltekin, and J. Elf, *The synchronization of replication and division cycles in individual E. coli cells,* Cell **166**, 729 (2016).

[58] M. Campos, I. V. Surovtsev, S. Kato, A. Paintdakhi, B. Beltran, S. E. Ebmeier, and C. Jacobs-Wagner, *A constant size extension drives bacterial cell size homeostasis,* Cell **159**, 1433 (2014), arXiv:NIHMS150003 .

[59] A. Amir, *Cell size regulation in bacteria,* Physical Review Letters **112**, 1 (2014), arXiv:1312.6562 .

[60] M. Osella, E. Nugent, and M. Cosentino Lagomarsino, *Concerted control of Escherichia coli cell division.* Proc. Natl. Acad. Sci. U.S.A. **111**, 3431 (2014).

[61] H. Suzuki, J. Pangborn, and W. W. Kilgore, *Filamentous Cells of Escherichia coli Formed in the Presence of Mitomycin,* Journal of Bacteriology **93**, 683 (1967).

[62] G. N. Rolinson, *Effect of beta-lactam antibiotics on bacterial cell growth rate.* Journal of general microbiology **120**, 317 (1980).

[63] C. Miller, L. E. Thomsen, C. Gaggero, R. Mosseri, H. Ingmer, and S. N. Cohen, *SOS response induction by beta-lactams and bacterial defense against antibiotic lethality.* Science (New York, N.Y.) **305**, 1629 (2004).

[64] P. Domadia, S. Swarup, A. Bhunia, J. Sivaraman, and D. Dasgupta, *Inhibition of bacterial cell division protein FtsZ by cinnamaldehyde,* Biochemical Pharmacology **74**, 831 (2007).

[65] S. S. Justice, D. a. Hunstad, P. C. Seed, and S. J. Hultgren, *Filamentation by Escherichia coli subverts innate defenses during urinary tract infection.* Proc. Natl. Acad. Sci. U.S.A. **103**, 19884 (2006).

[66] J. Möller, T. Luehmann, H. Hall, and V. Vogel, *The race to the pole: How high-aspect ratio shape and heterogeneous environments limit phagocytosis of filamentous Escherichia coli bacteria by macrophages,* Nano Letters **12**, 2901 (2012).

[67] T. Jones, C. O. Gill, and L. M. McMullen, *The behaviour of log phase Escherichia coli at temperatures that fluctuate about the minimum for growth,* Letters in Applied Microbiology **39**, 296 (2004).

[68] S. Linn and J. a. Imlay, *Toxicity, mutagenesis and stress responses induced in Escherichia coli by hydrogen peroxide.* J. Cell Sci. **6**, 289 (1987).

[69] T. Kawarai, M. Wachi, H. Ogino, S. Furukawa, K. Suzuki, H. Ogihara, and M. Ya-masaki, *SulA-independent filamentation of Escherichia coli during growth after release from high hydrostatic pressure treatment,* Appl. Microbiol. Biotechnol. **64**, 255 (2004).

[70] M. Wainwright, L. T. Canham, K. Al-Wajeeh, and C. L. Reeves, *Morphological changes (including filamentation) in Escherichia coli grown under starvation conditions on silicon wafers and other surfaces,* Letters in Applied Microbiology **29**, 224 (1999).

[71] B. Rosenberg, E. Renshaw, L. VanCamp, J. Hartwick, and J. Drobnik, *Platinum-Induced Filamentous Growth in Escherichia coli1,* Journal of Bacteriology **93**, 716 (1967).

[72] H. I. Adler and a. a. Hardigree, *Growth and Division of Filamentous Forms of Escherichia coli.* Journal of bacteriology **90**, 223 (1965).

[73] G. J. Kantor and R. A. Deering, *Ultraviolet radiation studies of filamentous Escherichia coli B.* Journal of Bacteriology **92**, 1062 (1966).

[74] M. Radman, *SOS repair hypothesis: Phenomenology of an inducible DNA repair which is accompanied by mutagenesis,* in *Molecular Mechanisms for Repair of DNA, Part A,* edited by P. Hanawalt (Springer US, 1975) pp. 355–367.

[75] B. Michel, *After 30 years of study, the bacterial SOS response still surprises us,* PLoS Biol. **3**, 1174 (2005).

[76] T. H. Jones, K. M. Vail, and L. M. McMullen, *Filament formation by foodborne bacteria under sublethal stress,* Int. J. Food. Microbiol. **165**, 97 (2013).

[77] S. S. Justice, D. a. Hunstad, L. Cegelski, and S. J. Hultgren, *Morphological plasticity as a bacterial survival strategy.* Nat. Rev. Microbiol. **6**, 162 (2008).

[78] R. Pulvertaft, *The effect of antibiotics on growing cultures in Bacterium coli,* J Pathol Bacteriol **64**, 75 (1952).

[79] A. Dajkovic, A. Mukherjee, and J. Lutkenhaus, *Investigation of regulation of FtsZ assembly by SulA and development of a model for FtsZ polymerization,* Journal of Bacteriology **190**, 2513 (2008).

[80] W. Donachie and K. Begg, *Growth of the Bacterial Cell,* Nature **227** (1970).

[81] P. E. Taschner, P. G. Huls, E. Pas, and C. L. Woldringh, *Division behavior and shape changes in isogenic ftsZ, ftsQ, ftsA, pbpB, and ftsE cell division mutants of Escherichia coli during temperature shift experiments.* Journal of Bacteriology **170**, 1533 (1988).

[82] S. G. Addinall, C. Cao, and J. Lutkenhaus, *Temperature shift experiments with an ftsZ84(Ts) strain reveal rapid dynamics of FtsZ localization and indicate that the Z ring is required throughout septation and cannot reoccupy division sites once constriction has initiated,* Journal of Bacteriology **179**, 4277 (1997).

[83] H. A. Arjes, A. Kriel, N. A. Sorto, J. T. Shaw, J. D. Wang, and P. A. Levin, *Failsafe mechanisms couple division and DNA replication in bacteria,* Current Biology **24**, 2149 (2014).

[84] E. Mileykovskaya, Q. Sun, W. Margolin, and W. Dowhan, *Localization and function of early cell division proteins in filamentous Escherichia coli cells lacking phosphatidylethanolamine.* Journal of Bacteriology **180**, 4252 (1998).

[85] A. Typas, M. Banzhaf, C. A. Gross, and W. Vollmer, *From the regulation of peptidoglycan synthesis to bacterial growth and morphology,* Nat. Rev. Microbiol. **10**, 123 (2012).

[86] W. D. Donachie and K. J. Begg, *'Division potential' in Escherichia coli,* Journal of Bacteriology **178**, 5971 (1996).

[87] P. Y. Ho and A. Amir, *Simultaneous regulation of cell size and chromosome replication in bacteria,* Frontiers in Microbiology **6**, 1 (2015), arXiv:arXiv:1507.07032v1 .

[88] H. Zheng, P.-Y. Ho, M. Jiang, B. Tang, W. Liu, D. Li, X. Yu, N. E. Kleckner, A. Amir, and C. Liu, *Interrogating the Escherichia coli cell cycle by cell dimension perturbations,* Proceedings of the National Academy of Sciences , 201617932 (2016).

[89] F. Si, D. Li, S. E. Cox, J. T. Sauls, O. Azizi, C. Sou, A. B. Schwartz, M. J. Erickstad, Y. Jun, X. Li, and S. Jun, *Invariance of initiation mass and predictability of cell size in Escherichia coli,* Current Biology **27** (2017), 10.1016/j.cub.2017.03.022.

[90] S. Cooper and C. E. Helmstetter, *Chromosome Replication and the Division Cycle of Escherichia coli B/r,* J. Mol. Biol. **31** (1968), 10.1016/0147-619X(81)90042-1.

[91] H. P. Erickson, D. E. Anderson, and M. Osawa, *FtsZ in bacterial cytokinesis: cytoskeleton and force generator all in one.* Microbiol. Mol. Biol. Rev. **74**, 504 (2010).

[92] J. Lutkenhaus, S. Pichoff, and S. Du, *Bacterial cytokinesis: From Z ring to divisome,* Cytoskeleton **69**, 778 (2012).

[93] R. Tsukanov, G. Reshes, G. Carmon, E. Fischer-Friedrich, N. S. Gov, I. Fishov, and M. Feingold, *Timing of Z-ring localization in Escherichia coli,* Physical Biology **8** (2011), 10.1088/1478-3975/8/6/066003.

[94] M. Loose, K. Kruse, and P. Schwille, *Protein self-organization: lessons from the min system.* Annual review of biophysics **40**, 315 (2011).

[95] D. M. Raskin and P. A. de Boer, *Rapid pole-to-pole oscillation of a protein required for directing division to the middle of Escherichia coli.* Proceedings of the National Academy of Sciences of the United States of America **96**, 4971 (1999).

[96] H. Meinhardt and P. A. J. de Boer, *Pattern formation in Escherichia coli: A model for the pole-to-pole oscillations of Min proteins and the localization of the division site,* Proceedings of the National Academy of Sciences of the United States of America **98**, 14202 (2001).

[97] K. C. Huang, Y. Meir, and N. S. Wingreen, *Dynamic structures in Escherichia coli: Spontaneous formation of MinE rings and MinD polar zones,* Proceedings of the National Academy of Sciences **100**, 12724 (2003).

[98] J. Hattne, D. Fange, and J. Elf, *Stochastic reaction-diffusion simulation with MesoRD,* Bioinformatics **21**, 2923 (2005).

[99] D. Fange and J. Elf, *Noise-induced Min phenotypes in E. coli,* PLoS computational biology **2**, 637 (2006).

[100] H. I. Adler, W. D. Fisher, A. Cohen, and A. A. Hardigree, *Miniature Escherichia coli cells deficient in DNA,* Proc. Natl. Acad. Sci. **57**, 321 (1967), arXiv:9810036 [physics] .

[101] L. J. Wu and J. Errington, *Nucleoid occlusion and bacterial cell division,* Nature reviews Microbiology **10**, 8 (2011).

[102] E. Mulder and C. L. Woldringh, *Actively replicating nucleoids influence positioning of division sites in Escherichia coli filaments forming cells lacking DNA.* Journal of Bacteriology **171**, 4303 (1989).

[103] E. Bi and J. Lutkenhaus, *FtsZ regulates frequency of cell division in Escherichia coli,* Journal of Bacteriology **172**, 2765 (1990).

[104] P. A. de Boer, R. E. Crossley, and L. I. Rothfield, *A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in E. coli,* Cell **56**, 641 (1989).

[105] M. Wery, C. L. Woldringh, and J. Rouviere-Yaniv, *HU-GFP and DAPI co-localize on the Escherichia coli nucleoid,* Biochimie **83**, 193 (2001).

[106] M. B. Elowitz, M. G. Surette, P. E. Wolf, J. B. Stock, and S. Leibler, *Protein mobility in the cytoplasm of Escherichia coli,* Journal of Bacteriology **181**, 197 (1999).

[107] M. Wehrens, F. Buke, P. Nghe, and S. J. Tans, *Stochasticity in cellular metabolism and growth: Approaches and consequences,* Current Opinion in Systems Biology **8**, 131 (2018).

[108] A. J. Walsh and M. C. Skala, *Optical metabolic imaging quantifies heterogeneous cell populations.* Biomedical optics express **6**, 559 (2015).

[109] J. L. Radzikowski, S. Vedelaar, D. Siegel, Á. D. Ortega, A. Schmidt, and M. Heinemann, *Bacterial persistence is an active σS stress response to metabolic flux limitation.* Molecular systems biology **12**, 882 (2016).

[110] E. Kussell, *Information in Fluctuating Environments Phenotypic Diversity , Population Growth , and Information in Fluctuating Environments,* Methods **309**, 2075 (2005).

[111] M. Acar, J. T. Mettetal, and A. van Oudenaarden, *Stochastic switching as a survival strategy in fluctuating environments,* Nature Genetics **40**, 471 (2008).

[112] Z. Wang and J. Zhang, *Impact of gene expression noise on organismal fitness and the efficacy of natural selection.* Proceedings of the National Academy of Sciences of the United States of America **108**, E67 (2011).

[113] M. J. Herrgård, M. W. Covert, and B. Ø. Palsson, *Reconstruction of microbial transcriptional regulatory networks,* Current Opinion in Biotechnology **15**, 70 (2004).

[114] F. C. Neidhardt, J. L. Ingraham, and M. Schaechter, *Physiology of the Bacterial Cell: A Molecular Approach* (Sinauer Associates Inc, Sunderland, 1990).

[115] M. Rodríguez, T. A. Good, M. E. Wales, J. P. Hua, and J. R. Wild, *Modeling allosteric regulation of de novo pyrimidine biosynthesis in Escherichia coli,* Journal of Theoretical Biology **234**, 299 (2005).

[116] Y. Hart, D. Madar, J. Yuan, A. Bren, A. E. Mayo, J. D. Rabinowitz, and U. Alon, *Robust control of nitrogen assimilation by a bifunctional enzyme in E. coli.* Molecular cell **41**, 117 (2011).

[117] S. Klumpp, Z. Zhang, and T. Hwa, *Growth rate-dependent global effects on gene expression in bacteria.* Cell **139**, 1366 (2009).

[118] H. S. Yun, J. Hong, and H. C. Lim, *Regulation of ribosome synthesis in Escherichia coli: Effects of temperature and dilution rate changes,* Biotechnology and Bioengineering **52**, 615 (1996).

[119] E. M. T. El-Mansi and W. H. Holms, *Control of Carbon Flux to Acetate Excretion During Growth of Escherichia coli in Batch and Continuous Cultures,* Journal of General Microbiology **135** (1989).

[120] W. A. Wilson, P. J. Roach, M. Montero, E. Baroja-Fernández, F. J. Muñoz, G. Eydallin, A. M. Viale, and J. Pozueta-Romero, *Regulation of glycogen metabolism in yeast and bacteria,* FEMS Microbiology Reviews **34**, 952 (2010), arXiv:NIHMS150003 .

[121] T. Esaki and T. Masujima, *Fluorescence Probing Live Single-cell Mass Spectrometry for Direct Analysis of Organelle Metabolism.* Anal Sci **31**, 1211 (2015).

[122] S. Yue and J.-X. Cheng, *Deciphering single cell metabolism by coherent Raman scattering microscopy,* Current Opinion in Chemical Biology **33**, 46 (2016).

[123] J. Nakai, M. Ohkura, and K. Imoto, *A high signal-to-noise Ca(2+) probe composed of a single green fluorescent protein.* Nature biotechnology **19**, 137 (2001).

[124] T. Nagai, A. Sawano, E. S. Park, and A. Miyawaki, *Circularly permuted green fluorescent proteins engineered to sense Ca2+,* Proceedings of the National Academy of Sciences **98**, 3197 (2001).

[125] H. Yaginuma, S. Kawai, K. V. Tabata, K. Tomiyama, A. Kakizuka, T. Komatsuzaki, H. Noji, and H. Imamura, *Diversity in ATP concentrations in a single bacterial cell population revealed by quantitative single-cell imaging,* Sci Rep **4**, 6522 (2014).

[126] H. Imamura, K. P. Huynh Nhat, H. Togawa, K. Saito, R. Iino, Y. Kato-Yamada, T. Na-gai, and H. Noji, *Visualization of ATP levels inside single living cells with fluorescence resonance energy transfer-based genetically encoded indicators,* Proceedings of the National Academy of Sciences **106**, 15651 (2009), arXiv:arXiv:1408.1149 .

[127] J. Klarenbeek, J. Goedhart, A. van Batenburg, D. Groenewald, and K. Jalink, *Fourth-Generation Epac-Based FRET Sensors for cAMP Feature Exceptional Brightness, Photostability and Dynamic Range: Characterization of Dedicated Sensors for FLIM, for Ratiometry and with High Affinity,* Plos One **10**, e0122513 (2015).

[128] B. Munsky and M. Khammash, *Transient analysis of stochastic switches and trajectories with applications to gene regulatory networks,* IET systems biology (2008), 10.1049/iet-syb, arXiv:0604006 [q-bio] .

[129] J. M. Pedraza and A. van Oudenaarden, *Noise propagation in gene networks.* Science (New York, N.Y.) **307**, 1965 (2005).

[130] V. Chubukov, L. Gerosa, K. Kochanowski, and U. Sauer, *Coordination of microbial metabolism,* Nature Reviews Microbiology (2014), 10.1038/nrmicro3238.

[131] J. H. Van Heerden, H. Kempe, A. Doerr, T. Maarleveld, N. Nordholt, and F. J. Bruggeman, *Statistics and simulation of growth of single bacterial cells: Illustrations with B. subtilis and E. coli,* Scientific Reports **7**, 1 (2017).

[132] I. Arganda-Carreras, V. Kaynig, C. Rueden, K. W. Eliceiri, J. Schindelin, A. Cardona, and H. S. Seung, *Trainable Weka Segmentation: A machine learning tool for microscopy pixel classification,* Bioinformatics **33**, 2424 (2017).

[133] M. Mir, Z. Wang, Z. Shen, M. Bednarz, R. Bashir, I. Golding, S. G. Prasanth, and G. Popescu, *Optical measurement of cycle-dependent cell growth,* Proceedings of the National Academy of Sciences **108**, 13124 (2011).

[134] S. Son, A. Tzur, Y. Weng, P. Jorgensen, J. Kim, M. W. Kirschner, and S. R. Manalis, *Direct observation of mammalian cell growth and size regulation,* Nature Methods **9**, 910 (2012).

[135] S. Tsuru, J. Ichinose, A. Kashiwagi, B.-W. Ying, K. Kaneko, and T. Yomo, *Noisy cell growth rate leads to fluctuating protein concentration in bacteria.* Physical biology **6**, 036015 (2009).

[136] G. Balázsi, A. Van Oudenaarden, and J. J. Collins, *Cellular decision making and biological noise: From microbes to mammals,* Cell **144**, 910 (2011).

[137] H. Maamar, A. Raj, and D. Dubnau, *Noise in gene expression determines cell fate in Bacillus subtilis.* Science **317**, 526 (2005), arXiv:NIHMS150003 .

[138] N. Walker, P. Nghe, and S. J. Tans, *Generation and filtering of gene expression noise by the bacterial cell cycle.* BMC biology **14**, 11 (2016).

[139] M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, and T. Hwa, *Interdependence of Cell Growth and Gene Expression: Origins and Consequences,* Science **330** (2010).

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

[141] J. Stewart-Ornstein, J. S. Weissman, and H. El-Samad, *Cellular Noise Regulons Underlie Fluctuations in Saccharomyces cerevisiae,* Molecular Cell **45**, 483 (2012), arXiv:NIHMS150003 .

[142] J. H. van Heerden, M. T. Wortel, F. J. Bruggeman, J. J. Heijnen, Y. J. M. Bollen, R. Planqué, J. Hulshof, T. G. O'Toole, S. A. Wahl, and B. Teusink, *Lost in Transition: Startup of Glycolysis Yields Subpopulations of Nongrowing Cells.* Science (New York, N.Y.) (2014), 10.1126/science.1245114.

[143] P. Murima, M. Zimmermann, T. Chopra, F. Pojer, G. Fonti, M. Dal Peraro, S. Alonso, U. Sauer, K. Pethe, and J. D. McKinney, *A rheostat mechanism governs the bifurcation of carbon flux in mycobacteria,* Nature Communications **7**, 12527 (2016).

[144] A.-K. Gustavsson, C. B. Adiels, B. Mehlig, and M. Goksör, *Entrainment of heterogeneous glycolytic oscillations in single cells.* Scientific reports **5**, 9404 (2015), arXiv:1504.0131 .

[145] A. Papagiannakis, B. Niebel, E. C. Wit, and M. Heinemann, *Autonomous Metabolic Oscillations Robustly Gate the Early and Late Cell Cycle,* Molecular Cell **65**, 285 (2017).

[146] D. Van Dijk, R. Dhar, A. Missarova, L. Espinar, W. Blevins, B. Lehner, and L. Carey, *Slow growing subpopulations have increased DNA damage and RNA polymerase error rates,* Nature Communications , 1 (2015).

[147] A. Adiciptaningrum, M. Osella, M. C. Moolman, M. Cosentino Lagomarsino, and S. J. Tans, *Stochasticity and homeostasis in the E. coli replication and division cycle.* Scientific reports **5**, 18261 (2015).

[148] S. Iyer-Biswas, C. S. Wright, J. T. Henry, K. Lo, S. Burov, Y. Lin, G. E. Crooks, S. Crosson, a. R. Dinner, and N. F. Scherer, *Scaling laws governing stochastic growth and division of single bacterial cells,* Proceedings of the National Academy of Sciences (2014), 10.1073/pnas.1403232111.

[149] A. S. Kennard, M. Osella, A. Javer, J. Grilli, P. Nghe, S. J. Tans, P. Cicuta, and M. Cosentino Lagomarsino, *Individuality and universality in the growth-division laws of single E. Coli cells,* Physical Review E **93**, 1 (2016), arXiv:1411.4321 .

[150] M. Osella, S. J. Tans, and M. Cosentino Lagomarsino, *Step by Step, Cell by Cell: Quantification of the Bacterial Cell Cycle,* Trends in Microbiology **25**, 250 (2017).

[151] S. Tănase-Nicola and P. R. ten Wolde, *Regulatory control and the costs and benefits of biochemical noise.* PLoS computational biology **4**, e1000125 (2008).

[152] B. Cerulus, A. M. New, K. Pougach, and K. J. Verstrepen, *Noise and Epigenetic Inheritance of Single-Cell Division Times Influence Population Fitness,* Current Biology **26**, 1138 (2016).

[153] B. D. Towbin, Y. Korem, A. Bren, S. Doron, R. Sorek, and U. Alon, *Optimality and sub-optimality in a bacterial growth law,* Nature Communications **8**, 14123 (2017).

[154] A. Kashiwagi, I. Urabe, K. Kaneko, and T. Yomo, *Adaptive response of a gene network to environmental changes by fitness-induced attractor selection,* PLoS ONE **1** (2006), 10.1371/journal.pone.0000049.

[155] A. M. New, B. Cerulus, S. K. Govers, G. Perez-Samper, B. Zhu, S. Boogmans, J. B. Xavier, and K. J. Verstrepen, *Different Levels of Catabolite Repression Optimize Growth in Stable and Variable Environments,* PLoS Biology **12**, 17 (2014).

[156] J. Miot, L. Remusat, E. Duprat, A. Gonzalez, S. Pont, and M. Poinsot, *Fe biomineralization mirrors individual metabolic activity in a nitrate-dependent Fe(II)-oxidizer,* Frontiers in Microbiology **6**, 1 (2015).

[157] D. D. Martino, F. Capuani, and A. D. Martino, *Growth against entropy in bacterial metabolism: the phenotypic trade-off behind empirical growth rate distributions in E. coli,* Physical Biology **13**, 036005 (2016), arXiv:1601.03243 .

[158] F. Schreiber, S. Littmann, G. Lavik, S. Escrig, A. Meibom, M. Kuypers, and M. Ackermann, *Phenotypic heterogeneity driven by nutrient limitation promotes growth in fluctuating environments,* Nature Microbiology **in press**, 1 (2016).

[159] B. P. H. Metzger, D. C. Yuan, J. D. Gruber, F. D. Duveau, and P. J. Wittkopp, *Selection on noise constrains variation in a eukaryotic promoter,* Nature **521**, 344 (2015), arXiv:NIHMS150003 .

[160] R. A. T. Mars, P. Nicolas, M. Ciccolini, E. Reilman, A. Reder, M. Schaffer, U. Mäder, U. Völker, J. M. van Dijl, and E. L. Denham, *Small Regulatory RNA-Induced Growth Rate Heterogeneity of Bacillus subtilis,* PLoS Genetics **11**, 1 (2015).

[161] H. J. E. Beaumont, J. Gallie, C. Kost, G. C. Ferguson, and P. B. Rainey, *Experimental evolution of bet hedging,* Nature **462**, 90 (2009).

[162] T. Morisaki, K. Lyon, K. F. DeLuca, J. G. DeLuca, B. P. English, Z. Zhang, L. D. Lavis, J. B. Grimm, S. Viswanathan, L. L. Looger, T. Lionnet, and T. J. Stasevich, *Real-time quantification of single RNA translation dynamics in living cells,* Science **352**, 1425 (2016).

[163] C. Cadart, S. Monnier, J. Grilli, R. Attia, E. Terriac, B. Baum, M. Cosentino-Lagomarsino, and M. Piel, *Size control in mammalian cells involves modulation of both growth rate and cell cycle duration,* bioRxiv **152728** (2017).

[164] B. Okumus, D. Landgraf, G. C. Lai, S. Bakhsi, J. C. Arias-Castro, S. Yildiz, D. Huh, R. Fernandez-Lopez, C. N. Peterson, E. Toprak, M. El Karoui, and J. Paulsson, *Mechanical slowing-down of cytoplasmic diffusion allows in vivo counting of proteins in individual cells,* Nature Communications **7**, 11641 (2016).

[165] I. Georgakoudi and K. P. Quinn, *Optical Imaging Using Endogenous Contrast to Assess Metabolic State,* Annual Review of Biomedical Engineering **14**, 351 (2012).

[166] D. Bray, *Protein molecules as computational elements in living cells,* (1995).

[167] U. Alon, *An introduction to systems biology, design principles of biological circuits* (Chapman & Hall/CRC, Boca Raton, 2006).

[168] U. Alon, *Network motifs: theory and experimental approaches.* Nature reviews. Genetics **8**, 450 (2007).

[169] J. J. Tyson and B. Novák, *Functional Motifs in Biochemical Reaction Networks,* Annual Review of Physical Chemistry **61**, 219 (2010), arXiv:NIHMS150003 .

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

[171] I. M. Keseler, A. Mackie, A. Santos-Zavaleta, R. Billington, C. Bonavides-Martínez, R. Caspi, C. Fulcher, S. Gama-Castro, A. Kothari, M. Krummenacker, M. Latendresse, L. Muñiz-Rascado, Q. Ong, S. Paley, M. Peralta-Gil, P. Subhraveti, D. A. Velázquez-Ramírez, D. Weaver, J. Collado-Vides, I. Paulsen, and P. D. Karp, *The EcoCyc database: reflecting new knowledge about Escherichia coli K-12.* Nucleic acids research **45** (2017), 10.1093/nar/gkw1003.

[172] E. Fic, P. Bonarek, A. Gorecki, S. Kedracka-Krok, J. Mikolajczak, A. Polit, M. Tworzydlo, M. Dziedzicka-Wasylewska, and Z. Wasylewski, *cAMP receptor protein from escherichia coli as a model of signal transduction in proteins–a review.* Journal of molecular microbiology and biotechnology **17**, 1 (2009).

[173] J. Deutscher, *The mechanisms of carbon catabolite repression in bacteria,* Current Opinion in Microbiology **11**, 87 (2008).

[174] B. Görke and J. Stülke, *Carbon catabolite repression in bacteria: many ways to make the most out of nutrients.* Nature reviews. Microbiology **6**, 613 (2008).

[175] K. Kochanowski, L. Gerosa, S. F. Brunner, D. Christodoulou, Y. V. Nikolaev, and U. Sauer, *Few regulatory metabolites coordinate expression of central metabolic genes in Escherichia coli,* Molecular Systems Biology **13**, 903 (2017).

[176] J. Green, M. R. Stapleton, L. J. Smith, P. J. Artymiuk, C. Kahramanoglou, D. M. Hunt, and R. S. Buxton, *Cyclic-AMP and bacterial cyclic-AMP receptor proteins revisited: adaptation for different ecological niches.* Current opinion in microbiology **18**, 1 (2014).

[177] T. Shimada, N. Fujita, K. Yamamoto, and A. Ishihama, *Novel roles of cAMP receptor protein (CRP) in regulation of transport and metabolism of carbon sources.* PloS one **6**, e20081 (2011).

[178] D. C. Grainger, D. Hurd, M. Harrison, J. Holdstock, and S. J. W. Busby, *Studies of the distribution of Escherichia coli cAMP-receptor protein and RNA polymerase along the E. coli chromosome,* Proceedings of the National Academy of Sciences **102**, 17693 (2005).

[179] D. Zheng, C. Constantinidou, J. L. Hobman, and S. D. Minchin, *Identification of the CRP regulon using in vitro and in vivo transcriptional profiling,* Nucleic Acids Research **32**, 5874 (2004).

[180] D. L. Nelson and M. M. Cox, *Lehninger principles of biochemistry*, 4th ed. (W.H. Freeman and Company, New York, USA, 2005).

[181] K. Hantke, K. Winkler, and J. E. Schultz, *Escherichia coli exports cyclic AMP via TolC,* Journal of Bacteriology **193**, 1086 (2011).

[182] P. R. Jensen, O. Michelsen, and H. V. Westerhoff, *Control analysis of the dependence of Escherichia coli physiology on the H(+)-ATPase,* Proceedings of the National Academy of Sciences of the United States of America **90**, 8068 (1993).

[183] E. Dekel and U. Alon, *Optimality and evolutionary tuning of the expression level of a protein.* Nature **436**, 588 (2005).

[184] J. Berkhout, E. Bosdriesz, E. Nikerel, D. Molenaar, D. de Ridder, B. Teusink, and F. J. Bruggeman, *How biochemical constraints of cellular growth shape evolutionary adaptations in metabolism,* Genetics **194**, 505 (2013).

[185] J. C. J. Ray, M. L. Wickersheim, A. P. Jalihal, Y. O. Adeshina, T. F. Cooper, and G. Balázsi, *Cellular Growth Arrest and Persistence from Enzyme Saturation,* PLoS Computational Biology **12**, 1 (2016).

[186] H. Aiba, T. Nakamura, H. Mitani, and H. Mori, *Mutations that alter the allosteric nature of cAMP receptor protein of Escherichia coli.* The EMBO journal **4**, 3329 (1985).

[187] S. Garges and S. Adhya, *Sites of allosteric shift in the structure of the cyclic AMP receptor protein.* Cell **41**, 745 (1985).

[188] B. D. Towbin, *Personal Communications,* (2017).

[189] L. Landis, J. Xu, and R. C. Johnson, *The cAMP receptor protein CRP can function as an osmoregulator of transcription in Escherichia coli,* Genes and Development (1999).

[190] J. Johansson, C. Balsalobre, S. Y. Wang, J. Urbonaviciene, D. J. Jin, B. Sondén, and B. E. Uhlin, *Nucleoid proteins stimulate stringently controlled bacterial promoters: a link between the cAMP-CRP and the (p)ppGpp regulons in Escherichia coli.* Cell **102**, 475 (2000).

[191] D. W. Jackson, J. W. Simecka, and T. Romeo, *Catabolite Repression of Escherichia coli Biofilm Formation Catabolite Repression of Escherichia coli Biofilm Formation,* Journal of Bacteriology **184**, 3406 (2002).

[192] M. Baga, M. Goransson, S. Normark, and B. E. Uhlin, *Transcriptional activation of a Pap pilus virulence operon from uropathogenic Eschenichia coli,* The EMBO Journal **4**, 3887 (1985).

[193] X. J. Mao, Y. X. Huo, M. Buck, A. Kolb, and Y. P. Wang, *Interplay between CRP-cAMP and PII-Ntr systems forms novel regulatory network between carbon metabolism and nitrogen assimilation in Escherichia coli,* Nucleic Acids Research **35**, 1432 (2007).

[194] L. Paul, P. K. Mishra, R. M. Blumenthal, and R. G. Matthews, *Integration of regulatory signals through involvement of multiple global regulators: Control of the Escherichia coli gltBDF operon by Lrp, IHF, Crp, and ArgR,* BMC Microbiology **7**, 1 (2007).

[195] Z. Zhang, G. Gosset, R. Barabote, S. Gonzalez, W. A. Cuevas, M. H. S. Jr, C. S. Gonzalez, and M. H. Saier, *Functional Interactions between the Carbon and Iron Utilization Regulators , Crp and Fur , in Escherichia coli Functional Interactions between the Carbon and Iron Utilization Regulators , Crp and Fur , in Escherichia coli,* Journal of Bacteriology **187**, 980 (2005).

[196] S. Sinha, A. D. S. Cameron, and R. J. Redfield, *Sxy Induces a CRP-S Regulon in Escherichia coli,* Journal of Bacteriology **191**, 5180 (2009).

[197] H. Hirakawa, Y. Inazumi, Y. Senda, A. Kobayashi, T. Hirata, K. Nishino, and A. Yamaguchi, *N-acetyl-D-glucosamine induces the expression of multidrug exporter genes, mdtEF, via catabolite activation in Escherichia coli,* Journal of Bacteriology **188**, 5851 (2006).

[198] N. De Lay and S. Gottesman, *The crp-activated small noncoding regulatory RNA CyaR (RyeE) links nutritional status to group behavior,* Journal of Bacteriology **191**, 461 (2009).

[199] J. M. Hudson and M. G. Fried, *Co-operative interactions between the catabolite gene activator protein and the lac repressor at the lactose promoter,* Journal of Molecular Biology **214**, 381 (1990).

[200] C. L. Lawson, D. Swigon, K. S. Murakami, S. A. Darst, H. M. Berman, and R. H. Ebright, *Catabolite activator protein: DNA binding and transcription activation,* Current Opinion in Structural Biology **14**, 10 (2004).

[201] D. Gillespie, *Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral.* Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics **54**, 2084 (1996).

[202] T. M. Ramseier, S. Bledig, V. Michotey, R. Feghali, and M. H. Saier, *The global regulatory protein FruR modulates the direction of carbon flow in Escherichia coli,* Molecular Microbiology **16**, 1157 (1995).

[203] K. Kochanowski, B. Volkmer, L. Gerosa, B. R. Haverkorn van Rijsewijk, A. Schmidt, and M. Heinemann, *Functioning of a metabolic flux sensor in Escherichia coli.* Proceedings of the National Academy of Sciences of the United States of America **110**, 1130 (2013).

[204] A. F. Alvarez and D. Georgellis, *In Vitro and In Vivo Analysis of the ArcB/A Redox Signaling Pathway,* Methods in Enzymology **471**, 205 (2010).

[205] M. Wang, C. J. Herrmann, M. Simonovic, D. Szklarczyk, and C. von Mering, *Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines,* Proteomics **15**, 3163 (2015).

[206] S. S. Chen and J. R. Williamson, *Characterization of the ribosome biogenesis landscape in E. Coli using quantitative mass spectrometry,* Journal of Molecular Biology **425**, 767 (2013).

[207] T. E. Norris and A. L. Koch, *Effect of growth rate on the relative rates of synthesis of messenger, ribosomal and transfer RNA in Escherichia coli,* Journal of Molecular Biology **64**, 633 (1972).

[208] B. S. Schuwirth, M. A. Borovinskaya, C. W. Hau, W. Zhang, A. Vila-Sanjurjo, J. M. Holton, and J. H. D. Cate, *Structures of the Bacterial Ribosome at 3.5 A Resolution,* Science (2005).

[209] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, *The protein data bank.* Nucleic acids research **28**, 235 (2000).

[210] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin, *UCSF Chimera - A visualization system for exploratory research and analysis,* Journal of Computational Chemistry **25**, 1605 (2004), arXiv:arXiv:1011.1669v3 .

[211] F. H. Crick, *On protein synthesis.* Symposia of the Society for Experimental Biology **12**, 138 (1958), arXiv:arXiv:1011.1669v3 .

[212] E. Lawrence, *Henderson's Dictionary of Biology,* 13th ed. (Pearson Education Limited, Harlow, 2005).

[213] D. E. Koshland Jr., *The Seven Pillars of Life,* Science **295**, 2215 (2002).

[214] G. E. Palade, *A small particulate component of the cytoplasm.* The Journal of biophysical and biochemical cytology **1**, 59 (1955).

[215] N. A. Campbell and J. B. Reece, *Biology,* sixth edit ed. (Pearson Education, Inc, San Fransisco, 2002).

[216] R. Iizuka, M. Yamagishi-Shirasaki, and T. Funatsu, *Kinetic study of de novo chromophore maturation of fluorescent proteins.* Analytical biochemistry **414**, 173 (2011).

[217] T. Preiss, *All Ribosomes Are Created Equal. Really?* Trends in Biochemical Sciences (2016).

[218] N. Slavov, S. Semrau, E. Airoldi, B. Budnik, and A. van Oudenaarden, *Differential Stoichiometry among Core Ribosomal Proteins,* Cell Reports **13**, 865 (2015), arXiv:1406.0399 .

[219] I. Shachrai, A. Zaslaver, U. Alon, and E. Dekel, *Cost of Unneeded Proteins in E. coli Is Reduced after Several Generations in Exponential Growth,* Molecular Cell **38**, 758 (2010).

[220] C. Condon, S. French, C. Squires, and C. L. Squires, *Depletion of functional ribosomal RNA operons in Escherichia coli causes increased expression of the remaining intact copies,* Embo J **12**, 4305 (1993).

[221] C. Condon, D. Liveris, C. Squires, I. Schwartz, and C. L. Squires, *Ribosomal-Rna Operon Multiplicity in Escherichia-Coli and the Physiological Implications of Rrn Inactivation,* Journal of Bacteriology **177**, 4152 (1995).

[222] S. Quan, O. Skovgaard, R. E. McLaughlin, E. T. Buurman, and C. L. Squires, *Markerless Escherichia coli rrn Deletion Strains for Genetic Determination of Ribosomal Binding Sites,* G3: Genes, Genomes, Genetics **5**, 2555 (2015).

[223] T. Bollenbach, S. Quan, R. Chait, and R. Kishony, *Nonoptimal Microbial Response to Antibiotics Underlies Suppressive Drug Interactions,* Cell **139**, 707 (2009), arXiv:NIHMS150003 .

[224] Z. Gyorfy, G. Draskovits, V. Vernyik, F. F. Blattner, T. Gaal, and G. Posfai, *Engineered ribosomal RNA operon copy-number variants of E. coli reveal the evolutionary trade-offs shaping rRNA operon number,* Nucleic Acids Research **43**, 1783 (2015).

[225] S. Quan, O. Skovgaard, and C. L. Squires, *Escherichia coli Ribosomal RNA deletion strain*, Ph.D. thesis, Stanford University (2013).

[226] M. Kaczanowska and M. Ryden-Aulin, *Ribosome Biogenesis and the Translation Process in Escherichia coli,* Microbiology and Molecular Biology Reviews **71**, 477 (2007).

[227] Z. Shajani, M. T. Sykes, and J. R. Williamson, *Assembly of bacterial ribosomes.* Annual review of biochemistry **80**, 501 (2011).

[228] H. M. Salis, E. A. Mirsky, and C. A. Voigt, *Automated design of synthetic ribosome binding sites to control protein expression.* Nature biotechnology **27**, 946 (2009).

[229] H. M. Salis, *The ribosome binding site calculator,* Methods in Enzymology **498**, 19 (2011).

[230] M. Maeda, T. Shimada, and A. Ishihama, *Strength and Regulation of Seven rRNA Promoters in Escherichia coli,* PLoS ONE **10**, 1 (2015).

[231] D. Zaporojets, S. French, and C. L. Squires, *Products Transcribed from Rearranged rrn Genes of Escherichia coli Can Assemble To Form Functional Ribosomes Products Transcribed from Rearranged rrn Genes of Escherichia coli Can Assemble To Form Functional Ribosomes,* Journal of bacteriology **185**, 6921 (2003).

[232] R. Milo, P. Jorgensen, U. Moran, G. Weber, and M. Springer, *BioNumbers The database of key numbers in molecular and cell biology,* Nucleic Acids Research **38**, 750 (2010).

[233] S. K. Lee, H. Chou, T. S. Ham, T. S. Lee, and J. D. Keasling, *Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels,* Current Opinion in Biotechnology **19**, 556 (2008).

[234] P. E. Savakis, S. A. Angermayr, and K. J. Hellingwerf, *Synthesis of 2,3-butanediol by Synechocystis sp. PCC6803 via heterologous expression of a catabolic pathway from lactic acid- and enterobacteria,* Metabolic Engineering **20**, 121 (2013).

[235] M. Wehrens, P. R. ten Wolde, and A. Mugler, *Positive feedback can lead to dynamic nanometer-scale clustering on cell membranes,* Journal of Chemical Physics **141** (2014), 10.1063/1.4901888.

[236] T. R. Sokolowski, J. Paijmans, L. Bossen, M. Wehrens, T. Miedema, N. B. Becker, K. Kaizu, K. Takahashi, M. Dogterom, and P. R. ten Wolde, *eGFRD in all dimensions,* arXiv (2017), arXiv:1708.09364 .

Acknowledgements

Not only the nature of the work, but also the environment in which you work is important for your motivation and happiness. I'm glad to have done my PhD at AMOLF, which provided a lively social scene and is populated by friendly and helpful people.

First and foremost I'd like to thank Sander here, for providing me the opportunity to work on the fascinating topics described in this thesis and being my mentor the past years. In our many discussions, perhaps after a slightly convoluted argument from me, you have often and repeatedly asked "but what is the key question?". I have more and more come to appreciate the wisdom of asking this question, and the importance of answering it. Why are we doing what we are doing and what is the use of it? Being able to precisely and concisely formulate the answer to this question, is a key quality to practice science well. You have also taught me to carefully look at data, and ask what answers are in there — answers that not necessarily answer our original question, but perhaps another, more interesting one. Cynics might call this quality salesmanship, but I think that undervalues the importance of being able to assess the value of features that we observe in our scientific data and place them in a larger context.

I am also glad to have worked with so many great people. My (former) colleagues from the lab who worked on the same topic: Noreen, Dmitry, Ferhat, Katja, Sebastian and Rutger R.. Noreen, thank you for your rigorous explanations of how to do experiments, and for being an example of how to work neatly and accurately. Dmitry, thank you for showing me how to operate microfluidic devices, I especially enjoyed (and learned from) our macGyver sessions trying to improve the device, I also greatly enjoyed our political discussions. Ferhat, without you the device that I mundanely refer to as the "microfluidic device 2" would never have been a success, I really appreciate our collaboration, and your "let's do it" mentality. Glorious. Vanda, thanks for all your hard work on the constructs, but more importantly I appreciate our conversations about miscellaneous life matters. Katja, I deeply appreciate your sharp mind and insightful observations; I am grateful for our discussions (usually at the coffee corner) where you often had great advice on professional/scientific matters. Sebastian, thank you for your insightful input, and for being an example of how it sometimes works very well to be headstrong. Rutger R., I thank you for your patience in explaining me experimental protocols; also, I have learned from your people skills.

A special thanks goes out to the two students I had the privilege to supervise, Nick and Giulia B. I hope you have learned from me. I have learned a tremendous deal from supervising you. It was very insightful to see projects through the eyes of a supervisor. I also like to think that I have grown as a teacher, because I had to think about the best way to convince you that approach X or Y was the best. Nick, I greatly admire your work ethics and skill in the lab, and greatly appreciate your positive attitude. Giulia B., I really enjoyed our discussions, both on science and life in general. Thanks for sharing your sharp mind.

Fatemeh, I enjoyed our insightful (political) discussions. Mario, thanks for making sure there was always the right music to dance to. Eline, wherever we end up in our future careers, at least we can have lab outing event planning as backup career. Xuan, Anne-Bart, Florian, take good care of the legendary office 1.37! And thanks to all the other (former) colleagues in the Tans lab for making the lab a bit more smart and colorful: Rebecca, Roeland, Moshin and David.

I also like to thank some people from outside AMOLF. Laurens, I enjoyed our brain-storm sessions, it was a great pleasure to think about cellular dynamics with you. Rutger H., thank you for sharing your sharp insights. Benjamin, thanks for helpful discussions, it was great meeting you in Heidelberg. Also thanks to Brian Munsky, who put his heart in organizing a great, fun summer school at CSU in Fort Collins, CO, which helped greatly in developing theoretical aspects of my research. Petra, not often does a PhD student get a professor as student; it was great fun showing you our setups and analysis. Moreover thanks for the opportunity to visit your lab and hosting Sonja and me in St. Louis; this was a great experience. I hope I can become as open minded and enterprising in both family and lab life as you. Thanks for all your helpful input and advice.

AMOLF, providing friendly people, is an excellent place to make friends. I am very greatful that I become good friends with Joris, Noreen, Adithya, Stephen, Thomas and Johannes. With the first five, I did not only spend many boardgame evenings, escaped from many escape rooms, and discovered great Dutch nature. Moreover we often enjoyed lively conversations about many topics, and shared many frustrations and accomplishments. Is it done Joris? No comrade premier, it has only begun. (Thanks Joris, you were a huge factor in making my years at AMOLF so great!) Noreen, I secretly think you're the best boardgame player. I admire your resolve and discipline. Adithya, you make the magic happen. I'm jealous of that ability. Thanks for all our extensive conversations, your advice is invaluable. Stephen, I could not have wished for a better person to share a home with. Thomas, I envy your ability to listen, analyze and make concise arguments in split seconds. Johannes, once *sjaars* (freshman) in office 1.37, I'm happy you were quickly part of our office's old boys club. Though superficially it might appear we have different views on many matters, we have a shared passion for political debate, sound reasoning based on facts, and getting to the bottom of things. I'm sure killing plants is unethical though.

And of course I thank all other people that gave office 1.37 its unique atmosphere. Mark, thanks for being the office mate with whom I could finally talk to on *vpro gracht-engordel elite* level. And also Rik, Robin, Esther, and Roland — it was great sharing the office with you!

Furthermore, importantly, Agata, we had similar experiences. We started together at AMOLF and ended together. Thanks for the great conversations about both our journeys. And many thanks for helping me ski a little bit better! Nicola, many thanks for your hospitality, and thanks for tolerating my slow skiing, especially when hung over. Thanks also to Harmen, Yuval, Simone K., Federica, Bart, Christina, Magdalena, Jeanette, Matthijs, Viktoria, Michele, Jacopo, Giulia M., Anders, Simone B., Marjolein and Yvonne — who each uniquely contributed to making life in and around AMOLF great.

A special thanks also to Floortje, Niels, Tom and Clyde, with whom I greatly enjoyed organizing AMOLF events in the *peevee* (PV). Floortje, I think you were a field marshal in

a previous life, because organizing comes so natural to you.

A great thing about AMOLF is that interaction between people is encouraged. Also research leaders often join for coffee break or lunch break, and contribute to lively discussions about the topic of the day. Pieter Rein, Bela, Jeroen, and others, I hope you keep up this tradition.

Also thanks to all great support staff, especially Hinco and Marko, who were always willing to help with any technical issues I experienced, and always decently made sure the issue was solved properly.

A special thanks goes to Andrew Mugler, my supervisor during my MSc internship, whose good example of dealing with the ups and downs of scientific progress helped me to stay balanced during this PhD project.

A very special thanks also goes to the late At (Ton) de Waart, who applied the belief that anyone can understand anything at a children's technical club, where fantastical electronic, mechanical engineering and programming ideas came alive. I am indebted to your commitment.

There are likely people that I have not mentioned here who have also contributed to making my PhD a great experience. Also to you, I am grateful.

Finally, I thank my good friend Rob, for providing me with sometimes hard-needed distraction and your sympathetic ear.

Then, I thank my parents, who taught me the value of good education, always wholeheartedly supported me throughout any endeavor and encouraged my enthusiasm for science. I thank my aunt Wilma, who also always encouraged me in my studies. I also thank my brother Erik. Our career paths are very different, but we think (and talk) very much alike, which explains why we can have such comfortable conversations in mutual understanding. And lastly, I thank my girlfriend Sonja. You understood what support I needed during setbacks and challenges, often tolerated my quirks (both general and PhD-related), and made sure accomplishments were celebrated accordingly. I am also impressed with the level of insight you showed in my scientific projects. You even pointed out a mistake in one of my presentations. I conclude these acknowledgments with the observation that we met at the end of this adventure of mine, but we are just starting a new adventure together.

CURRICULUM VITAE

Martijn Wehrens (April 3rd, 1986) was born in Nijmegen, where he attended high school at the *Nijmeegse Scholengemeenschap Groenewoud.* He always had an interest in science and technology, and was at that time for example working on building and programming moving little machines at an engineering club. To understand the most complicated machine of all, life, he started to study biology at the Radboud University in Nijmegen in 2004. In 2007 he realized that to fully grasp fundamental processes in the biological cell, one should delve deeper into the biochemical processes that go on in the cell. Therefore, he began to study chemistry. In addition to his studies he also participated in the Honours program, with elective courses on free will, holy texts, community, and economics; worked as editor for the student magazine ANS; and was a member of the student council. After finishing both the biology and chemistry bachelor program in 2010, he wanted to further understand the cell at the fundamental level. He therefore joined the MSc program in chemistry at the University of Amsterdam. During this theoretical program he focused on computer simulations of chemical processes. This was also the topic of his MSc thesis, where he extended and used the eGFRD reaction-diffusion simulation algorithm to show that positive feedback in chemical reactions can contribute to clustering of receptor molecules on the surface of biological cells. This research was done in the Biochemical Networks research group of Prof. Pieter Rein ten Wolde at the AMOLF research institute in Amsterdam. After completing the MSc program in chemistry in 2013, he further pursued his interest in biochemical networks by starting to work as a PhD student in the Prof. Sander J. Tans Biophysics group at AMOLF in January of 2014. The work on dynamical regulation in single cells he performed as a PhD student is described in this thesis.

Martijn Wehrens.

Photo by Wouter Harmsen.

PUBLICATIONS BY THE AUTHOR

RELATED TO THIS THESIS

- M. Wehrens, D. Ershov, R. Rozendaal, N. Walker, D. Schultz, R. Kishony, P. A. Levin, and S. J. Tans, *Size Laws and Division Ring Dynamics in Filamentous Escherichia coli cells*, Current Biology , 1 (2018). [51]
  https://doi.org/10.1016/j.cub.2018.02.006.
  See chapter 3.

- M. Wehrens, F. Buke, P. Nghe, and S. J. Tans, *Stochasticity in cellular metabolism and growth: Approaches and consequences*, Current Opinion in Systems Biology 8, 131 (2018). [107]
  http://doi.org/10.1016/J.COISB.2018.02.006.
  See chapter 4.

OTHER PUBLICATIONS

- M. Wehrens, P. R. ten Wolde, and A. Mugler, *Positive feedback can lead to dy- namic nanometer-scale clustering on cell membranes*, Journal of Chemical Physics 141 (2014), 10.1063/1.4901888. [235]

- T. R. Sokolowski, J. Paijmans, L. Bossen, M. Wehrens, T. Miedema, N. B. Becker, K. Kaizu, K. Takahashi, M. Dogterom, and P. R. ten Wolde, *eGFRD in all dimensions*, arXiv (2017), arXiv:1708.09364. [236]
  Manuscript in preparation.