CENTRO DE INVESTIGACIÓN Y ESTUDIOS AVANZADOS DEL INSTITUTO POLITÉCNICO NACIONAL

Unidad Monterrey
Ingeniería y Física Biomédica

Study on the dynamic behavior of *Escherichia coli*'s tryptophanase operon

Thesis dissertation by

David Israel Orozco Gómez

to obtain the degree of

Doctor in Sciences
in

Biomedical Physics and Engineering

Thesis supervisor: Dr. Moisés Santillán Zerón

Apodaca, Nuevo León

August 2018

Operons are always more complicated than you think

Kevin D. Young, 2016

Acknowledgments

First and foremost, I want to thank CONACYT for the National Postgraduate Studies Grant N° 394293. I also want to acknowledge Cinvestav–Monterrey for giving me the opportunity and aid to pursue my formation towards a PhD in Biomedical Physics and Engineering.

I cannot find the proper words to express gratitude to my advisor, Dr. Moisés Santillán. During the development of this project there were many dead ends and misleading clues. Ironically, I—a biologist—was always on the look for the ideal solutions and absolute results; Dr. Moisés—a physicist—had the patience to always remind me of the implausibility of spherical cows. His guidance, counsel and teachings helped me to lay what I hope to be the solid foundations of an upcoming scientific career. I can say, without a hint of doubt, I am proud of having been his student.

The project's success is due largely to the support and kindness of Dr. Kevin D. Young and his lab team at UAMS in Little Rock, AR. He gave me the opportunity to work a short time at his lab and went above and beyond to provide me with all the means necessary to carry out the experiments I needed to perform. I want to thank Dr. Suresh Kannan for his guidance during my stay there; his knowledge of cell and molecular biology, technical skill and patience assured the best possible outcome of this experience. I need to also mention William, Matt and Dina for their warm welcome and the great work environment they provided. I wish you all the best.

The collaboration with Dr. Eduardo Sosa was essential at many points during the project, from the early design proposals for the microreactor to the final image analysis and data acquisition.

Undoubtedly, his contributions paved the way to reach more than one milestone along this road. His analytical abilities, his eagerness to help, and his sincere friendship were pivotal at times that certainly felt critical. Thank you, Lalo; I hope to someday have the opportunity to pay you back.

Material and technical support in microfludics, kindly provided by Dr. José Luis García and his team at bioARTS lab, were an important key to the development of this work. Overall, I want to acknowledge Manolo de Hoyos and Alan González for their helpfulness and willingness to share their time and experience.

QBio lab was my workplace during four years. I saw its beginnings and early growth. I'll always remember the first members: Óscar Gallardo, Jorge Arias and Daniela Nevárez. We somehow managed to help establish a research lab and keep it running without major issues. I guess I can say we are not a bad team at all, even if we all do very different kinds of research. I have to bring a spotlight to Daniela, she had to learn from me; luckily she's got a brilliant mind and managed to adapt and produce a great work despite her inexperience and my terrible teaching skills.

Finally, I want to thank Dr. Daniel Sánchez, Dr. Jesús Rodríguez, Dr. Jesús Santana from Cinvestav–Monterrey, and Dr. Víctor Treviño from ITESM—my thesis committee—for their notes, comments and guidance to the final version of this thesis dissertation.

Contents

List of Tables

List of Figures

List of Terms and Abbreviations

AC

Adenylate cyclase, membrane bound enzyme that catalyzes ATP from cAMP. 11, 69

AroP

Phenylalanine/tyrosine/tryptophan transporter from the APC transporter superfamily. 16, 69

ATP

Adenosine triphosphate. 11, 13

cAMP

Cyclic adenosine monophosphate. 11, 40, 46, 49, 69

CAP

Catabolite activator protein. 11, 40, 46, 49, 69

DNA

Deoxyribonucleic acid. 11, 13

EI

Enzyme I, second enzymatic element in the PTS. 10

EII

EII enzyme group, includes all the remaining enzymatic elements in the PTS. Contrary to the first two elements, there is one specific EII complex per each carbohydrate involved in the catabolite repression. 10

EIIA

Enzyme II A, element of the EII group. 10, 11

EIIA^{gluc} ∼P

Phosphorilated EIIA^{gluc} enzyme. 11

EIIA^{gluc}

Glucose–specific EIIA enzyme. 10–12, 69

EIIB^{gluc}

Glucose–specific membrane–bound II B enzyme, element of the EII group. 10, 12

EIIBC^{gluc}

Glucose–specific membrane–bound II BC enzymatic complex, it is in charge of importing exogenous glucose into the cytoplasm and transferring it a phosphate molecule–thus beginning the glycolysis pathway. 10, 11

EIIC^{gluc}

Glucose–specific membrane–bound II C enzyme, element of the EII group. 10, 12

FACS

Fluorescence activated cell sorting. 44

GFP

Green fluorescent protein. 22, 29

HPr

Phosphohistidine carrier protein, first enzymatic element in the PTS. 10

IEI

Inducer exclusion inhibition. 10, 11, 69

IPTG

Isopropyl β-D-1-thiogalactopyranoside. Gratuitous inducer of the *lac* operon.. 47

M9

M9 minimal medium. 30, 31, 37, 40, 41

mRNA

messenger RNA. 12, 13, 35

Mtr

Tryptophan specific transport protein, does not belong to the *tna* operon regulatory pathway. 16, 69

OD_{600}

Optical density at 600 nm. 30, 37, 42, 43

ON

Over night. 31

PBS

Phosphate–buffered saline. 29, 31

PDF

Probability distribution function. 30, 42, 45, 46, 48, 50, 51

PDMS

Polydimethylsiloxane. 28, 31

PEP

Phosphoenolpyruvate molecule. 10, 11

PTS

Phosphoenolpyruvate–carbohydrate phosphotransferase system. 10, 69

PVP–40

Polyvinylpyrrolidone. 28, 31

RNA

Ribonucleic acid. 11

RNApol

RNA polymerase. 11, 13, 69

sfGFP

super folder GFP. 29

Si

Silicon chemical symbol. 30

TnaA

Tryptophanase enzyme, product of the *tnaA* gene in the *tna* operon. 9, 14, 15, 24, 35, 36, 45, 49, 54

TnaB

Tryptophan specific transport protein, product of the *tnaB* gene in the *tna* operon. 9, 16, 17, 35, 69

tRNA^{trp}

Transfer RNA unit loaded with a tryptophan molecule. 13

Summary

The proteins encoded by the genes in *Escherichia coli*'s *tna* operon are in charge of metabolizing tryptophan as a carbon source under certain medium conditions. It is regulated at the promoter level by catabolite repression and, at the transcription elongation level, by Rho-mediated termination through a binding site at *tnaC*—a coding sequence for a leader peptide. The latter mechanism can be stalled by the presence of tryptophan. This regulatory arrangement implies that a hierarchy of sugars, mainly glucose, act as repressors of the system and that tryptophan acts as system inducer.

It is relevant to the scope of this project to point that this kind of architecture—namely, the combination of catabolite repression and self up-regulation—is shared among many known systems in *E. coli* in charge of the metabolism of carbon sources alternative to glucose. One of them in particular, the *lac* operon has been one of the most widely studied genetic expression systems in bacteria and it has been proposed to be bistable; however, bistability has never been convincingly proven in an experimental approach.

The similarity between the *lac* and *tna* operons gives rise to the possibility that the latter may also be bistable, however, there is scarce information on its dynamics and the bistability hypothesis cannot still be addressed. We intend to add to the discussion by gathering experimental data that can lead toward accurate predictions on whether bistable behavior can arise in biological systems.

The *tna* operon of *E . coli* comprises two genes, *tnaA* and *tnaB*. The first codes for tryptophanase and the second for a low affinity tryptophan specific permease, TnaB. Tryptophanase degrades the

tryptophan imported by TnaB into three byproducts of metabolic importance. The dynamics of this system imply the emergence of two self regulation loops, a positive loop related to TnaB and a negative loop related to tryptophanase activity. The positive one makes the operon a candidate for bistability. Bistability arises when two stable fixed points coexist in a system. A fixed point is that in which a system presents no changes over time. From a biological point of view, a fixed point corresponds to a steady state, in which the response of an organism to constant conditions in its medium remain unchanged.

Recently, it has been reported that inactive tryptophanase forms a single focus at the cell poles during lag and exponential growth phases. Upon reaching the stationary phase, the foci disaggregate, tryptophanase forms tetramers and activates. The *E. coli* GL69 strain allows to observe this phenomenon through a *tnaA-sfGFP* genomic fusion. Further inquiries on the post-translational regulation of this system evidenced a previously undisclosed, and still barely studied, role of glucose on the activation of the tryptophanase enzyme. It is still to be determined if the presence of tryptophan also affects its activation.

Here, we aim to experimentally determine the existence of a bistability region given by the environmental conditions of the bacteria. To do so, we intended to study the dynamic behavior of the *tna* operon, driving it to steady states under different growth medium compositions. To do so, we developed a microreactor to cultivate *E. coli* GL69 under constant glucose and tryptophan concentrations to drive it to stationary states under several different conditions and observe its expression under fluorescence microscopy. Results suggest that bistability emerges under a certain range of tryptophan and glucose combinations. Furthermore, we found evidence of an unknown post-translational regulation mechanism. In this regard, our evidence implies that both the presence of glucose and tryptophan regulate the disgregation of the TnaA foci and the subsequent enzymes activation. Possible biological implications of these findings are further discussed.

Resumen

Las proteínas codificadas en el operón *tna* de *Escherichia coli* están a cargo de metabolizar triptófano como una fuente de carbono bajo ciertas condiciones de crecimiento. A nivel promotor se encuentra regulado por represión catabólica y, a nivel transcripcional, por represión mediada por Rho. Esta última ocurre a través de un sitio de unión contenido en la secuencia *tnaC*, que codifica para un péptido líder, y puede ser impedida por la presencia de triptófano en el citoplasma. Este arreglo regulatorio implica que existe una jerarquía de azúcares, siendo la glucosa el más importante, que actúan como represores del sistema y, al mismo tiempo, el triptófano juega un papel de inductor.

Dado el enfoque del proyecto, es relevante señalar que este tipo de arquitectura (la combinación de represión catabólica y autoinducción) es compartido por varios sistemas de expresión genética de *E. coli*; varios de ellos a cargo del metabolismo de fuentes de carbono alternativas a la glucosa. Uno de ellos en particular, el operón *lac*, es uno de los sistemas de expresión y regulación genética bacterianos más estudiados. Se ha propuesto que este es biestable; sin embargo, la biestabilidad no ha podido ser comprobada convincentemente mediante una estrategia experimental.

La semejanza entre los operones *lac* y *tna* sugiere que este último también podría ser biestable; sin embargo, existe poca información sobre su comportamiento dinámico y la hipótesis de biestabilidad no puede ser abordada satisfactoriamente. Nuestro objetivo es obtener datos experimentales que puedan aportar evidencias precisas respecto a la posibilidad del comportamiento biestable en sistemas biológicos.

El operón *tna* de *E. coli* incluye dos genes: *tnaA* y *tnaB*. El primero codifica para la triptofanasa y el segundo para una permeasa específica de triptófano de baja afinidad, TnaB. La triptofanasa degrada el triptófano importado por TnaB en tres subproductos de importancia biológica. La dinámica de este sistema implica la existencia de dos ciclos de autoregulación: uno positivo relacionado a la actividad de TnaB y uno negativo relacionado a la triptofanasa. El ciclo positivo de autoregulación sugiere que este operón es candidato a biestabilidad. Esta ocurre cuando dos puntos fijos estables coexisten en un sistema. Un punto fijo es aquel en el cual dicho sistema no cambia en el tiempo, permanece estático. Desde un punto de vista biológico, un punto fijo corresponde a un estado estacionario, en el cual la respuesta de un organismo a condiciones constantes en su medio permanece igualmente constante.

Recientemente, se ha reportado que la triptofanasa inactiva forma un gránulo en uno de los polos celulares durante la fases de latencia y exponencial del crecimiento celular. Una vez que las células bacterianas alcanzan la fase estacionaria, dicho gránulo se disgrega, la triptofanasa libre forma homotetrámeros y se activa. La cepa GL69 de *E. coli* permite observar este fenómeno mediante una fusión genómica *tnaA-sfGFP*. Investigaciones posteriores sobre la regulación post-traduccional del operón reportaron evidencias sobre el papel de la glucosa en la activación de la triptofanasa, una cuestión hasta entonces desconocida y, en consecuencia, poco investigada actualmente. Aún queda por determinar si el triptófano posee igualmente un rol en el mecanismo de activación de dicha enzima.

En este proyecto, nos propusimos determinar experimentalmente la existencia de una región de biestabilidad en la dinámica de regulación del operón *tna*, dada por las condiciones ambientales de *E. coli*. Para lograrlo, implementamos un estudio experimental del estudio dinámico de este operón, forzándolo a alcanzar estados estacionarios en medios de crecimiento de diferentes composiciones. Desarrollamos un microrreactor para cultivar la cepa *E. coli* GL69 en diferentes concentraciones constantes de glucosa y triptófano. Observamos la expresión del operón mediante microscopía de fluorescencia.

Los resultados sugieren que la biestabilidad emerge en un rango delimitado de combinaciones concentraciones de ambos reactivos. Además, encontramos evidencia de la un mecanismo de

regulación post-traaduccional; en este, la disolución de los gránulos de triptofanasa (así como su subsecuente activación) requieren la presencia tanto de glucosa como de triptófano. Las posibles implicaciones biológicas de estos hallazgos son discutidas a lo largo de este texto.

1 Introduction

Genetic regulation is a product of species evolutive history. It stems from the instructions recorded in each cell's genome. At the same time, it encompasses all the mechanisms required to maintain it and properly implement those instructions. It is a complex phenomenon that spans several levels of living matter organization. Describing and understanding it has proven to be a difficult task (Jacob and Monod, 1961; Tyson et al., 2003).

The structure and function of genes have long been discussed and investigated, sometimes even redefined (Pearson, 2006). However, in the realm of genetic regulation they are mere elements of different systems, notwithstanding their relevance to cell survival. This is because a gene by itself seldom presents emergent behavior, possibly the most conspicuous property of any kind of system (Trewavas, 2006).

Operons are genetic regulation systems. Their existence and structure was hypothesized to explain phenomena like diauxic growth in *Escherichia coli*, in which bacteria growing in a mixture of two different carbon sources would first consume one and then another (Monod, 1949). On the molecular level, this is not a simple behavior; the sudden switch between carbon sources implies a drastic change in the metabolic pathways used by the cell to generate energy and ensure survival. Monod found several pairs of carbohydrates with which *E. coli* behaved this way; in later works Jacob et al. (1960) found evidence of multiple genes working in coordination to trigger these metabolic changes. They deemed the expression of these genes was coordinated by a specific element called 'operator'. They called operons these units of coordinated expression. In other words, an operon is a set of genes whose expression is controlled by a single promoter.

Operons were first described and believed to be exclusive of prokaryotic cells, however, they were later proven to also be present in eukaryotic genomes, although with slightly different structures (Blumenthal, 2004). They are important because much of the cellular sensing and adaption mechanisms are products of operons (Jacob and Monod, 1961). As such, they provide a deep insight into the dynamics of genetic regulation (Yanofsky, 1981).

In terms of complexity, operons present a compromise between full scale genomes and single genes under constitutive expression. In the case of genomes, their sheer amount of information renders them hard to study efficiently; on the other side, it is possible to make very detailed and precise observations from constitutively expressed genes, although the scantness of the information they offer turns them useless as a model for the more abundant complex cases of genetic regulation. This project is a study on the dynamic behavior of *Escherichia coli*'s tryptophanase operon.

1 *Escherichia coli*'s *tna* operon

E. coli's *tna* operon comprises a sequence called *tnaC* and two genes: *tnaA* and *tnaB*. The schematics in Fig. 1.1 describe the dynamics of its elements. The *tna* operon is in charge of processing extracellular tryptophan as a carbon and nitrogen source in the absence of glucose and other sugars. Its promoter activity is mainly regulated by catabolite repression (Gong and Yanofsky, 2002). *tnaC* is a coding sequence whose product is called 'leader peptide'. It interacts with tryptophan and the Rho (ρ) enzyme to regulate the expression of *tnaA* and *tnaB* (Konan and Yanofsky, 1997).

The *tnaA* gene encodes tryptophanase, an enzyme that degrades tryptophan into three by-products: indole, pyruvate and ammonia. Pyruvate and ammonia are used in different metabolic pathways involving energy generation and protein biosynthesis. Indole is a low molecular weight compound involved in intercellular communication mechanisms, like quorum sensing (Piñero-Fernandez et al., 2011).

The *tnaB* gene expresses a low affinity permease that brings extracellular tryptophan into the cell. This is not the only permease involved in the tryptophan uptake; however, it has been reported that it is the only one expressed under the medium conditions in which the *tna* operon usually works (Li and Young, 2013).

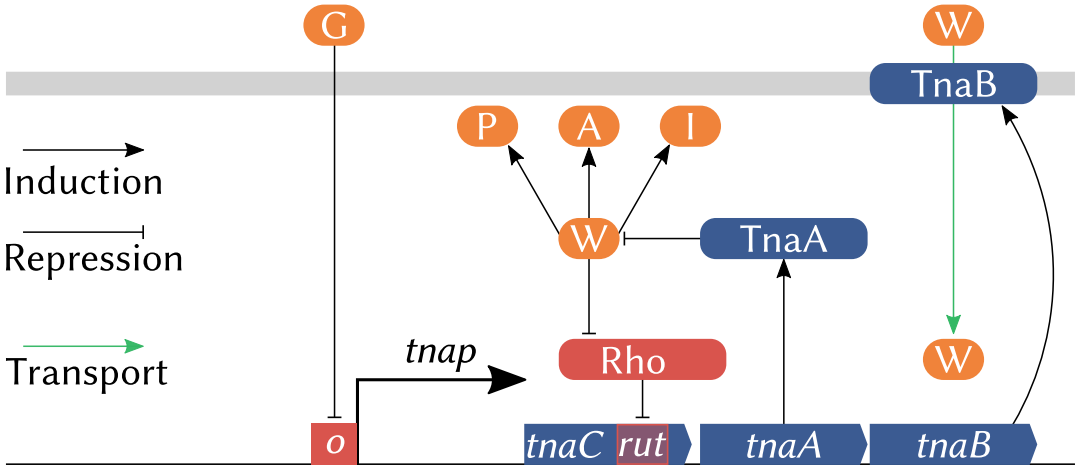


Figure 1.1: The *Escherichia coli*'s *tna* operon. The operon's regulation pathway begins at the promoter level (*tnap*), here denoted by a folded black arrow; the presence of glucose (G) in the cytoplasm inhibits the promoter activation by the means of catabolite repression—the black flat-headed arrow denotes repression and the black spear-headed ones stand for induction. The first element downstream of *tnap* is the coding sequence *tnaC*. It holds a Rho-binding site (*rut*); Rho is an enzyme involved in the early translation termination of many *E. coli*'s genes and operons, including the *tna* operon. Tryptophan (W) inhibits Rho's termination activity over the operon. The last coding elements in the operon are the genes *tnaA* and *tnaB*. The former's product is the enzyme tryptophanase (TnaA), whose function is to process tryptophan into three byproducts of biological importance: pyruvate (P), ammonia (A), and indole (I). *tnaB* codes for a tryptophan-specific low-affinity permease (TnaB), which imports extracellular tryptophan—transport into the cytoplasm is denoted by a green arrow—and is the only tryptophan permease active in the conditions in which the *tna* operon is expressed.

1.1 Catabolite repression and the *tna* operon

Catabolite repression is one of the most important genetic regulation mechanisms in several bacterial species, including *E. coli* (Brückner and Titgemeyer, 2002; Escalante et al., 2012). It originates in the phosphoenolpyruvate–carbohydrate phosphotransferase system (PTS), a group of transporter proteins that participate in the translocation and phosphorylation of different sugars. The details of this process will be discussed in the next paragraphs; resort to Fig. 1.2 for graphical aid on the details.

PTS begins when a phosphoryl group from a phosphoenolpyruvate molecule (PEP) is transferred to the EI enzyme, which then transfers it to the phosphohistidine carrier protein (HPr). These two enzymes are non-specific and their activity occurs in the cytoplasm, just as that of the EIIA enzyme. When the available carbon source processed by the PTS is glucose, HPr phosporylates a glucose specific EIIA (EIIA^{gluc}), which then cedes the phosphoryl group and activates the EIIBC^{gluc} complex. This is formed by the membrane bound enzymes EIIB^{gluc} and EIIC^{gluc} and it is in charge of translocating glucose from the periplasm into the cytoplasm (Gosset, 2005).

After the EIIBC^{gluc} complex is activated, dephosphorylated EIIA^{gluc} presents affinity to other similar complexes and permeases. However, in these its function is of inhibition rather than induction—it blocks the transport of other sugars into the cell.

It has been reported the existence of at least 21 EII complexes, each of those matching different carbon sources. Thus, PTS establishes a hierarchical consumption order which allows an efficient energy expenditure (Aidelberg et al., 2014), avoiding the synthesis of unnecessary molecular machinery. In a medium with different available carbohydrates, the one at the top position of such hierarchy will block the activities related to all the other ones. This phenomenon is called inducer exclusion inhibition (IEI). Usually, glucose is the preferred carbon source because it is the one which induces a greater growth, except in very specific cases such as in nitrogen depletion conditions (Bren et al., 2016).

IEI does not repress directly the synthesis of enzymes and proteins related to the catabolism of alternative carbon sources. That is the function of catabolite repression, which is a negative

regulation mechanism dependent on the EIIA enzymes phosphorylation state. When these enzymes are dephosphorylated, there is no induction of the promoters dependent on the catabolite activator protein (CAP), so the expression of genes under their control remains repressed.

Assuming abundance of glucose in the medium, the PEP cellular concentration is high, the EIIBC^{gluc} complex is active and the phosphorylated EIIA^{gluc} (EIIA^{gluc} $\sim$P) concentration is low. When the glucose availability decreases, the EIIA^{gluc} $\sim$P concentration soars and induces the activity of adenylate cyclase (AC). This enzyme is adhered to the internal side of the plasmatic membrane. Its function is to synthesize cyclic adenosyne monophosphate (cAMP) from ATP. cAMP then binds to CAP and triggers a structural change in it.

The cAMP-CAP complex binds to a 22 bp consensus region in about 100 different promoters in the bacterial genome (Lawson et al., 2004). Once bound to DNA, cAMP-CAP works as a structural anchor for the RNA polymerase (RNApol) while interacting with the polymerase's αCTD subunit, thus providing a better stability during the transcription process.

The promoters regulated by catabolite repression are categorized in three different classes depending on the final arrangement of the cAMP-CAP-RNApol complex (Busby and Ebright, 1999). The features of the *tna* operon's promoter region (Ward and Yudkin, 1976; Deeley and Yanofsky, 1982) suggest this is a class I CAP-dependent promoter (Ebright, 1993; Zhou et al., 2014), which is the simplest class known. In summary, the presence of any carbon source related to IEI represses the activity of the *tna* operon promoter.

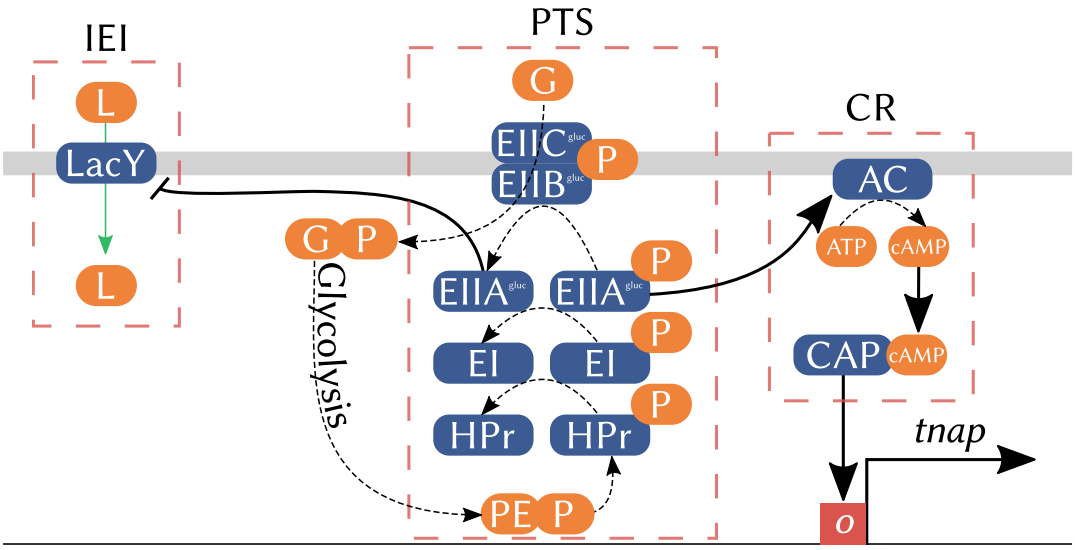


Figure 1.2: Catabolite repression pathways. The phosphotransferase system (PTS) is a regulatory pathway in which a phosphate group (P) is transferred—dashed spear-headed arrows denote transfer—from a phosphoenolpyruvate (PE) molecule to different enzymes, beginning with the Histidine carrier protein (HPr), which passes it down to the enzyme EI. The next enzyme in the pathway ($EIIA^{gluc}$) is specific to the process of the glucose (G) metabolism. $EIIA^{gluc}$ molecules finally transfer the phosphate group to a membrane bound complex comprised of the $EIIB^{gluc}$ and $EIIC^{gluc}$ enzymes. If glucose is present, the phosphate group is transfered from said complex to glucose itself—beginning the glycolysis process—and $EIIA^{gluc}$ remains phosphorylated, triggering the Inducer exclusion inhibition (IEI), preventing the cell from using another sugar such as lactose (L) as carbon source—the green spear-headed arrow stands for lactose transport into the cytoplasm. In the absence of glucose, $EIIA^{gluc}$ remains phosphorilated and activates the Adenylate cyclase protein (AC), which in turn forms cAMP from ATP—thus ending the Catabolite repression (CR). Finally, the Catabolite activator protein (CAP) binds to a cAMP molecule to induce the formation of the transcription complex at the promoter (*tnap*) operator site (*o*).

1.2 Rho mediated termination

The *tna* operon has a second regulatory mechanism other than catabolite repression: the Rho (ρ) mediated termination (Fig. 1.3). The former acts directly over the promoter sequence and the latter over the *tnaC* leader sequence (Stewart and Yanofsky, 1985; Konan and Yanofsky, 1997). This sequence is a 24–bp short sequence downstream the operon's promoter with a ρ-dependent termination site called *rut*. After transcription, ρ binds to the resulting mRNA's *rut* site (Konan

and Yanofsky, 2000).

Protein ρ is a homohexameric ring shaped helicase. Once bound to mRNA, it uses ATP to power its helicase function and move over the transcript toward the polymerase. Upon reaching the transcription complex, it forces the disengagement of mRNA from both DNA and RNApol. This is a complex mechanism involved in the regulation of several *E. coli*'s genes (Kaplan and O'Donnell, 2003). Recently, ρ has been found to be implied in intragenical regulation related to perturbations on the coupling of translation and transcription mechanisms (Boudvillain et al., 2013), although this does not appear to be the case in the *tna* operon.

Tryptophan inhibits the ρ-mediated termination. This is due to a single Trp codon in the *tnaC* sequence. When the ribosome reaches this codon in the mRNA transcript, the binding of the corresponding tRNA^{trp} molecule to the translation complex triggers a structural change related to the ribosome's exit tunnel and its peptydil transferase center. This in turn leads to a stalemate of the whole complex over the mRNA, effectively blocking ρ's binding (Cruz-Vera et al., 2007; Yang et al., 2009). In this way, tryptophan up-regulates the expression of both genes downstream of *tnaC*.

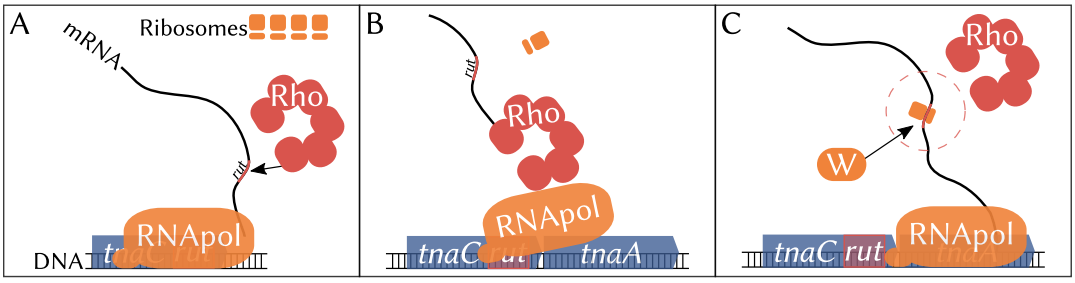


Figure 1.3: Rho mediated termination. Once the polymerase (RNApol) reaches and transcripts the *tnaC* sequence into mRNA (A), the enzyme Rho may bind to its *rut* site: If there's no tryptophan (W) present in the cytoplasm (B), *rut* will remain free of ribosomes, thus facilitating the binding of Rho, which in turn will unravel the mRNA from the transcription complex, leading to the disengagement of RNApol from the DNA. When there's tryptophan available (C), ribosomes stall over a Trp codon on the *rut* site because its corresponding tRNA^{trp} triggers a structural change on the ribosome that prevents its disengagement—thus allowing RNApol to continue the transcription of the sequence downstream of *tnaC*.

1.3 Tryptophanase dynamics

Indole is an important compound synthesized by different bacterial species. In *E. coli* it is involved in several functions such as genetic regulation, plasmid maintenance, quorum sensing, biofilm formation and acts as an interspecies signaling molecule (Mueller et al., 2009).

Early studies on indole biosynthesis evidenced a relation between indole production and the tryptophan metabolism. *E. coli* and some other microorganisms were able to survive in carbon depleted media supplemented with tryptophan while producing indole. Several mechanisms were proposed to explain this phenomenon and further studies proved the existence of a particular enzyme responsible of such reaction: tryptophanase (TnaA). This enzyme degrades tryptophan into indole, ammonia and pyruvate using pyridoxal 5'–phosphate as coenzyme (Wood et al., 1947).

Tryptophanase can also synthesize tryptophan from indole and serine, although this reaction is less likely to occur in the conditions required for the *tna* operon expression (Newton and Snell, 1964). Indole can also result from the tryptophan synthase enzyme function. However, tryptophan represses the activity of this enzyme, so it is safe to assume that most indole produced by *E. coli* is a by-product of TnaA operation (Snell, 1975).

When fully induced, TnaA can account for 10% of the total water soluble proteins in *E. coli* (Snell, 1975). Li and Young (2012) found that it localizes at the cell poles after expression, where it forms a well delimited and tight focus. This kind of structures are usually assumed to be inclusion bodies. These are commonly formed by over expressed or misfolded proteins and are often deposited at the cell poles by nucleoid inclusion.

TnaA foci were ruled out to be inclusion bodies through different tests involving isolation, purification and *in vivo* observation techniques (Li and Young, 2012). They could not be isolated or purified, since they disintegrated as soon as they were perturbed. They were not made up of misfolded proteins; this was verified through two genomic fusions, *tnaA-sfGFP* and *IbpA-mCherry* in *E. coli* MG1655. The first resulted in a green fluorescent and fully functional TnaA and the second in a red fluorescent IbpA protein which binds to misfolded proteins that are often clustered throughout the whole cell. Fluorescence microscopy analysis verified no overlap between both products.

Further studies on the TnaA foci behavior attempted to clear their role on the *tna* operon dynamics. Using the *E. coli* MG1655 GL69 strain, which harbors a *tnaA-sfGFP* genomic fusion, Li and Young (2013) performed several experiments comparing changes in growth, proportion of cells with a focus, and indole concentration in cell cultures over time, under proper conditions for *tna* operon induction.

They found that in the first 3 hours, during which the cultures are still in the exponential growth phase, 100% to 80% of the observed cells presented a focus and indole concentrations remained very low. After the 4^{th} hour, once cultures reached the stationary phase, the percentage of cells with foci dropped steeply to almost 0% and remained there for the rest of the experiment, while indole concentration soared and remained at high values.

This behavior suggests that during lag and exponential growth phases TnaA is inactive and clustered in foci. Once in the stationary phase, foci disperse and indole production begins, thus suggesting that tryptophanase activates at dispersion. The active form of this enzyme is a tetramer formed by four identical 52.8 kDa subunits. In each monomer a molecule of pyridoxal 5'phosphate covalently links to the Lys270 residue in their active site (Rety et al., 2015). This, along with the results of further experiments with the GL69-derived strains with directed mutations on specific codons of the *tnaA* gene (Li and Young, 2015), led to the conclusion that foci are mainly formed by inactive TnaA dimers and monomers. Said mutations targeted the protein's active site and showed that it is regulated by occlusion. A loop near the base of the catalytic pocket blocks tryptophan from accessing the site when the enzyme is not in a tetramer conformation.

Attempts to disclose more information on the foci regulation dynamics included experiments with *E. coli* GL69 strains modified with deletions or constitutive expression of the *toiC* and *cpdA* genes, both related to the cAMP-CAP regulation system discussed in previous sections. Results suggest the existence of an unknown cAMP-CAP independent post translational mechanism related to glucose. The addition of glucose to growth medium after the formation of foci resulted in disintegration of these and an apparent inactivation of the already synthesized tryptophanase, even in cases where catabolite repression was suppressed (Li and Young, 2014). It remains unclear if the presence of tryptophan or variations on its medium concentration affect tryptophanase

activation. Further work in this regard is required to disclose its possible role on the mechanism.

It is important to highlight the fact that all the experiments cited in this section were performed in batch method cultures. Such kind of cultures allow only to control the experiments' initial conditions. This happens because byproducts of cellular growth and depletion of medium nutrients change media composition along time. These considerations imply that the observed results do not account for constant reagent concentrations over the experiment time frame (Schulze and Lipe, 1964).

1.4 The TnaB permease

The operon's last element is the *tnaB* gene, which is located 90 bp downstream of *tnaA*. This gene codes for the TnaB permease: a low affinity tryptophan porter protein. Its high hydrophobicity, total length (415 aminoacids) and the position of several arginine and lysine clusters, suggest that it possesses 11 transmembrane domains. Its sequence holds a 52% similarity to Mtr, a high affinity tryptophan specific transporter (Sarsero et al., 1991).

Mtr, along with AroP are the only other known tryptophan permeases in *E. coli*, although the latter is not specific and also imports tyrosine and phenylalanine into the cytoplasm (Edwards and Yudkin, 1982; Sarsero et al., 1991; Gu et al., 2013). Tryptophan is a valuable and expensive aminoacid, so it is expected to find several mechanisms to provide the cell with it, either by transportation or biosynthesis (Yanofsky, 1960; Sarsero et al., 1991); hence the different transporters. Despite this, TnaB has been proven the only main tryptophan transporter present under the conditions needed for the operon expression, since the *mtr* gene is repressed by tryptophan itself (Edwards and Yudkin, 1982; Heatwole and Somerville, 1991) and studies on *E. coli* with repressed or deleted *aroP* gene have proven no incidence of its expression over the TnaB function (Gu et al., 2013; Li and Young, 2013).

2 Bistability

The function of the *tna* operon is to metabolize tryptophan as a carbon source alternative to glucose. The operon regulatory pathway features two main mechanisms: down-regulation by catabolite repression and up-regulation by exogenous tryptophan. The latter implies a positive feedback loop; this means that there is a self-induction loop tied to the activity of TnaB—the tryptophan specific permease encoded by the *tnaB* gene. This architecture is shared by many systems in control of the consumption of alternative carbon sources. A well known example in *E. coli* is the lactose (*lac*) operon; it has been studied for a long time and its mechanisms are now well understood (Jacob et al., 1960; Babloyantz and Sanglier, 1972; Wanner et al., 1978; Reznikoff, 1992; Yildirim and Mackey, 2003; Ozbudak et al., 2004; Zander et al., 2017). The proteins encoded within this operon are in charge of the transport and metabolism of extracellular lactose in the absence of glucose. In a similar fashion to the *tna* operon and tryptophan dynamics, extracellular lactose increases the *lac* operon expression levels by inhibiting a specific repressor for the *lac* promoter (Beckwith, 1987).

The positive feedback loop in the *lac* operon—driven by the activity of the β-galactosidase enzyme—gives rise to a phenomenon called '*all-or-none response*', wherein the expression of the operon genes remains relatively low until an inducer concentration threshold is reached (Novick and Weiner, 1957). Bistability has been proposed many times to explain this behavior (Yildirim and Mackey, 2003; Santillán et al., 2007); although experimentally it has only been proven under the effects of gratuitous inducers —compounds analogous to lactose that cannot be metabolized by β-galactosidase (Ozbudak et al., 2004; Zander et al., 2017). The fact that the *tna* and *lac* operons cover similar functions (and employ similar mechanisms to fulfill them) opens the possibility that the first one is bistable. To the best of our knowledge, this possibility has not been explored neither mathematically nor experimentally; doing so would add to the discussion on whether bistability can arise in natural systems with positive feedback loops. The features of bistability are explained in more detail in the following sections.

2.1 Mathematical origins of bistability in gene expression

Mathematical approaches to the study of genetic regulation systems offer the possibility to find patterns and behaviors underlying their dynamics. These systems tend to present two distinct traits: they change along time and are complex. Hence, differential equations systems intending to describe them are usually nonlinear. This means that most of them are impossible to solve analytically. However, there are still dynamic analysis techniques that offer qualitative and useful information on the system properties. Case in point, bifurcation diagrams and phase planes drawn from fixed points analysis (Strogatz, 1994).

Fixed points represent equilibrium solutions that are constant over time. For a differential equation like $\dot{x} = f(x)$, fixed points—denoted as x^*—are defined by the condition $\dot{x} = 0$. Fixed points can be stable or unstable. The former are those in which small disturbances away from them damp out in time, whereas in the latter those disturbances grow in time. A stable fixed point represents a steady state of the system that tends to be conserved along time—in an unstable fixed point the system presents the opposite behavior.

Assuming that equation 1.1 describes the regulation dynamics of a gene product denoted P, its accumulation rate ($\dot{P}$) is given by a synthesis term—an undefined function ($f(P)$)—and a degradation term (γP). In this case, fixed points (P^*) happen when the accumulation rate equals 0, thus $\dot{P} = 0$, therefore $f(P) = \gamma P$.

$$\dot{P} = f(P) - \gamma P. \tag{1.1}$$

A simple solution to find the fixed points of an equations system is to plot its terms separately over the same axes. Functions $f(P)$ and γP are plotted this way in Fig. 1.4 under the assumption that equation 1.1 describes the expression of a gene subject to positive feedback regulation. The solid curves are different forms of the synthesis function $f(P)$, corresponding to different maximal expression levels of P; while in this specific case $f(P)$ remains undefined, we can assume—for the sake of discussion—its parameters allow it to take the forms readily plotted in this figure. The dashed line corresponds to the degradation function γP. Fixed points P^* —represented by translucent dots—appear at the intersections of both functions, thus when $f(P) = \gamma P$. The blue and red curves only intersect once with the dashed line; in both cases, the resulting fixed points

would be stable for this kind of systems—those with positive feedback regulation loops (Alon, 2007). In the remaining instance, $f(P)$ intersects thrice with γP; the middle fixed point is unstable and the other two are stable. Hence a bistable system.

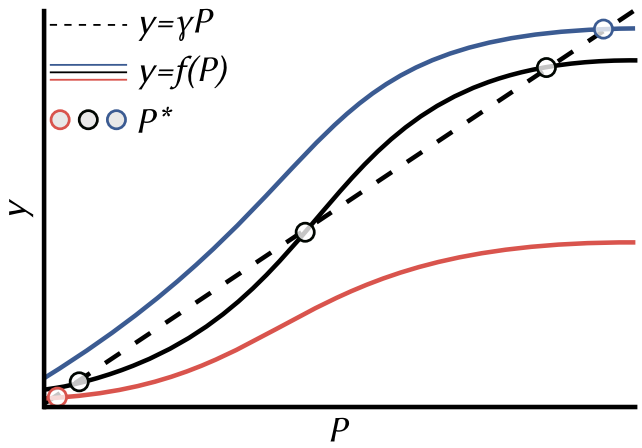


Figure 1.4: Fixed points in an expression system with positive feedback regulation. A graphical solution of equation 1.1. The dashed line is the system's degradation function in solved the form $y = \gamma P$. The solid curves are different instances of the synthesis function solved as $y = f(P)$. Fixed points P^*, denoted by translucent dots, happen at the intersections between degradation and synthesis functions.

In figure 1.4 we can assume $f(P)$ is an increasing function with a parameter denoted as h that determines the maximal expression levels of P—hence the different instances plotted. Therefore, the values of all existing fixed points P^* in the system can be plotted as a function of h—this is known as a bifurcation diagram.

The bifurcation diagram in Fig. 1.5 illustrates the behavior of the system's fixed points P^* in function of a parameter h. The solid blue lines denote stable fixed points and the dashed red line does so for unstable fixed points. The gray shaded box frames a so-called bistable region located at medium h values. In this area there are two blue solid lines and a red dashed one; consequently, for each point over the h axis there will be an unstable fixed point and two simultaneous stable ones; therefore the bistability region. Observe that at low and high h values, outside the bistable area, there are only single blue lines: only one stable fixed point appears per each h in these zones; hence they are deemed monostable regions.

The synthesis function $f(P)$ stands for the maximal expression levels of P, thus the curves plotted in Fig. 1.4 illustrate the expression dynamics of P. Every instance of $f(P)$ in Fig. 1.4 is a monotonically increasing function and, thereof, we can infer that the expression of P increases along with $f(P)$. As so, keeping this and the origin of fixed points in consideration, it is safe to assume that those fixed points denoted by the lower blue line plotted in Fig. 1.5 stand for low expression steady states of the system. On the other side, the higher blue line in the bifurcation diagram denotes high expression steady states. Therein, the monostable regions correspond to low and high expression regions.

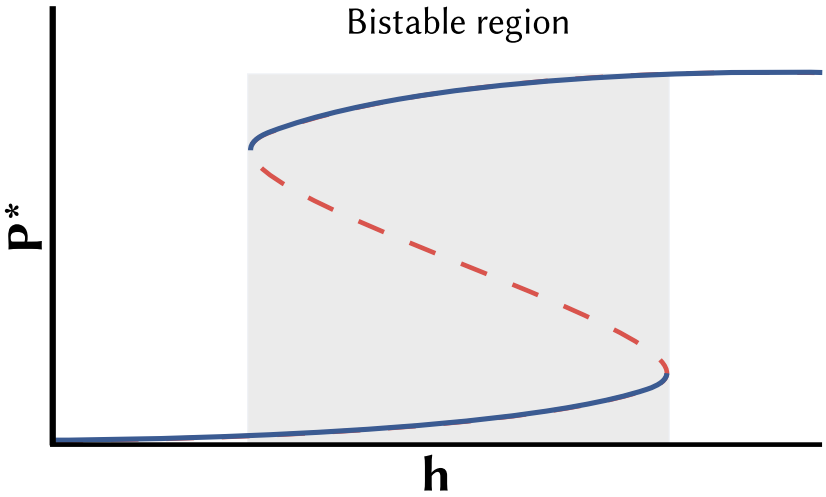


Figure 1.5: Bifurcation diagram of a bistable system. Fixed points P^* vs parameter h. Solid blue lines represent stable fixed points; dashed red line stand for unstable fixed points. Two monostable regions form at low and high h values. The shaded area is a bistable region in which two stable and one unstable fixed points appear for each h.

To infer the system behavior, suppose its current state is a dot over the lines plotted in Fig. 1.5; also suppose this dot to be located close to the graph origin—recalling the discussion in previous paragraphs, this would imply the system is at a low expression state. If the dot starts to travel along the h axis toward higher values, it would remain at low expression upon reaching the right end of the bistable region; then, there would be a sudden switch and the system would immediately enter to a high expression state. The same applies backwards, if the system state goes from high to low h it will remain on the high stable fixed points and switch suddenly to the low state at the bistable

region's left limit. This is known as a bistable switch and has relevant biological implications. This phenomenon is known as hysteresis and it is an important bistability feature. It is defined as the dependence of the state of a system on its history, in other words, the system's refusal to change its state (Strogatz, 1994).

2.2 Biological and experimental implications of bistability

Bistability implies that a system will not switch states unless its parameter values reach a defined threshold. This is called an '*all-or-none response*' and originates directly from hysteresis. Systems presenting this behavior have been studied for a long time. At first it was attributed mostly to enzyme dynamics (Novick and Weiner, 1957; Monod, 1949), later genetic regulation was deemed to play a central role (Jacob and Monod, 1961).

The mechanisms from which bistability arises originate in autoregulation—regulation of a gene by its own gene product. It can be negative, when it turns in self-repression, or positive, in the case of self-induction. Positive autoregulation systems are bistable when they present hypersensitivity. This results in a very slow response to stimuli, albeit extremely stable once it is established. When a bistable switch is reached the system's new state will remain for a long time, often spanning several cell generations. As a result, bistable systems are classified as developmental transcription networks where cells need to make irreversible changes in slow timescales (Alon, 2007; Tiwari et al., 2011).

Two well known genetic regulation systems with positive feedback loops and switch behaviors are the previously mentioned *lac* operon in *E. coli*, and the *GAL* regulon, which controls the metabolism and consumption of galactose in *Saccharomyces cerevisiae* (Apostu and Mackey, 2012); both have been proposed to be bistable based on their regulation pathways and experimental observations. A quick analysis of the schematics in figure 1.1 reveals two different self-regulation loops embedded in the *tna* operon dynamics. The first one is related to the *tnaA* gene. Tryptophanase degrades tryptophan, a system inducer, thus creating a negative autoregulation loop. The second one arises from *tnaB*; the expressed permease imports tryptophan into the cytoplasm, which represses ρ by itself, resulting in self induction. This operon is indeed a candidate to bistability.

The behavior outlined in the previous section, specifically regarding figure 1.5, results from a deterministic equations system. However, biological systems are not deterministic, they are in fact stochastic, noise prone and chaotic (Alon, 2007). Experimentally, several considerations arise from this situation. First, cell populations are heterogeneous; they do not express genes in an uniform way even if all the cells are genetically identical. Second, phenotypic heterogeneity implies different subpopulations with different genetic expression levels at different experimental conditions. Third, since the average filters out expression differences, the best approach is to measure each single cell's expression level (Dubnau and Losick, 2006; Magdanova and Golyasnaya, 2013). Experiments performed under the conditions stated previously would render thousands of data points, one per cell. Thus, the resulting distributions would display the system of interest's state at the exact moment of measurement.

The work by Gardner et al. (2000) is a clear example of the conditions listed above. Briefly, Gardner et al. designed a synthetic genetic circuit based on a simple mathematical model of a bistable genetic expression system. Although the origins of the bistable behavior in this case were basically different to those of the natural examples previously discussed, their results remain useful to illustrate the experimental side of bistability. The molecular pieces of the circuit were carefully chosen to fulfill the model parameter requirements for bistability. The circuit was constructed, inserted in an expression plasmid and introduced into *E. coli* MG1655. They included the *gfpmut3* gene as a reporter; thus they used the fluorescence intensity of expressed green fluorescent protein (GFP) as a measure of the circuit expression level. One of the experiments they performed to prove bistability was to induce the toggle switch behavior at different induction levels and record the fluorescence intensity of each individual cell within the induced culture.

Results of this experiment are summarized in Fig. 1.6. The left panel of each plot shows scatter plots of GFP fluorescence intensity vs cell count. Right panels show the resulting fluorescence distribution within the population. First, they induced a culture with an inducer close to the switch-triggering concentration (A), note that most cells stay within a low fluorescence level and the resulting distribution is unimodal. Then, they used the inducer threshold concentration (B); in this case, there are two distinct subpopulations, one at low expression level and another at high expression level—here, the population distribution turns bimodal. The last plot (C) shows the

results of induction at a higher concentration than the previous ones: once again, almost all the cells are clustered at a single population but in high expression level and the resulting distribution is unimodal. Bimodality evidences bistability (Shu et al., 2011). This is because the stochastic nature of cells pushes them towards one of the stable fixed points inside the bistable region. In a deterministic system, all cells in a population would remain in the low expression state until the switch threshold is surpassed.

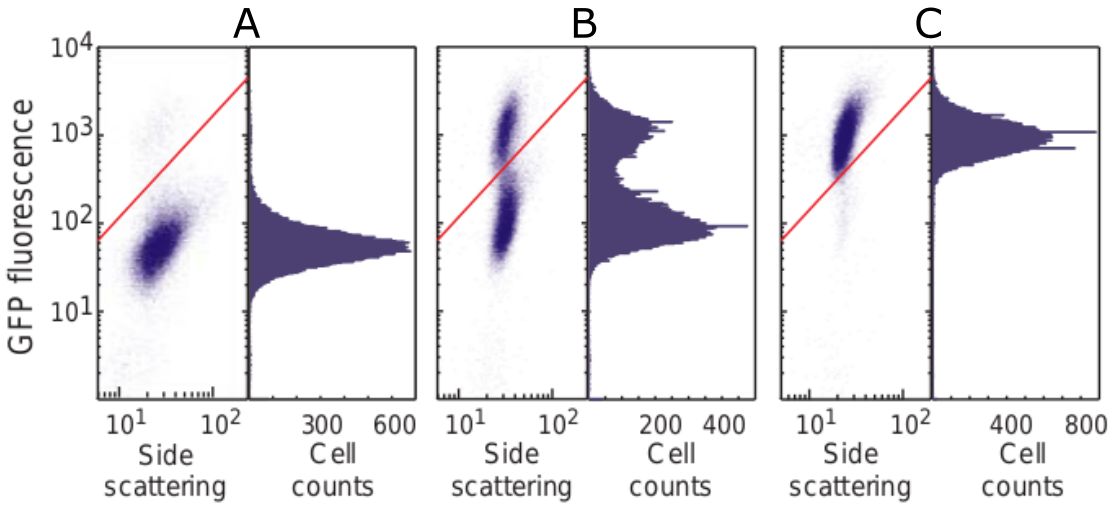


Figure 1.6: Experimental data of a bistable switch. A synthetic bistable switch was tested at different inducer concentrations. These graphs show the reporter gene expression—measured from fluorescence intensity—around threshold inducer concentrations. Left panels show scatter plots of fluorescence intensity vs cell count. Right panels are the resulting fluorescence distributions. A) The system remains at a low expression level right before the induction of the bistable switch; all the measured cells appear to form a single almost homogeneous population. B) Two distinct subpopulations appear, one at high expression and another at low expression level; this is the bistable region, as evidenced by the resulting bimodal distribution. C) At an inducer concentration outside the bistable area the system reaches a high expression level and, once again, there seems to be only one homogeneous population. Modified from Gardner et al. (2000).

It is important to note the experiments by Gardner et al. were planned according to their mathematical model of the toggle switch. As such, they were able to accurately drive their system to its bistable region. Mathematical models allow to explore known systems beyond experimental limits. The predictions obtained from them are often extremely accurate and widen the general

comprehension of natural processes (Bialek, 2018). Ideally, these models are built from available experimental data and their predictions proven within the limits of current research. Sometimes this process is not straightforward, the information already available may not suffice to build an accurate model—this appears to be the case with the *tna* operon.

A strategy to deal with the lack of experimental data would be to explore a system behavior under several environmental conditions to find the limits of an hypothetical bistable region. There have been attempts to apply this idea; the works by Ozbudak et al. (2004) and Zander et al. (2017) are solid examples. Both tried to demonstrate bistability in the *lac* operon and the latter was able to convincingly prove it to arise only in the presence of gratuitous inducers.

Recalling Fig. 1.5, the bistable region is well defined within a bifurcation diagram. As previously discussed, a bifurcation diagram results from plotting the fixed points of a system in function of its parameters. This implies that the experimental strategy proposed to find bistability in a genetic expression system aims to retrieve data on the system's fixed points. Steady states are biological analogues to fixed points. These are achieved in a bacterial culture when the response to constant medium conditions remains equally constant along time (Schulze and Lipe, 1964). Consequently, the system of interest must be driven to a steady state under several different medium conditions corresponding to different parameter values.

From the technical side of the experiments, it is easier and more precise to control tightly medium concentrations of the system inducers. In the case of the *tna* operon, those parameters are the glucose and tryptophan medium concentrations. Both because their concentrations are controllable and play important roles in the operon's regulation pathway; glucose because it is the system's main repressor and it plays an important yet undetermined function in tryptophanse activation; and tryptophan because it is the operon's main inducer and the possibility that it may interfere in or induce the TnaA activation has not been ruled out.

2 Project definition

The architecture and expression dynamics of *E. coli*'s *tna* operon suggest that it is bistable. However, current available data is insufficient to support bistability. In order to do so, it is necessary to explore the operon's behavior in several stationary states to draft a bifurcation diagram. In this way, we would be able to not only find if there is bistability, but the range of conditions in which the system assumes this behavior. Also, recent research suggests the presence of complex and fairly unexplored regulation mechanisms related to the presence of glucose; the steady states analysis would contribute to a better comprehension on their dynamics and to find out if tryptophan is also related to them. To achieve this, we propose the following hypothesis and objectives.

1 Hypothesis

The *Escherichia coli*'s *tna* operon is bistable and the concentrations of both tryptophan and glucose in its medium regulate tryptophanase activation.

2 Objectives

The main objective of this project is:

To study the dynamic behavior of *E. coli*'s *tna* operon in response to changes on the concentrations of tryptophan and glucose in its medium.

To achieve it, we proposed three specific objectives:

1. To design, test and implement a steady state bacterial culture protocol.

2. To analyze the *tna* operon's behavior during induced steady states in several medium concentrations of tryptophan and glucose.

3. To analyze how different medium concentrations of tryptophan and glucose affect tryptophanase activation.

3 Methodology

1 Experimental design

The necessity to drive the *tna* operon to a fixed point was stated in previous sections. A fixed point implies no changes in a system over time: a steady state. The objectives of this project require to retrieve data at several fixed points. Mathematically, a way to change the state of a system from one fixed point to another is to manipulate the system's parameters values while keeping constant the remaining variables. However, biological systems are inherently noisy and this poses an important challenge to achieve a steady state under experimental conditions.

In microbiology, '*steady state*' means to reach and keep a point in a culture in which the cellular mass and medium conditions remain constant. This cannot be achieved through the traditional method of batch culture, in which conditions change slowly over time due to by-products of bacterial growth and nutrient depletion along time. Consequently, it is unfit to explore the relation between microorganisms' metabolic states and its environmental conditions due to an unavoidable lack of control over medium changes (Schulze and Lipe, 1964).

Achieving a steady state is not a new idea, it has been researched previously and one of the solutions proposed then has become an established standard ever since: the continuous flow culture. In this method, the culture is kept inside a reaction chamber fitted with inlets for fresh medium, outlets for waste disposal and sampling, and different sensors and actuators that control variables such as temperature, oxygen levels and optical density. This system is called bioreactor

(Schulze and Lipe, 1964). Bioreactors are complicated devices usually designed to operate with large culture volumes, from 500 ml up to several liters, so every experiment is costly and it is impractical to perform several experiments with different sets of conditions. An obvious solution to this problem is to scale-down a bioreactor enough to work with just milliliters and even microliters of medium (Weibel et al., 2007).

Scaling down a bioreactor to micron dimensions can be achieved through microfluidics techniques. In this regard, soft lithography allows the design and assembly of very small devices with high precision. Furthermore, there is a wide range of materials with different properties which can cater to different culture requirements (Weibel et al., 2007; Sackmann et al., 2014).

Polydimethylsiloxane (PDMS) is the material of choice in this project to build a micro-scale bioreactor. Its features include gas permeability, elasticity, optical transparency, flexible surface chemistry, low permeability to water, and low electrical conductivity (Ng et al., 2002). The main disadvantage it presents is protein fouling; which is easily avoidable with a biocompatible anti-fouling coating such as polyvinylpyrrolidone (PVP−40) (Zhang and Chiao, 2015), that also inhibits the formation of biofilms on the device inner surfaces. Thus, we designed a PDMS microreactor capable of keeping a healthy cell culture with constant conditions, simple enough to be easily fabricated, and small enough to require just microliters of medium per experiment.

The microreactor was designed to trap cells in a growth chamber that allows culture replenishment with fresh medium and avoids overpopulation; while keeping growth conditions constant and stress free for bacteria. We based our design on the work of Danino et al. (2010). Their arrangement consists of a main nutrient delivery channel that feeds a rectangular trapping chamber on the side of the channel itself. Once seeded, a monolayer of bacterial cells grows in the chamber and excess cells are eventually pushed back into the channel where they then flow to the waste port. This device allowed for a constant supply of nutrients and inducers while discarding depleted medium and excess cells. In our case, we opted for bigger dimensions and avoided the formation of biofilms.

Figure 3.1 shows the microreactor design. It has two inlets (a, b), a channel connecting both to a growth chamber (c), and an outlet (d) connected to the chamber through another channel. The

device's inner height is 50 µm. Medium flows from inlet (a) to outlet (d). The flow velocity in the entrance channel is high enough to prevent bacteria swimming against the flow and reaching the medium reservoirs. Due to the device design, the velocity decreases to a tenth of its initial value inside the growth chamber, letting bacteria population grow and take advantage of the medium nutrients. Finally, the flow velocity increases again towards the exit channel, taking depleted medium and excess cells—which can be retrieved for its subsequent observation.

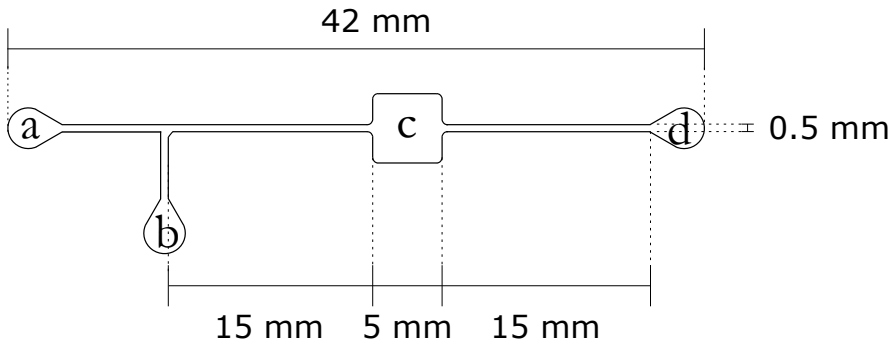


Figure 3.1: Microreactor design. Upper view of the microreactor. Fresh medium kept in a syringe is fed into the device from inlet (a) and flows to the growth chamber (c) at 0.3 $^{\mu l}/_{min}$. A solution of *E. coli* cells/PBS is introduced into the device at inlet (b), the flow from the medium inlet carries them to the growth chamber, were the flow velocity is low enough to account for a normal growth. The dimensions of the growth chamber (c) allow for a sudden decrease on flow speed to allow growth while medium is replenished. Speed increases again close to the exit channel, thus discarding excess cells and depleted medium. Discarded culture exits the device from outlet (d), which also allows to take samples at any time. The microreactor inner height is 50 µm.

The microreactor was thoroughly tested to verify the cultures viability and the possibility of them reaching steady states inside it. This was done with the *E. coli* GL69 strain (Li and Young, 2012). This strain is a derivative from *E. coli* MG1655 with an *sfGFP* gene fused to the *tnaA* gene within the genome, which results in tryptophanase monomers fused to super-folder GFP, thus the resulting fluorescence can be used as a direct measure of the *tnaA* gene expression.

After verifying the attainment of steady states inside the microreactor and defining the glucose and tryptophan concentrations to be tested, we performed several experiments in steady state. From the resulting samples, we analyzed the sfGFP fluorescence intensity at single cell level via image analysis. Finally, we drawn conclusions on the dynamic behavior of the *tna* operon

by contrasting the resulting probability distribution functions (PDF) of single cell fluorescence intensities corresponding to each bacterial culture.

In the following section, we describe in detail the experimental materials and methods we employed.

2 Methods

2.1 Strains and growth conditions

The *E. coli* GL69 strain (Li and Young, 2012) was kindly provided by Kevin Young's lab at University of Arkansas for Medical Sciences. Stocks were prepared in 15% glycerol and kept at -80°C. Working stocks were kept in solid LB-Agar plates at 4°C. Experiments were performed in M9 minimal medium, supplemented as stated by Li and Young (2015), with 18 mM sodium pyruvate instead of 1% Bacto Casaminoacids. All media and stock solutions were filter sterilized.

Growth curves were performed for 12 h in 96 wells culture microplates at 37°C. Optical density (OD_{600}) was measured every 15 min. Tested glucose concentrations were $\{500, 250, 125, 62.5, 31.25, 15.62, 7.81, 3.9, 1.95, 0.97$ and $0.48\}$ mM. Tryptophan concentrations were $\{48.97, 24.48, 12.24, 6.12, 3.06, 1.53, 0.76, 0.38, 0.19, 0.09, 0.04$ and $0.02\}$ mM. Maximum concentrations in both cases were chosen to ensure medium saturation. Each experiment was performed in triplicate.

2.2 Microreactor fabrication

The reactor design was plotted in AutoCAD (AutoDesk ®) at a 1:1 scale. This design was used later to create a master mold to manufacture the devices through soft lithography.

A master mold was created over a Si wafer coated with SU-8 3050 (Microchem®), an epoxy based negative photoresist. It was prepared as per manufacturer's instructions to achieve a 50 µm height. The device design was etched via focused laser writing at 70 mW × 100%.

The devices were prepared pouring degassed PDMS, mixed with curing agent in a 10:1 proportion, over the master and left to polymerize at 80° C for 2 hours. After polymerization, devices were cut off from the resulting PDMS mold and cored at the inlet and outlet locations with a 1.5 mm Uni-Core® tool. They were washed twice in MiliQ water for 30 minutes, then 15 minutes in 97% isopropanol and once again 15 minutes in MiliQ water. All washes were performed in an ultrasonic bath sonicator. Finally, the devices were left to dry in a desiccator for 48 hours.

Dry reactors were assembled over microscope slides previously cleaned with Micro-90 cleaning solution (Cole-Parmer®). Both PDMS molds and slides' surfaces were treated with oxygen plasma generated by a BD-20AC Laboratory Corona Treater (Electro-Technic®) by 30 to 60 seconds and, immediately after assembly, heated up to 2 hours at 80° C.

Assembled devices were perfused with 97% isopropanol, dried at 95° C, perfused with 0.04% PVP–40, and dried again at 95° C. Once dry, devices were sterilized by autoclaving.

2.3 Continuous flow culture setup

Bacteria were grown overnight (ON) at 37°C in 5 ml of M9 with sodium pyruvate, then harvested by centrifugation at 1 G per 5 min, washed and centrifuged twice in 10 mM phosphate-buffered saline (PBS). After the last centrifugation, supernatant was discarded and cells resuspended in 10 mM PBS. This cell suspension was later used to feed the microreactor.

All the media used to perform the continuous flow experiments were based in the previously described M9 medium, further supplemented with combinations of different concentrations of glucose ($\{0, 3.7, 7.5, 15, \text{and } 30\}$ mM), and tryptophan ($\{0, 6, 12, 24, \text{and } 48\}$ mM). To avoid the formation of cell clumps inside the microreactor, all media were added Tween20 to a final concentration of 0.02%.

Sterile microreactors were filled with medium using a KDS-210 syringe pump (kdScientific®) through outlet c at 150 $^{\mu l}/_{min}$. Then, they were inoculated with 60 µl of PBS/cells suspension at 120 $^{\mu l}/_{min}$ through inlet b (see Fig. 3.1) while occluding the inlet a to avoid

syringe contamination. After inoculation, medium flowed from inlet a at 0.3 $\mu l/min$. After 24 h, 5 µl samples were taken from the outlets and loaded on 1% agarose coated slides for imaging.

2.4 Data acquisition

2.4.1 Microscopy

Bright field and fluorescence 10 sec videos from at least 10 different regions were obtained from each sample with a CCD `WAT-902H2 Supreme` camera (`Watec®`) mounted on an `Olympus®` BX51 microscope with an 100× oil objective (1.3 NA PH3) and illuminated by the `Olympus®` fluorescence source `BH2-RFL-T3`. sfGFP was visualized through an `Olympus®` `U-MNIB2` filters set (480/20 excitation / 510 LP emission). Videos were recorded at 30 fps with a 480 × 720 pixel resolution.

2.4.2 Image analysis

200 frames of each video, both bright field and fluorescence, were averaged in order to improve sharpness, reduce background fluorescence and noise, and acquire a representative single image of each recorded region. To quantify fluorescence intensity within each cell, images were analyzed in `MatLab R2017b®`(The MathWorks, Inc.). In the next paragraphs, we include a brief description on this software. Resort to section A.B to consult the source code.

First, we created mask images from bacteria regions in bright field images. To determine if a region belonged to a cell body, we performed a segregation operation analyzing mean contrast and intensity. If contrast was good enough we applied a threshold at half intensity to a simple binary operation (`imbinarize` in `MatLab`) to isolate the areas of interest. In the opposite case, we first summed the bright field image to its corresponding fluorescence image to improve contrast. Then, the binary segmentation function was performed with an adaptive sensitivity of 0.35. This process rendered images with empty areas corresponding to bacteria bodies, hence mask images.

The resulting masks were subtracted from fluorescence images to keep only the pixels corresponding to bacteria. The pixels within each isolated region were quantified to obtain cell

areas. The average area value was 190 pixels. To avoid taking debris and extraneous objects as bacteria, we applied area thresholds, a lower one of 114 pixels and a higher one of 7600 pixels. The former was applied to account for smaller than average cells; the latter was applied to include areas with up to 40 clustered cells.

Once the bacteria areas were isolated in the fluorescence images, we performed a fluorescence intensity analysis to each region. First, the intensity of each pixel was quantified from the absolute difference between the masked and the original fluorescence images. Intensity values between 5 and 90 accounted for spread fluorescence. The lower threshold value, 5, was established to disregard background intensity. Values above 90 were considered as corresponding to foci fluorescence. From these data, we computed the spread and foci fluorescence densities in each bacterium by adding the intensities of the corresponding pixels and dividing by the cell area.

4 Results and discussion

The proteins encoded in the *tna* operon of *E. coli* are in charge of importing and metabolizing exogenous tryptophan as a carbon source in the absence of glucose and any other sugar capable of triggering catabolite repression. Their dynamics comprise two self regulation loops: a self-repression loop related to tryptophanase activity, which is encoded by the *tnaA* gene and catabolizes tryptophan into three by-products of biological importance; and a positive feedback loop supported by the importation of tryptophan into the cytoplasm by TnaB—tryptophan blocks the activity of the ρ-protein over the operon's nascent mRNA transcript, thus preventing its early termination.

This architecture, most specifically the positive feedback loop, is shared by many known systems related to the hierarchical consumption of carbon sources alternative to glucose. It has been theorized that it gives rise to a switch behavior and bistability has been proposed to explain its origin. However, bistability has never been proven convincingly in an experimental setup. In the case of the *tna* operon, there is little information available to prove bistability beyond its regulation pathway.

The biochemical dynamics of the *tna* operon were explored in a series of papers by Li and Young (2012, 2013, 2014, 2015). They built the *E. coli* GL69 strain, which harbors a *tnaA-sfGFP* genomic fusion, resulting in a functional tryptophanase with a sfGFP molecule attached that allowed to track its position within the cell and measure its expression level through fluorescence intensity. They found out that TnaA monomers formed tight foci, usually located at the cell poles and one per cell. When most cells in the culture presented a focus, little to no indole was detected

in the medium—this happened during the lag and exponential growth phases. After the cultures reached the stationary growth phase, the foci dissolved, TnaA formed active tetramers and indole production began. This suggests the existence of a post-translational regulation mechanism. However, these experiments were performed with batch method cultures, and no control over changes in the media was possible.

We used *E. coli* GL69 to perform several experiments driving the operon activity to a steady state, which implied no changes in operon's expression state due to constant medium conditions over long periods of time—from 24 h to 48 h. To achieve this we designed and developed a simple micro-scale bioreactor able to operate with microliters of medium.

The microreactor design is depicted in Figure 3.1 (ch. 3, sec. 1). It keeps culture inside a growth chamber were fresh medium is replenished at low flow velocity to avoid stressing bacteria, while, at the same time, depleted medium and excess bacteria are washed away. Medium is delivered from a feeding channel at a flow velocity ten times higher than that inside the growth chamber, thus keeping bacteria from reaching the medium reservoir. Samples are collected from an outlet linked to the chamber by another high flow velocity channel. The samples in these experiments were later analyzed via microscopy and image analysis.

1 Growth media

We needed to define the widest possible range of tryptophan and glucose concentrations with which to perform experiments to gather data on the behavior of the *tna* operon in steady state. As so, we needed to find the highest concentrations of both reagents that were soluble in water which do not compromise bacterial growth. We measured growth curves of *E. coli* GL69—these require to perform batch cultures while measuring cell growth at specific time intervals; in this project measurements were performed each 15 min—in M9 media supplemented with different concentrations of glucose $\{500, 250, 125, 62.5, 31.25, 15.62, 7.81, 3.9, 1.95, 0.97$ and $0.48\}$ mM and tryptophan $\{48.97, 24.48, 12.24, 6.12, 3.06, 1.53, 0.76, 0.38, 0.19, 0.09, 0.04$ and $0.02\}$ mM. In both cases, maximum concentrations were chosen to ensure medium saturation. Each experiment was performed in triplicate.

The resulting data points from the growth curves ($^{OD_{600}}/_{time}$) were fitted to a logistic growth curve. Logistic growth occurs when a population's growth rate decreases as population size approaches a maximum imposed by available resources. It can be described by this equation:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = rN\left(1 - \frac{N}{K}\right),$$

Where N is the population size and its rate of change is determined by the rate of growth r and the medium carrying capacity K. Data were fitted to its solution below, where N_0 denotes the initial value of N in the experiment:

$$N(t) = \frac{K}{1 + \left(\frac{K}{N_0} - 1\right)\mathrm{e}^{-rt}}.$$

Given that N stands for the population size, it can be interchangeable with OD_{600}, since that was the variable measured to track population growth over time. Figure 4.1 displays an example of the fitting operation described above. The white dots stand for the data gathered from three measured growth curves in M9 medium with 30 mM glucose; each of them were fitted to a logistic curve and their resulting growth rates (r) and medium carrying capacities (K) were averaged. The red curve is the resulting logistic growth curve from the average r and K values.

Data from growth curves performed in glucose are summarized in figure 4.2. Red triangles represent mean r values and blue dots correspond to mean K values, error bars denote standard deviation. Note that there is little change in the growth rates from 3.9 up to 250 mM glucose. However, $\bar{K}$ values peak at 31.2 mM and begin a slow descent toward higher glucose concentrations. This means the medium carrying capacity is decreasing beyond this point and most probably medium conditions are becoming toxic for bacterial growth.

Previous works on *E. coli*'s growth in glucose have concluded that concentrations beyond 50 mM have little effect on cellular mass in continuous flow cultures (Shehata and Marr, 1971)—which are the kind of cultures needed to observe steady states. Noting this and the behavior depicted in Fig. 4.2, and to avoid the possibility of an altered metabolism and physiology due to to osmotic stress (Lange et al., 1993), we chose to use 30 mM glucose as maximum concentration, as such,

{3.7, 7.5, 15, and 30} mM glucose were the concentrations used in the upcoming experiments.

Similar to glucose, tryptophan growth curves data are illustrated in figure 4.3. In contrast to the previous plot, growth rate (blue dots) achieves a maximum value at the lowest concentration, 0.02 mM. Then, after a steep fall and what appears to be a plateau, $\bar{r}$ begins a slow increasing trend after 6.12 mM. On the other hand, $\bar{K}$ (red triangles) is a monotonic increasing function of tryptophan concentration. Error bars also denote standard deviation.

The above described behavior can be due to the metabolic cost of the *tna* operon activation. If there is little to no tryptophan available as carbon source, it may not be cost effective to metabolize it. The fact that $\bar{K}$ values increase along with the concentration hints to better growth conditions and low toxicity in a tryptophan saturated media. Given these conclusions, we chose 48 mM as maximum tryptophan concentration, since it is close to the maximum tryptophan solubility in water at room temperature. The final tryptophan concentrations used in the steady state experiments were {6, 12, 24, and 48} mM.

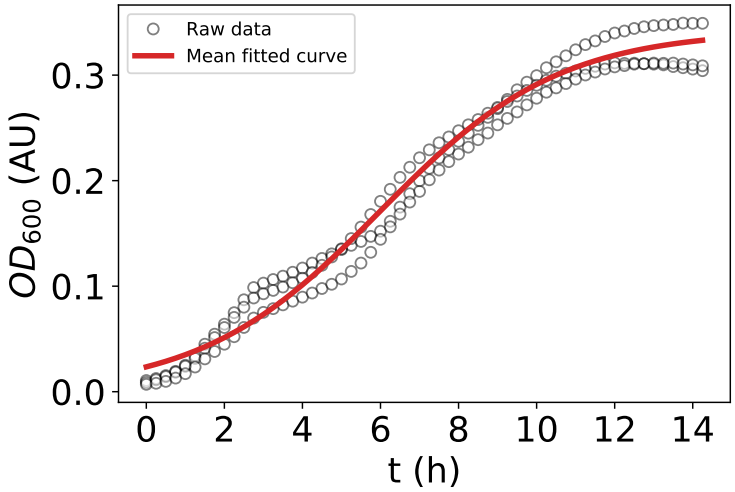


Figure 4.1: Mean logistic growth curve. White dots denote data points from three growth curves performed in M9 medium with 30 mM glucose. After fitting data from the three experiments to a logistic growth curve, their resulting r and K values were averaged and used to calculate the red curve.

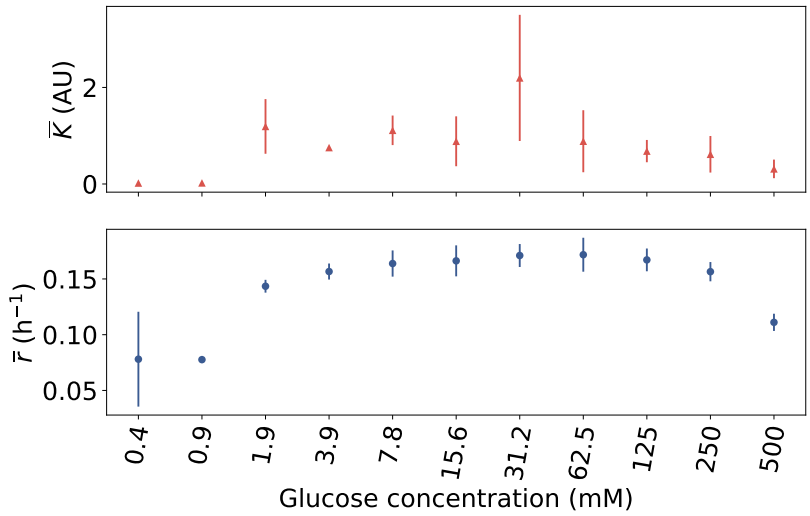


Figure 4.2: Mean logistic growth fit parameter values for different concentrations of glucose. Red triangles stand for mean medium carrying capacity K and blue dots for mean growth rate r. Each experiment was performed in triplicate. Error bars indicate standard deviation.

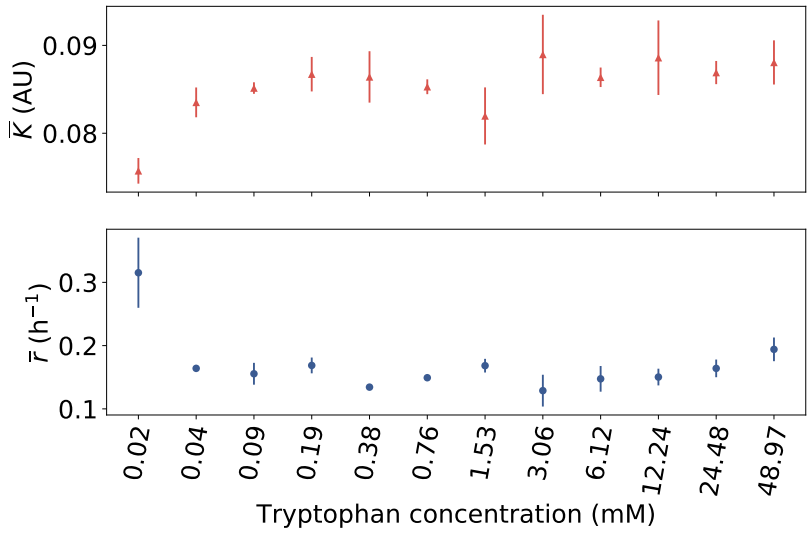


Figure 4.3: Mean logistic growth fitting parameter values for different concentrations of tryptophan. Red triangles stand for mean medium carrying capacity K and blue dots for mean growth rate r. Each experiment was performed in triplicate. Error bars indicate standard deviation.

We needed to test the *tna* operon in the absence of glucose and tryptophan. This implied to use minimal M9 medium supplemented with neither reagent. However, M9 formulation requires a carbon source to be added for growth to occur. Therefore, we required an alternative carbon source that would not trigger catabolite repression and we used sodium pyruvate.

E. coli can use pyruvate as a sole carbon source since it is the final product of glycolysis and it plays a central role in both aerobic and anaerobic growth. Furthermore, there are at least two different pyruvate uptake systems in *E. coli*, so the species is well prepared to survive on it as a carbon source (Kreth et al., 2013).

We tested sodium pyruvate by itself as an alternative carbon source to study the operon's behavior in absence of tryptophan without triggering cAMP-CAP dependent catabolite repression. To allow for a proper comparison with media with either glucose or tryptophan, we required to add it to all experimental media. Taking this into consideration, we also tested the possible interference of pyruvate as supplement in media with added glucose or tryptophan.

We measured *E. coli* GL69 growth in M9 media supplemented with 18 mM pyruvate, 18 mM pyruvate + 30 mM glucose, and 18 mM pyruvate + 48 mM tryptophan. All media were tested in triplicate. The retrieved data points were fitted to the previously described logistic growth curve. The resulting fit parameter values were averaged and compared.

These results are shown in Fig. 4.4. Blue bars stand for mean growth rates ($\bar{r}$) and red bars for mean medium capacities ($\bar{K}$). The error bars denote standard deviations. M9 media with either pyruvate + glucose (Pyr+Gluc) or pyruvate + tryptophan (Pyr+Trp) present substantially greater parameter values than those of M9 medium with just pyruvate (Pyr). In summary, *E. coli* can grow with just pyruvate as a carbon source; also, the addition of pyruvate to media with either glucose or tryptophan doesn't appear to have a significant effect in neither growth rate nor medium carrying capacity. All media were added 18 mM pyruvate after these experiments, even when not explicitly stated.

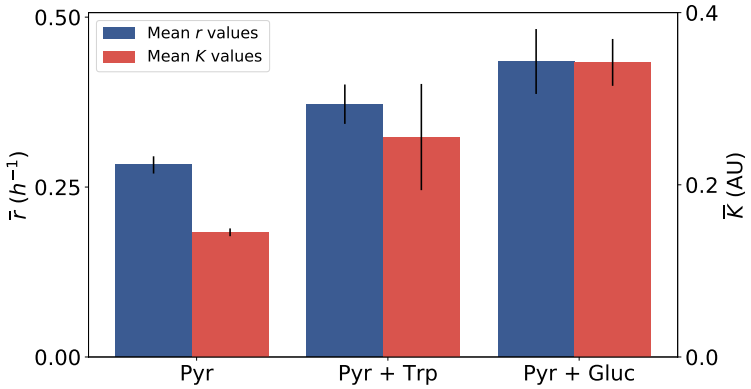


Figure 4.4: Mean logistic growth fit parameter values for different media with added pyruvate. Growth parameters comparison between M9 media supplemented with pyruvate (Pyr), pyruvate + tryptophan (Pyr+Trp), and pyruvate + glucose (Pyr+Gluc). Blue bars stand mean growht rate values, red bars stand for mean medium carrying capacity values. Error bars denote standard deviation.

2 Continuous flow culture protocol

We proposed the microreactor design depicted in Fig. 3.1 based on the work of Danino et al. (2010). They used a series of simple rectangular traps to keep monolayers of bacterial cells constantly supplied with fresh medium through an adjacent channel that also helped to discard excess bacteria. Their main objective was to study the synchronization of bacterial populations and their design was aimed to stimulate quorum sensing; to achieve this, they tested different trap sizes and geometries. They found that 100 × 100 µm was the best size to achieve a good enough cell density and delivery of inducer reagents for quorum sensing. They forced the formation of monolayers with a 1.65 µm trap height. In another experiment, they also tested a 100 × 2000 × trap with a 0.95 µm height surrounded by a larger chamber 10 µm tall; nutrients and inducers were also successfully delivered in this case.

We aimed to isolate the activity of the *tna* operon. This required to avoid triggering mechanisms like quorum sensing—the opposite to the experiments discussed above. Taking into account the design principles of their devices regarding the mechanisms of cell entrapment and media delivery, we proposed our device's dimensions intending to let bacteria swim freely inside the trap while avoiding to reach population densities similar to those of their experiments.

Our device is intended to be simple to construct and use. This is why both the channels and the growth chamber have an even height. In the same vein, the experimental protocol (see section 2.3) requires to introduce an already homogenized medium to a specific inlet to avoid complicated reagent mixing steps inside the microreactor.

Media were introduced to the microreactors at 0.3 $\mu l/min$, this way, the flow velocity in the delivery and exit channels was 200 $\mu m/s$ and 20 $\mu m/s$ inside the growth chamber. Darnton et al. (2007) reported a swimming speed of 25 ± 8 $\mu m/s$ for *E. coli*. Given these facts, bacteria were able to swim inside the growth chamber but could not overcome the flow velocity in the channels, thus remaining unable to reach the medium reservoirs.

We tested the microreactor to verify if it effectively allowed bacterial cultures to reach stationary states. We performed experiments with *E. coli* GL69 using two different media: M9 + 30 mM glucose and M9 + 48 mM tryptophan. Each experiment lasted 48 h and was repeated 5 times. We sampled the cultures at different time intervals, measured their OD_{600}, and calculated the fluorescence intensity of each individual cell within the sample—we did this following the methods outlined in section 2.4.

Results of growth measurements are displayed in Fig.4.5. Notice how mean OD_{600} values stay constant after 18 h, in either glucose (A) or tryptophan (B) media, meaning the cell growth reaches a stationary state at this point.

However, growth by itself is an insufficient evidence of steady states regarding genetic expression; that is why we calculated the fluorescence intensity PDFs of each culture sample taken at 18, 24, 42, and 48 h; we averaged the means of the resulting distributions and compared them to find out changes in expression along time. Resulting fluorescence data is plotted in Fig. 4.6. Error bars also denote standard deviation in these cases. Notice how mean fluorescence intensity stays at similar values 24 h after inoculation, both in glucose (A) and tryptophan (B) supplemented media. Expression in glucose remains noticeable lower than in tryptophan, which is expected due to the repressive role of glucose in the *tna* operon regulation pathways.

These results evidence the existence of a steady state in expression after 24 hours of constant

culture conditions inside the microreactor. All subsequent continuous flow cultures in this work lasted at least 24 h to ensure that each one attained a steady state before sampling.

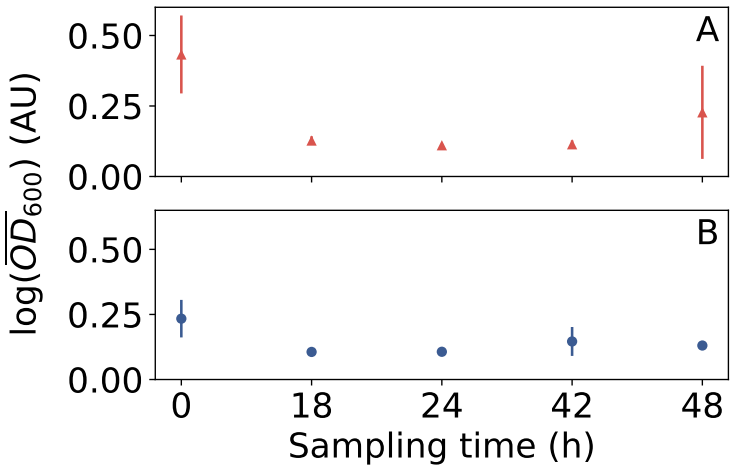


Figure 4.5: Mean OD_{600} of bacterial cultures inside the microreactor at different times. Averaged OD_{600} values of 5 experiments for each media used: M9 + 30 mM glucose (A) and M9 + 48 mM tryptophan (B). Error bars denote standard deviation.

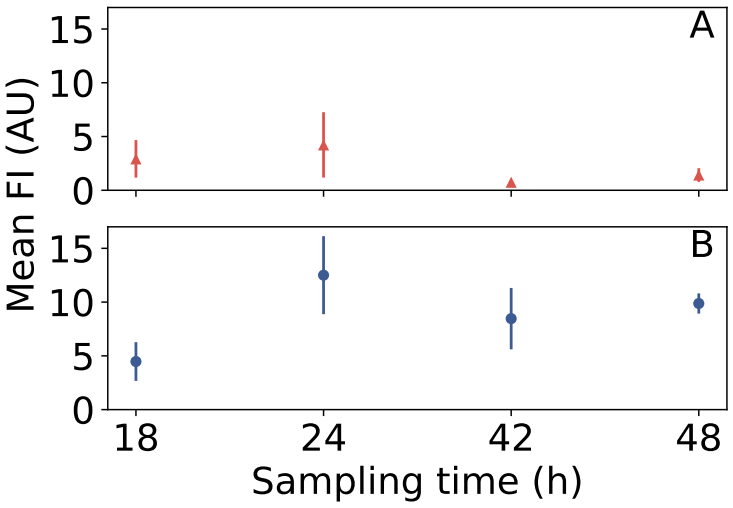


Figure 4.6: Mean fluorescence intensity of bacterial cultures inside the microreactor at different times. Averaged fluorescence intensity of 5 experiments for each media used: M9 + 30 mM glucose (A) and M9 + 48 mM tryptophan (B). Error bars denote standard deviation.

3 Experiments in steady state

After verifying the viability of the micro-scale continuous flow culture protocol and feasibility of steady states inside the microreactor, we performed three experiments with *E. coli* GL69 per each combination of glucose and tryptophan concentrations stated in the methodology section (ch. 3, sec. 2.3). Each culture lasted at least 24 h to ensure the attainment of steady states. It was important to take into account the intrinsic stochasticity of each cell to get an accurate description of the culture's current expression state. As such, we required to measure fluorescence intensity of each single cell within the retrieved samples.

A commonly used technique to study genetic expression at single cell level is the fluorescence activated cell sorting (FACS). This is a cytometry technique that allows the separation of cell populations within a sample in accordance to different criteria such as particle size, opacity, and fluorescence intensity (Herzenberg et al., 2002).

However, FACS is not a viable to this project because it is not precise enough to account for intracellular structures in bacteria, as required by the very localized nature of the inactive tryptophanase within the cells. Instead, we used fluorescence phase contrast microscopy to acquire short videos from static experiment samples, both in bright and fluorescence field. Later, we analyzed those videos to count down each cell within the samples and measured their fluorescence intensity.

We identified four kinds of fluorescence states within the cells: no fluorescence, spread fluorescence, foci fluorescence, and foci and spread fluorescence simultaneously (see Fig. 4.7). Spread fluorescence is assumed to happen when tryptophanase molecules assemble into tetramers and activate, thus dispersing away from the initial focus—which contains only inactive monomers and dimers.

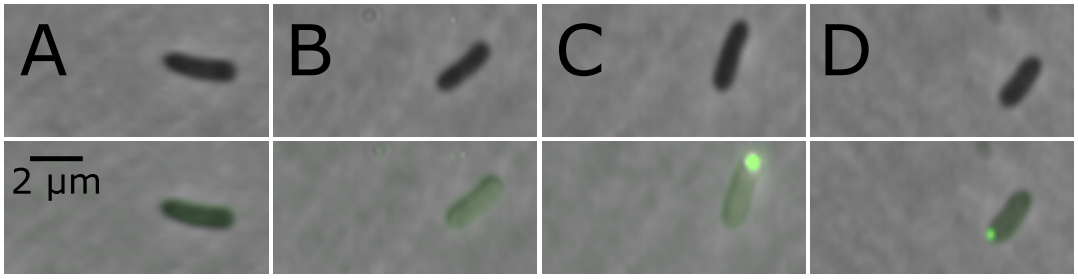


Figure 4.7: Micropraphs of *E.coli* GL69 with different levels of *tnaA-sfGFP* expression. Images in the higher row were taken in bright field illumination. Lower row shows the same bacteria under fluorescence excitation—background fluorescence was subtracted and bright field images were overlaid. These micrographs show four different kinds of fluorescence observed inside cells: A) No fluorescence, B) Spread fluorescence, C) Spread and focus fluorescence within the same cell, D) Fluorescence confined to a single focus.

The experimental strategy of this project was planned to retrieve data from multiple fixed points over the parameter space, which is the set of all possible combinations of values for all the different parameters contained in a particular mathematical model—In this case, those parameters are the tryptophan and glucose concentrations—this is because a bifurcation diagram with three dimensions will project over the parameter space (Strogatz, 1994), therefore we expect to explore the system's bifurcation diagram from our experimental results.

As stated previously, we took videos on the microscope under bright field and fluorescence excitation settings. Then, we analyzed each samples videos following the protocol outlined in section 2.4.2 (ch.3). We estimated each cell's fluorescence density by normalizing their total fluorescence by their area. This way, it is possible to picture *tnaA* gene expression regardless of the TnaA activation state.

Considering that the results discussed on the previous section evidence the repeatibility of the experiments in steady state, we lumped the data of all samples within each experiment for its subsequent analysis. Results are displayed in Fig. 4.8. Histograms in each plot stand for the fluorescence PDF and the red curves denote a gaussian kernel density estimation to illustrate the presence of subpopulations within each sample. Notice that the probability of observing cells with a higher fluorescence level increases along tryptophan concentration. On the other side, higher glucose concentrations imply that the system requires higher amounts of tryptophan to express.

Beyond 15 mM glucose, the operon appears to be barely expressed even at the highest tryptophan concentration (48 mM).

Interestingly, the effect of pyruvate over the *tna* operon expression seems to be negligible. Early works on tryptophanase regulation emphasized the role of catabolite repression and pyruvate was deemed one of the *tna* operon's main repressors (Botsford and DeMoss, 1971); after all, it plays a central role in the phosphoenolpyruvate transfer system, the main precursor mechanism to cAMP-CAP dependent catabolite repression (Deutscher et al., 2006). Data displayed in figure 4.8 contradict this statement—in the absence of glucose, *tnaA-sfGFP* expression increased along with tryptophan concentration. If pyruvate were effectively a repressor for the *tna* operon, most distributions would resemble the one at 30 mM glucose + 0 mM tryptophan since all the tested media included 18 mM pyruvate.

Several of the PDFs are bimodal, as illustrated by the kernel density estimation. Bimodality suggests this operon shows bistable behavior in some ranges of tryptophan and glucose concentrations. Namely, at 7.5 mM glucose and below, bimodality is observed at 6–48 mM tryptophan. At 15 mM glucose, bimodality is only observed at 48 mM tryptophan. Finally, at the highest glucose concentration (30 mM), there is no bimodality observed.

Bistability, from a cell population perspective, has been defined as phenotypical heterogeneity manifested by a bifurcation into distinct subpopulations. It differs from other heterogeneity events—like phase variations—in that the changes it brings usually last several generations and tend to be hysteretic (Dubnau and Losick, 2006). This phenomenon is due to stochastic fluctuations in cellular components and certain regulation architectures (Veening et al., 2008). This implies that, under proper conditions, the expression of a bistable system would render a bimodally distributed population—such as many data have shown in figure 4.8.

One of the proposed bistability mechanisms is the double negative feedback, in which one element represses another and vice versa, and the state switch is given by an external inducer to the system. A good example of this is the λ-phague lysis-lysogeny switch (Ptashne, 2004). Another mechanism is the two-component positive feedback loop, where the activity of the first component induces the other, which in turn enhances the first one. *E. coli* has a well known system

with this arrangement: the *lac* operon. This operon has been hypothesized to be bistable because it has all the traits required for bistability (Santillán and Mackey, 2008).

The *lac* operon is in charge of processing lactose as a carbon source in the absence of glucose. However, several works have shown under experimental conditions that bistability doesn't arise under induction by lactose but by the synthetic inducer IPTG, or by an artificial enhancement of its repressor (Ozbudak et al., 2004; Zander et al., 2017); which defeats the proposed purposes of bistability regarding survival and long term adaption to sudden medium composition changes. The *tna* operon not only presents a similar architecture to that of the *lac*, but it also holds a comparable function. Hence, in light of its previously reviewed traits and behavior shown in the data analysis, is an experimentally proven two-component positive feedback bistable system.

From the discussion in previous sections on the features of bistability (ch.1, sec2), it is relevant at this point to recall two specific traits that may have biological implications: hysteresis and hypersensitivity. Both are related to the velocity of the system's response to changes in its environment; the former implies a delay in the expression of the operon and the latter a sudden switch in its expression state. While they may appear to hold opposing functions, bistable behavior is the outcome of their combined effects. Experimentally, this implies both effects are predictable and observable at specific conditions. Those conditions, namely parameter values, need to be defined from available experimental data. In this case, as previously stated, our experimental strategy allows for the exploration of the parameter space; henceforth, the proposal of a mathematical model able to predict the boundaries of the system's bistability region within its bifurcation diagram. This would allow to design an experiment akin to those of (Gardner et al., 2000), where it would be possible to accurately switch the system state, thus defeating hysteresis and evidencing hypersensitivity.

Currently, the development of such a mathematical model falls outside the scope of this project; however, we may draw some preliminary conclusions regarding the conditions upon the system's bistable switch—and consequently the system's hypersensitivity parameter values. First, the bistability region is extremely wide. This is evidenced by the presence of bimodal distributions at 48 mM tryptophan because that is an extremely high concentration, extremely unlikely to be found in natural conditions. Bimodal distributions imply that the system still

remains within the boundaries of the bistable region. From an evolutionary stance, this situation may arise as a survival or fitness enhancement strategy (Magdanova and Golyasnaya, 2013); the population heterogeneity rooted in a bimodal distribution may ensure the survival of at least some individuals if the environmental conditions change abruptly: those individuals would have a higher probability to survive such scenario. Another preliminary conclusion is that the complexity of the operon's regulatory pathway may give rise to unexpected situations were the interactions between the operon and its effectors may blur the boundaries of the bistability region. We discuss these details in the subsequent paragraphs.

Fluorescence density by itself accounts for *tnaA-sfGFP* expression levels; however, as previously discussed, expression does not imply activity. To quantify the tryptophanase activity levels within each set of conditions, we plotted the PDF that a cell has a given focus fluorescence intensity and a given spread fluorescence. The results are shown in figure 4.9. This is a set of heatmaps following the same structure as the histograms grid in figure 4.8.

Each heatmap depicts a sample under specific glucose and tryptophan concentrations, and each one has its own fluorescence intensity maximum value. The color code allows for the comparison of population composition between samples. The color of each pixel within the heatmaps stands for the fraction of cells in a population with given values of both foci and spread fluorescence. At 0 mM glucose, the fraction of cells with foci fluorescence increases together with tryptophan concentration up to 24 mM and then it decreases again at 48 mM. This behavior pattern appears also at 3.75 and 7.5 mM glucose although it is less noticeable. Finally, at 15 mM and 30 mM glucose, the probability of finding cells with foci is negligible.

While figure 4.9 offers information on the distribution of different foci and spread fluorescence fractions due to different concentrations of glucose and tryptophan, figure 4.10 shows a simple summary on the effects of both reagents on the operon. The blue areas in the pie charts stand for the average of the spread fluorescence fraction of individual cells in the sample, i.e. the ratio of spread fluorescence to the total sum of both disperse and foci fluorescence. Red stands for the average of cells' foci fluorescence fraction.

At 0 mM glucose, the fraction of foci fluorescence increases along with tryptophan

concentration until 24 mM and decreases again at 48 mM. This pattern repeats in all rows but at 3.75 mM glucose, where there is no discernible pattern. As for the effects of glucose, there is a remarkable increase in the fraction of foci fluorescence up to 3.75 mM and a sharp decrease from 7.5 mM. This happens in all columns except in that corresponding to 24 mM tryptophan where there's only a sustained decrease on the foci fraction while glucose concentration increases.

Li and Young (2014) demonstrated the effect of glucose on the *tna* operon went beyond catabolite repression; it appeared to halt the expression of *tnaA* even when the cAMP-CAP complex formation remained untouched, thus avoiding promoter exclusion. Furthermore, glucose also appeared to prevent the activation of newly formed tryptophanase. They propose the existence of a post-translational regulation mechanism. Results plotted in figures 4.9 and 4.10 suggest this mechanism is even more complex, since tryptophan also appears to be involved. To the best of our knowledge, tryptophan involvement in TnaA activation has not been previously suggested nor reported.

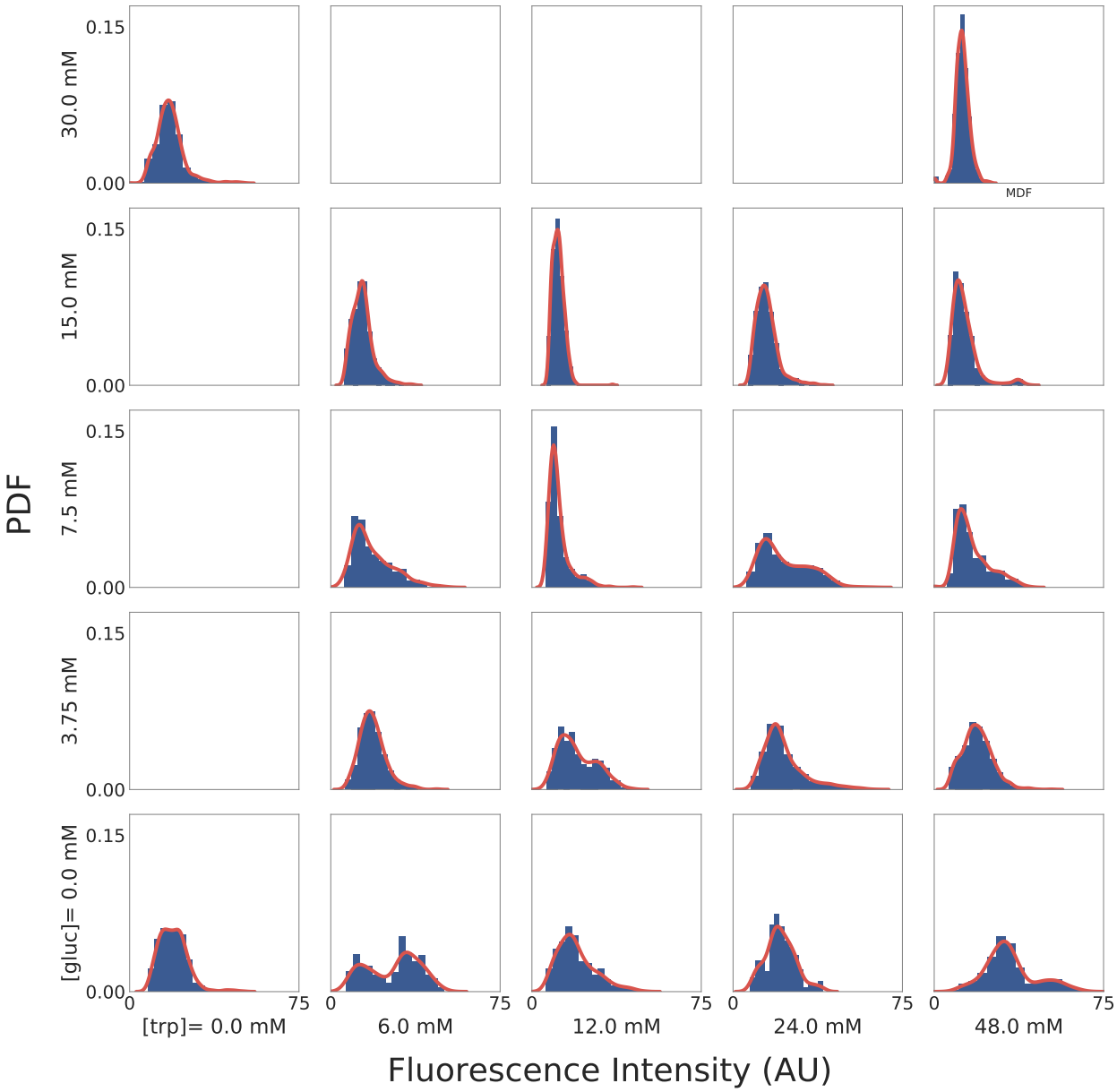


Figure 4.8: PDF OF FLUORESCENCE INTENSITY FOR EACH TESTED MEDIUM CONDITION. Blue areas are normalized histograms of fluorescence intensity for all samples tested in each set of conditions. Overlaid red curves result from gaussian kernel density estimations. No data were retrieved for blank sets.

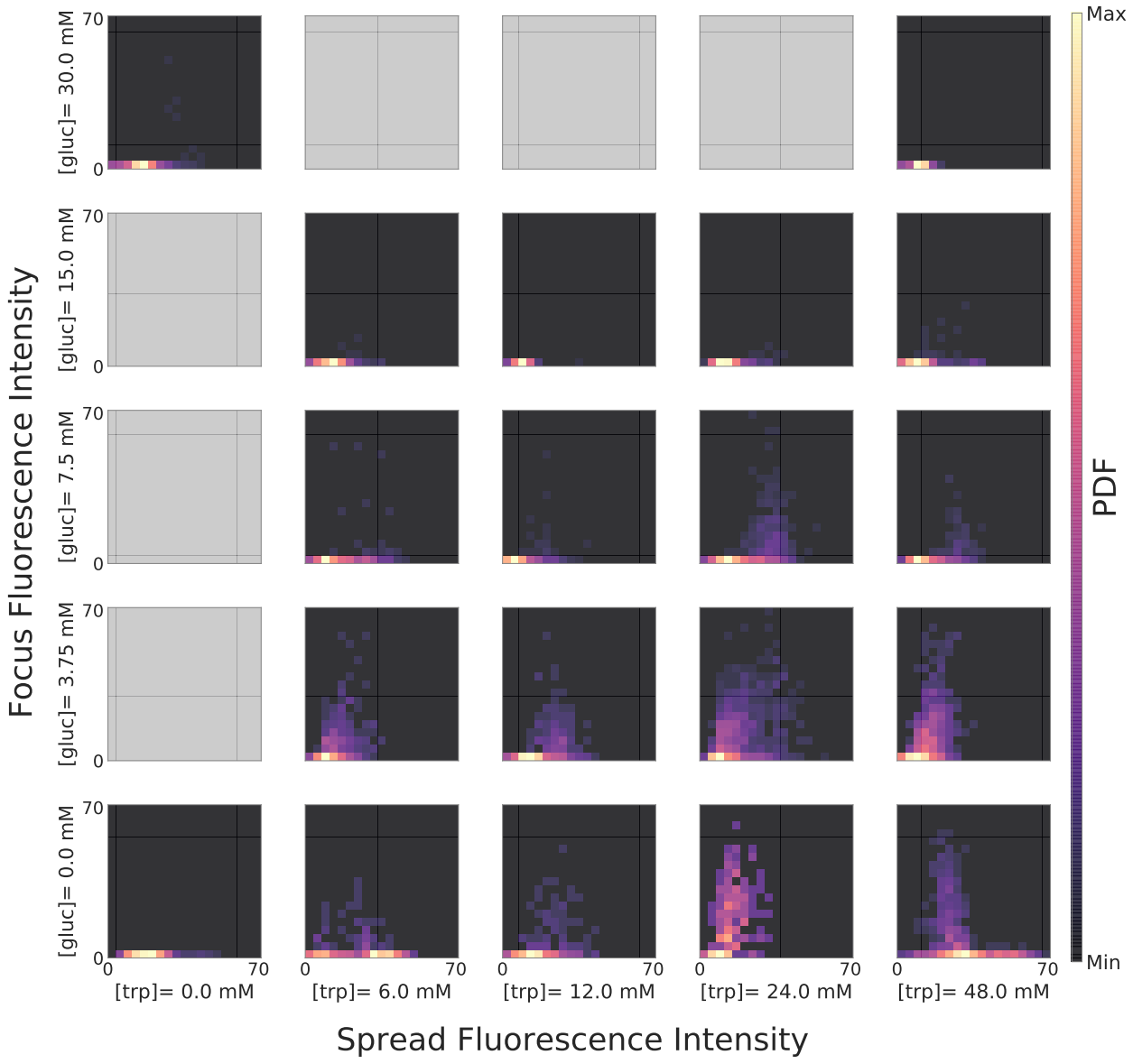


Figure 4.9: PDF of fluorescence intensity as a function of focus fluorescence and spread fluorescence for each tested medium condition. Each set has its own maximum value. Values closer to x axis mean higher probability of spread fluorescence. Values over y axis denote higher probability of focus fluorescence. There were no data gathered for sets shaded in gray.

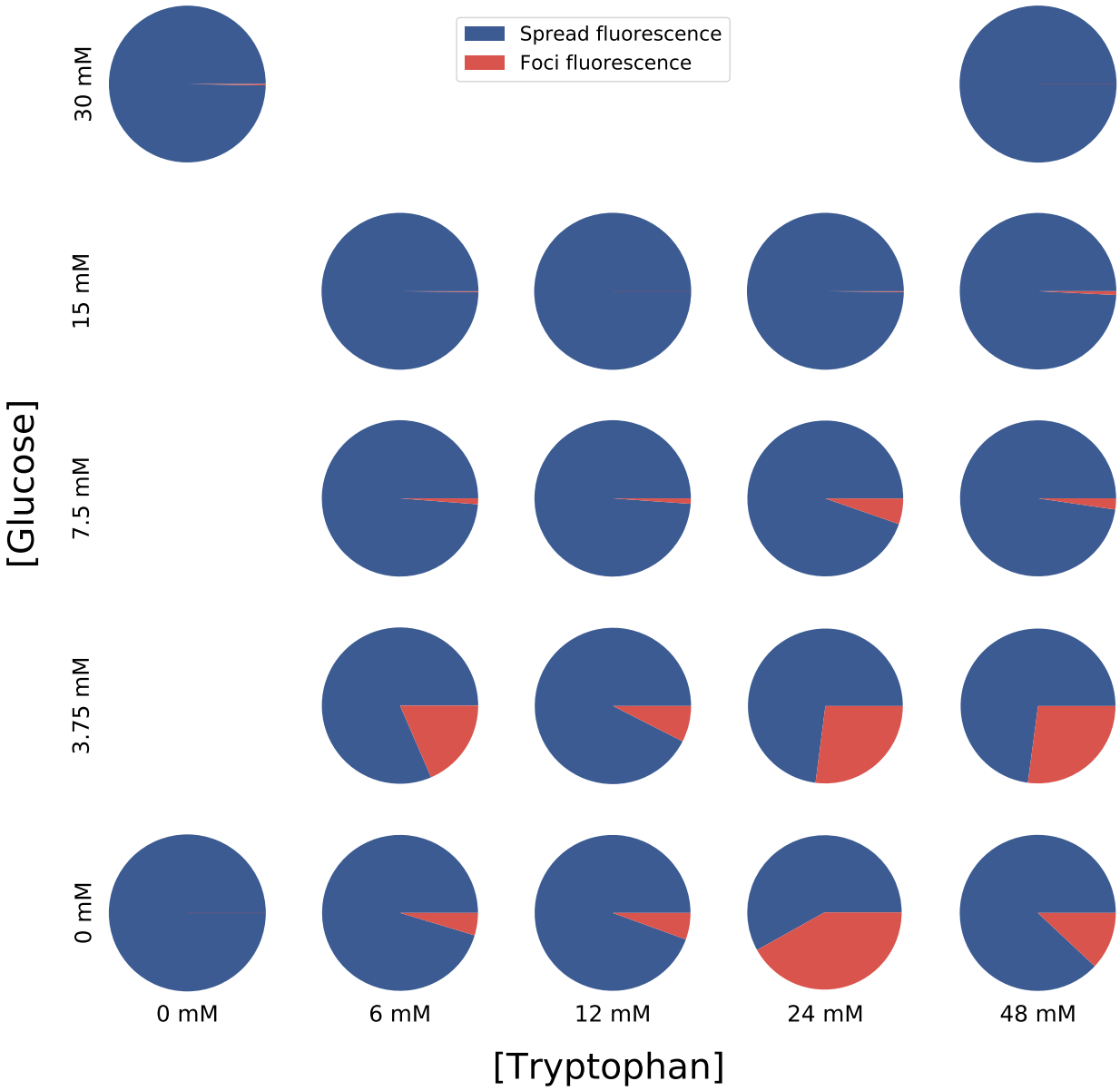


Figure 4.10: Mean fractions of spread and foci fluorescence within cells. Blue areas stand for the averaged fractions of spread fluorescence intensity in the sample cells, this also applies for red areas in the case of foci fluorescence intensisty.

5 Closing remarks

1 Conclusions

We studied the dynamic behavior of the *tna* operon in *E. coli*. We aimed to demonstrate bistability in its regulatory pathway. A classical approach to address these kind of objectives is mathematical modeling supported by data already reported in scientific literature. However, in the case of the *tna* operon, valuable information on the subject was scarce or barely appropriate to develop a precise mathematical model. Previous attempts to do so (Gomez-Schiavon, 2011), while mathematically solid, were unable to be proven experimentally; mostly due to the scantness of accurate information on the behavior of the operon. Our approach was purely experimental, yet intended to fill some of the extant gaps that disallow accurate mathematical approaches to explore this genetic expression system.

Regarding bystability, we conclude that the regulation dynamics of the *tna* operon suggested it to be bistable. TnaA expression data show bimodal distributions within a delimited range of tryptophan and glucose concentrations. Bimodality itself is a strong evidence of bistability. Thus, we conclude *E. coli*'s *tna* operon is bistable. However, there's still much work pending to accurately delimit the full range of the bistability region, which we have proven to extend well beyond natural conditions. We argue that this strategy poses evolutionary and survival advantages related to population heterogeneity, since it may increase the odds that at least some individuals survive sudden changes in their environmental conditions.

Finally, bistability alone is insufficient to explain the behaviors observed beyond the *tnaA* gene expression. Most specifically, those related to tryptophanase activation. The proportions of active and inactive TnaA within each individual cell of every population sampled suggest a yet unknown post-translational regulation mechanism. From data retrieved in this project, we conclude that glucose and tryptophan concentration regulate TnaA foci disaggregation through a complex yet unexplored process.

2 Perspectives

1. The experimental strategy was designed to gather information in a way that allowed us to draft the parameters space of a mathematical model describing the *tna* operon. While the intention behind this design was to unveil the presence of bistability in the operon dynamics, it also allows the data retrieved to be detailed and relevant enough to create an accurate deterministic model based on them. In this regard, we propose the development of such model to further aid in the study of this or any possible bistable genetic regulation systems.

2. During the development of this project, we generated several mutants based off the *E. coli* GL69 strain with strategic deletions of several genes related to tryptophan transportation and catabolite repression. We propose to perform experiments on steady state with those strains, since they may possibly reveal more details on the operon dynamics and its interaction with other *E. coli* genetic regulation systems.

3. Indole may have a relevant role on the regulation of TnaA activation. Since indole effects may appear on a timescale shorter than the onset of steady states—because its main functions are related to intercellular communication and fast population responses to medium changes—our experimental setup requires to be adjusted to take this into account. Hence, we propose to trap a single cell or colony of *E. coli* GL69 inside a microreactor to monitor changes on its expression over time under constant concentrations of glucose, tryptophan and exogenous indole.

Given the importance of indole and its roles in population and inter–kingdoms dynamics; if its presence triggers different foci disaggregation speeds, several questions would arise regarding the *tna* operon regulation pathway, the *E. coli* population dynamics, and the bacteria-host interaction.

Bibliography

Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A., and Alon, U. (2014). Hierarchy of non-glucose sugars in Escherichia coli. *BMC systems biology*, 8:133.

Alon, U. (2007). *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Taylor & Francis Group, Boca Raton, FL.

Apostu, R. and Mackey, M. C. (2012). Mathematical model of GAL regulon dynamics in Saccharomyces cerevisiae. *Journal of theoretical biology*, 293:219–35.

Babloyantz, A. and Sanglier, M. (1972). Chemical instabilities of "all-or-none" type in beta - galactosidase induction and active transport. *FEBS letters*, 23(3):364–366.

Beckwith, J. (1987). The lactose operon. In Neidhart, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., and Umbarger, H. E., editors, *Escherichia coli and Salmonella thyphymurium: cellular and molecular biology, Volume 2*, pages 1439–1443. American Society for Microbiology, Washington, DC.

Bialek, W. (2018). Perspectives on theory at the interface of physics and biology. *Reports on Progress in Physics*, 81(1):012601.

Blumenthal, T. (2004). Operons in eukaryotes. *Briefings in Functional Genomics and Proteomics*, 3(3):199–211.

Botsford, J. L. and DeMoss, R. D. (1971). Catabolite repression of tryptophanase in Escherichia coli. *Journal of Bacteriology*, 105(I):303–312.

Boudvillain, M., Figueroa-Bossi, N., and Bossi, L. (2013). Terminator still moving forward: expanding roles for Rho factor. *Current opinion in microbiology*, 16(2):118–24.

Bren, A., Park, J. O., Towbin, B. D., Dekel, E., Rabinowitz, J. D., and Alon, U. (2016). Glucose becomes one of the worst carbon sources for E.coli on poor nitrogen sources due to suboptimal levels of cAMP. *Scientific reports*, 6(April):24834.

Brückner, R. and Titgemeyer, F. (2002). Carbon catabolite repression in bacteria: Choice of the carbon source and autoregulatory limitation of sugar utilization. *FEMS Microbiology Letters*, 209:141–148.

Busby, S. and Ebright, R. H. (1999). Transcription activation by catabolite activator protein (CAP). *Journal of molecular biology*, 293(2):199–213.

Cruz-Vera, L. R., New, A., Squires, C., and Yanofsky, C. (2007). Ribosomal features essential for tna operon induction: tryptophan binding at the peptidyl transferase center. *Journal of bacteriology*, 189(8):3140–6.

Danino, T., Mondragón-Palomino, O., Tsimring, L., and Hasty, J. (2010). A synchronized quorum of genetic clocks. *Nature*, 463(7279):326–330.

Darnton, N. C., Turner, L., Rojevsky, S., and Berg, H. C. (2007). On torque and tumbling in swimming Escherichia coli. *Journal of bacteriology*, 189(5):1756–64.

Datsenko, K. a. and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America*, 97(12):6640–5.

Deeley, M. C. M. and Yanofsky, C. (1982). Transcription initiation at the tryptophanase promoter of Escherichia coli K-12. *Journal of Bacteriology*, 151(2):942–951.

Deutscher, J., Francke, C., and Postma, P. W. (2006). How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria. *Microbiology and Molecular Biology Reviews*, 70(4):939–1031.

Dubnau, D. and Losick, R. (2006). Bistability in bacteria. *Molecular microbiology*, 61(3):564–72.

Ebright, R. H. (1993). Transcription activation at Class I CAP-dependent promoters. *Molecular microbiology*, 8:797–802.

Edwards, R. M. and Yudkin, M. D. (1982). Location of the gene for the low-affinity tryptophan-specific permease of Escherichia coli. *The Biochemical journal*, 204(2):617–9.

Escalante, A., Cervantes, A. S., Gosset, G., and Bolívar, F. (2012). Current knowledge of the Escherichia coli phosphoenolpyruvate-carbohydrate phosphotransferase system: Peculiarities of regulation and impact on growth and product formation. *Applied Microbiology and Biotechnology*, 94(6):1483–1494.

Gardner, T. S., Cantor, C. R., and Collins, J. J. (2000). Construction of a genetic toggle switch in Escherichia coli. *Nature*, 403(6767):339–42.

Gomez-Schiavon, M. (2011). *Modelo del circuito de regulación del operón de triptofanasa en Escherichia coli*. Msc thesis, Cinvestav-Monterrey.

Gong, F. and Yanofsky, C. (2002). Analysis of tryptophanase operon expression in vitro: accumulation of TnaC-peptidyl-tRNA in a release factor 2-depleted S-30 extract prevents Rho factor action, simulating induction. *The Journal of biological chemistry*, 277(19):17095–100.

Gosset, G. (2005). Improvement of Escherichia coli production strains by modification of the phosphoenolpyruvate:sugar phosphotransferase system. *Microbial cell factories*, 4(1):14.

Gu, P., Yang, F., Li, F., Liang, Q., and Qi, Q. (2013). Knocking out analysis of tryptophan permeases in Escherichia coli for improving L-tryptophan production. *Applied Microbiology and Biotechnology*, 97(15):6677–6683.

Heatwole, V. M. and Somerville, R. L. (1991). The tryptophan-specific permease gene, mtr, is differentially regulated by the tryptophan and tyrosine repressors in Escherichia coli K-12. *Journal of bacteriology*, 173(11):3601–4.

Herzenberg, L. A., Parks, D., Sahaf, B., Perez, O., Roederer, M., and Herzenberg, L. A. (2002). The history and future of the fluorescence activated cell sorter and flow cytometry: a view from Stanford. *Clinical chemistry*, 48(10):1819–27.

Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, 3(3):318–356.

Jacob, F., Perrin, D., Sanchez, C., and Monod, J. (1960). Operon: a group of genes with the expression coordinated by an operator. *Comptes rendus hebdomadaires des seances de l'Academie des sciences*, 250(6):1727–9.

Kaplan, D. L. and O'Donnell, M. (2003). Rho factor: transcription termination in four steps. *Current biology : CB*, 13(18):R714–6.

Konan, K. V. and Yanofsky, C. (1997). Regulation of the Escherichia coli tna operon: nascent leader peptide control at the tnaC stop codon. *Journal of bacteriology*, 179(5):1774–9.

Konan, K. V. and Yanofsky, C. (2000). Rho-Dependent Transcription Termination in the tna Operon of Escherichia coli: Roles of the boxA Sequence and the rut Site. *Journal of Bacteriology*, 182(14):3981–3988.

Kreth, J., Lengeler, J. W., and Jahreis, K. (2013). Characterization of Pyruvate Uptake in Escherichia coli K-12. *PLoS ONE*, 8(6):e67125.

Lange, R., Barth, M., and Hengge-Aronis, R. (1993). Complex transcriptional control of the sigma s-dependent stationary-phase-induced and osmotically regulated osmY (csi-5) gene suggests novel roles for Lrp, cyclic AMP (cAMP) receptor protein-cAMP complex, and integration host factor in the stationary-phas. *Journal of bacteriology*, 175(24):7910–7.

Lawson, C. L., Swigon, D., Murakami, K. S., Darst, S. a., Berman, H. M., and Ebright, R. H. (2004). Catabolite activator protein: DNA binding and transcription activation. *Current Opinion in Structural Biology*, 14(1):10–20.

Li, G. and Young, K. D. (2012). Isolation and identification of new inner membrane-associated proteins that localize to cell poles in Escherichia coli. *Molecular Microbiology*, 84(2):276–295.

Li, G. and Young, K. D. (2013). Indole production by the tryptophanase TnaA in Escherichia coli is determined by the amount of exogenous tryptophan. *Microbiology*, 159(Pt_2):402–410.

Li, G. and Young, K. D. (2014). A cAMP-independent carbohydrate-driven mechanism inhibits tnaA expression and TnaA enzyme activity in Escherichia coli. *Microbiology*, 160(Pt_9):2079–2088.

Li, G. and Young, K. D. (2015). A new suite of tnaA mutants suggests that Escherichia coli tryptophanase is regulated by intracellular sequestration and by occlusion of its active site. *BMC microbiology*, 15(1):14.

Magdanova, L. A. and Golyasnaya, N. V. (2013). Heterogeneity as an adaptive trait of microbial populations. *Microbiology*, 82(1):1–10.

Monod, J. (1949). The Growth of Bacterial Cultures. *Annual Review of Microbiology*, 3(1):371–394.

Mueller, R. S., Beyhan, S., Saini, S. G., Yildiz, F. H., and Bartlett, D. H. (2009). Indole acts as an extracellular cue regulating gene expression in Vibrio cholerae. *Journal of bacteriology*, 191(11):3504–16.

Newton, W. A. and Snell, E. E. (1964). CATALYTIC PROPERTIES OF TRYPTOPHANASE, A MULTIFUNCTIONAL PYRIDOXAL PHOSPHATE ENZYME. *Proceedings of the National Academy of Sciences of the United States of America*, 51(3):382–9.

Ng, J. M. K., Gitlin, I., Stroock, A. D., and Whitesides, G. M. (2002). Components for integrated poly(dimethylsiloxane) microfluidic systems. *Electrophoresis*, 23(20):3461–73.

Novick, A. and Weiner, M. (1957). ENZYME INDUCTION AS AN ALL-OR-NONE PHENOMENON. *Proceedings of the National Academy of Sciences of the United States of America*, 43(7):553–66.

Ozbudak, E. M., Thattai, M., Lim, H. N., Shraiman, B. I., and Van Oudenaarden, A. (2004). Multistability in the lactose utilization network of Escherichia coli. *Nature*, 427(6976):737–40.

Pearson, H. (2006). Genetics: what is a gene? *Nature*, 441(7092):398–401.

Piñero-Fernandez, S., Chimerel, C., Keyser, U. F., and Summers, D. K. (2011). Indole transport across Escherichia coli membranes. *Journal of Bacteriology*, 193(8):1793–1798.

Ptashne, M. (2004). *A Genetic Switch: Phage Lambda Revisited*. Cold Spring Harbor Laboratory Press, 3rd edition.

Rety, S., Deschamps, P., and Leulliot, N. (2015). Structure of Escherichia coli tryptophanase purified from an alkaline-stressed bacterial culture. *Acta Crystallographica Section F Structural Biology Communications*, 71(11):1378–1383.

Reznikoff, W. S. (1992). The lactose operon-controlling elements: a complex paradigm. *Molecular microbiology*, 6(17):2419–22.

Sackmann, E. K., Fulton, A. L., and Beebe, D. J. (2014). The present and future role of microfluidics in biomedical research. *Nature*, 507(7491):181–189.

Santillán, M. and Mackey, M. C. (2008). Quantitative approaches to the study of bistability in the lac operon of Escherichia coli. *Journal of the Royal Society, Interface*, 5 Suppl 1(SUPPL. 1):S29–39.

Santillán, M., Mackey, M. C., and Zeron, E. S. (2007). Origin of bistability in the lac Operon. *Biophysical journal*, 92(June):3830–3842.

Sarsero, J. P., Wookey, P. J., Gollnick, P., Yanofsky, C., and Pittard, a. J. (1991). A new family of integral membrane proteins involved in transport of aromatic amino acids in Escherichia coli. *J Bacteriol*, 173(10):3231–3234.

Schulze, K. L. and Lipe, R. S. (1964). Relationship between Substrate Concentration, Growth Rate, and Respiration Rate of Escherichia coli in Continuous Culture. *Archiv fur Mikrobiologie*, 48(1):1–20.

Shehata, T. E. and Marr, A. G. (1971). Effect of nutrient concentration on the growth of Escherichia coli. *Journal of bacteriology*, 107(1):210–6.

Shu, C.-C., Chatterjee, A., Dunny, G., Hu, W.-S., and Ramkrishna, D. (2011). Bistability versus bimodal distributions in gene regulatory processes from population balance. *PLoS computational biology*, 7(8):e1002140.

Snell, E. E. (1975). Tryptophanase: Structure, Catalytic Activities, and Mechanism of Action. In Meister, A., editor, *Advances in enzymology and related areas of molecular biology*, volume 42, pages 287–333. John Wiley & Sons, Inc., Berkeley, California.

Stewart, V. and Yanofsky, C. (1985). Evidence for transcription antitermination control of tryptophanase operon expression in Escherichia coli K-12. *Journal of bacteriology*, 164(2):731–40.

Strogatz, S. H. (1994). *Nonlinear Dynamics And Chaos: with applications to physics, biology, chemistry and engineering*. Perseus Books.

Tiwari, A., Ray, J. C. J., Narula, J., and Igoshin, O. A. (2011). Bistable responses in bacterial genetic networks: Designs and dynamical consequences. *Mathematical Biosciences*, 231(1):76–89.

Trewavas, A. (2006). A brief history of systems biology. "Every object that biology studies is a system of systems." Francois Jacob (1974). *The Plant cell*, 18(10):2420–30.

Tyson, J. J., Chen, K. C., and Novak, B. (2003). Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology*, 15(2):221–231.

Veening, J.-W., Smits, W. K., and Kuipers, O. P. (2008). Bistability, epigenetics, and bet-hedging in bacteria. *Annual review of microbiology*, 62(1):193–210.

Wanner, B. L., Kodaira, R., and Neidhardt, F. C. (1978). Regulation of lac operon expression: Reappraisal of the theory of catabolite repression. *Journal of Bacteriology*, 136(3):947–954.

Ward, D. F. and Yudkin, M. D. (1976). Mutations in Escherichia coli that relieve catabolite repression of tryptophanase synthesis. Tryptophanase promoter-like mutations. *Journal of general microbiology*, 92(1):133–7.

Weibel, D. B., Diluzio, W. R., and Whitesides, G. M. (2007). Microfabrication meets microbiology. *Nature reviews. Microbiology*, 5(3):209–18.

Wood, W. A., Gunsalus, I. C., and Umbreit, W. W. (1947). Function of pyridoxal phosphate: resolution and purification of the tryptophanase enzyme of Escherichia coli. *Journal of Biological Chemistry*, 170(1):313–321.

Yang, R., Cruz-Vera, L. R., and Yanofsky, C. (2009). 23S rRNA nucleotides in the peptidyl transferase center are essential for tryptophanase operon induction. *Journal of bacteriology*, 191(11):3445–50.

Yanofsky, C. (1960). The tryptophan synthetase system. *Bacteriological reviews*, 24(2):221–45.

Yanofsky, C. (1981). Attenuation in the control of expression of bacterial operons. *Nature*, 289(5800):751–8.

Yildirim, N. and Mackey, M. C. (2003). Feedback regulation in the lactose operon: A mathematical modeling study and comparison with experimental data. *Biophysical Journal*, 84(5):2841–2851.

Zander, D., Samaga, D., Straube, R., and Bettenbrock, K. (2017). Bistability and Nonmonotonic Induction of the lac Operon in the Natural Lactose Uptake System. *Biophysical journal*, 112(9):1984–1996.

Zhang, H. and Chiao, M. (2015). Anti-fouling Coatings of Poly(dimethylsiloxane) Devices for Biological and Biomedical Applications. *Journal of Medical and Biological Engineering*, 35(2):143–155.

Zhou, Y., Kolb, A., Busby, S. J. W., and Wang, Y.-P. (2014). Spacing requirements for Class I transcription activation in bacteria are set by promoter elements. *Nucleic acids research*, 42(14):9209–16.

Appendices

A Project resources

A.A *E. coli* strains

Strains listed in table A.1 were either constructed at or retrieved from Dr. Kevin Young's lab at UAMS. We used the λ RED method (Datsenko and Wanner, 2000) to delete genes of interest that may interfere with the *tna* operon functions. Primers used to perform those deletions are listed in table A.2.

Figure A.1 illustrates the relationship between the deleted genes and the operon. CyaA, ToiC and CpdA regulate cAMP intracellular concentration, the first by synthesis and the other two by exportation and degradation, respectively. Thus, all three of them are related to the cAMP-CAP dependent catabolite repression.

Mtr and AroP are the two known tryptophan permeases besides TnaB. Mtr is specific and high affinity. AroP also imports tyrosine and phenylalanine, although it has a higher affinity to tryptophan than TnaB.

Table A.1: *E. coli* strains

Strain name	Description	Source or reference
GL40	**MG1655** tnaA-sfGFP::frt::Kan::frt	Li and Young (2012)
GL69	**MG1655** tnaA-sfGFP::frt	Li and Young (2012)
DO03	**GL69** ΔaroP::frt::Kan::frt	This study
DO06	**GL69** ΔmtR::frt::Kan::frt	This study
DO01	**GL40** ΔaroP::frt ΔmtR::frt	This study
GL583	**GL69** ΔtoIC::frt ΔcyaA::frt::Kan::frt	Li and Young (2014)
GL473	**GL69** ΔcyaA::frt::Kan::frt	Li and Young (2014)
GL500	**GL69** ΔcpdA::frt ΔtoIC::frt::Kan::frt	Li and Young (2014)
GL504	**MG1655** ΔtnaA	Li and Young (2015)

Table A.2: Primers used in strains construction

Primer	Sequence	Function
aropH1P1	GGGTGAGGGCGTAGAGAGATTAATGCGCTTTT ACGGCTTTTGTAGGCTGGAGCTGCTTCG	*aroP* deletion by λ RED
aropH2P2	CACTGCGTAGATCAAAAAAACAACCACCGCA CGAGGTTTCATTCCGGGGATCCGTCGACC	
aropchkF	TACGGGTGAGGGCGTAGAGA	*aroP* deletion check
aropchkR	CGCTGCCGCATACCATTATT	
mtrH1P1	TCGGATAAGGCACCGCTGATTACTGATACACC GGCAGTAATGTAGGCTGGAGCTGCTTCG	*mtr* deletion by λ RED
mtrH2P2	ACAACGCAGTCGCACTATTTTTCACTGGAGAG AAGCCCTCATTCCGGGGATCCGTCGACC	
mtrchkF	AGCCTTATCCGAGCTGGCAA	*aroP* deletion check
mtrchkR	CGCTGAACAGCGAACACAATCT	

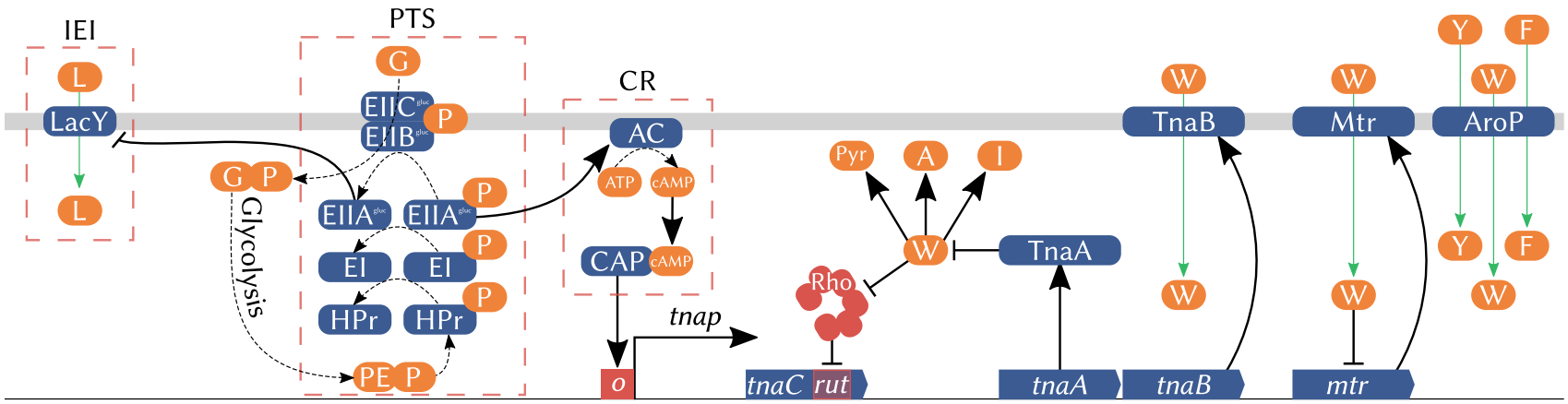
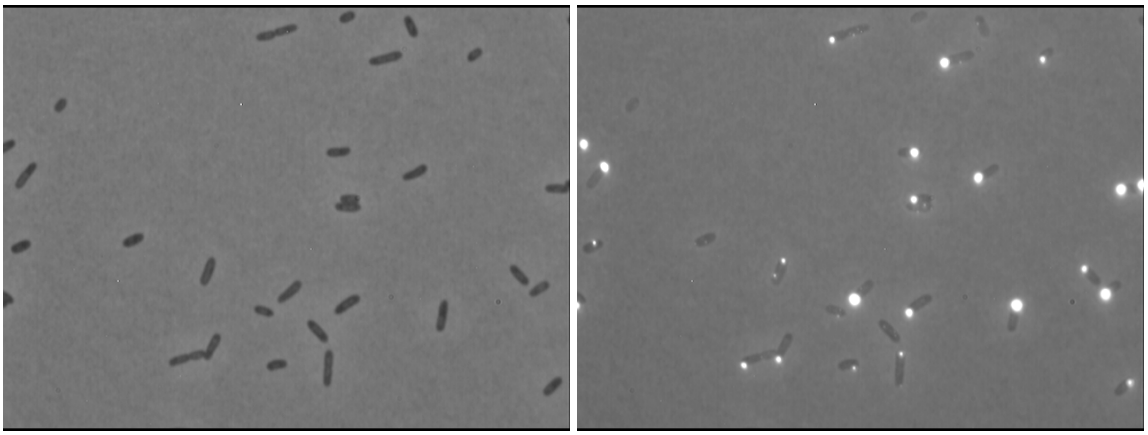


Figure A.1: Regulation elements that may interfere with the *tna* operon. The presence of extracellular sugars determines the outcome of the phosphotransferase system (PTS) (See Fig. 1.2). In this case, glucose (G) forces the phosphorilation (slim dashed arrows) of enzyme EIIA^{gluc}, which in turn represses (bold flat-headed arrows) the import of other carbon sources (IEI); on the other side, dephosphorilated EIIA^{gluc} cannot induce (bold pointed arrows) the activity of membrane bound AC, this is called catabolite repression (CR). Enzyme AC is in charge of producing cAMP, which forms the cAMP-CAP complex that induces RNApol to bind to promoter *tnap* (folded arrow), thus beginning the operon transcription. Downstream of the promoter, the *tnaC* sequence holds the *rut* binding site (See Fig. 1.3) whose function is to trigger the *r*-dependent termination of the operon expression. Then follow the *tnaA* and *tnaB* genes (their dynamics and interactions with exogenous tryptophan (W) are detailed in Fig. 1.1). Mtr and AroP are two permeases capable of importing (green arrows) tryptophan into the cytoplasm. Mtr expression is repressed by the presence of tryptophan; AroP is not a tryptophan exclusive permease—as opposed to TnaB and Mtr—and also imports phenylalanine (F) and tyrosine (Y). Neither Mtr nor AroP have been proven to interfere with the *tna* operon dynamics.

A.B Image analysis

Videos taken were averaged to a single image. Both phase contrast and fluorescence images were retrieved from every sample. An example of these images is shown in figure A.2.



(a) Phase contrast (b) Fluorescence

Figure A.2: Averaged images of recorded samples. (a) Averaged image of a 10 s video taken in a phase contrast microscope at 100X. (b) Same sample under fluorescence excitation.

All images were processed and analyzed with the following script. This was coded by Dr. Eduardo Sosa in MatLab R2017b®. The final product is a table with the data enlisted in the script's first section, per each bacterium found in all samples.

```
% BrightBact calculates the fluorescence for each bacteria and the corresponding foci if present.

% Targets
% 000. Hora
% 00. AreaBacteria
% 0. Bacteria(estimated)
% 1. Mean dispersed fluorescence
% 1.1 Total disperse fluorescence
% 2. Area in pixels with dispersed fluorescence
% 3. Mean fluorescence in foci (0 if no foci are present)
% 3.1 Total foci fluorescence
% 4. Total area in pixels of all foci within a cell (0 if no foci are present)
```

```
% 5. Sample number
% 6. Image number
% 07. Glucose
% 08. Tryptophan
% 09. Indole

% clear all;
% close all;
addpath tracking
warning('off','all');
warning

hlap = [0 1 0; 1 -4 1; 0 1 0];
se1 = strel('disk',3);
hm = 1/9.*[1 1 1;1 1 1;1 1 1];

res = '“ResultadosCD˙0';
directorio = [pwd,'“caract disp'];
mkdir([pwd,res]);

% Variable initialization in Hard Drive
data = [];
dirData = strcat(pwd,res,'“data.mat');
save(dirData,'data','-v7.3');
DataFile = matfile(dirData,'Writable',true);
contCnt = 1;
nameFolds=extractNames(directorio);
numFolds = size(nameFolds);

ThresD = 20; % Threshold for disperse fluorescence
ThresF = 90; % Threshold for foci fluorescence
BacteriaArea = 750; % Mean bacteria area
tic
for karp = 1:numFolds(1)%
    RutaFolder = dir(fullfile(directorio,strtrim(nameFolds(karp,:)),'*C.png')); %1:8 normal --- karp,1:18 para febrero
    nameFolds1 = –RutaFolder(˜[RutaFolder(:).isdir]).name˝'; %enter subfolder name
    nameFolds1(ismember(nameFolds1,–'.','..'˝)) = []; %removes . and ..
    nameFolds1 = char(nameFolds1); %char type so it is compatible with imread
    numFolds1 = size(nameFolds1);

    for vid = 1:numFolds1(1)%
        ruta = strcat(directorio,'“',strtrim(nameFolds(karp,:)),'“',nameFolds1(vid,1:end-5));
        imageFullFileNameC = strcat(ruta, 'C.png'); % Phase contrast image
        imageFullFileNameF = strcat(ruta, 'F.png'); % Phase contrast image with fluorescence
        IC = (imread(imageFullFileNameC)); %rgb2gray
        IF = (imread(imageFullFileNameF));
        icmedia = ceil(mean(mean(IC)));
        IC(IC¡2) = icmedia; IC(:,end-2:end) = icmedia;
        IF(IF¡2) = icmedia; IF(:,end-2:end) = icmedia;
```

```
    [optimizer, metric] = imregconfig('multimodal');
    tform = imregtform(IC, IF, 'translation', optimizer, metric);
    IC = imwarp(IC,tform,'OutputView',imref2d(size(IF)),'FillValues',icmedia);

    IF = imfilter(IF,hm);
    IC = imfilter(IC,hm);

    Io = imadjust(IC,[0 1],[1 0]);

%       if mean(max(IF)) ¡ 160 — mean(max(Io)) ¡ 160 — std(double(max(IF))) ¿ 10
%           Io = imadjust(Io);
        IBW = imbinarize(Io,.6);
%       else
%           Io = imadjust(imadd(Io,IF));
%           IBW = imbinarize(Io,'adaptive','Sensitivity',0.35);
%       end
    IBW = imdilate(IBW,se1);
    IBW(1:6,:) = 0; IBW(end-5:end,:) = 0; IBW(:,end-5:end) = 0; IBW(:,1:5) = 0;

    S0 = regionprops(IBW, 'area', 'centroid');
    cc = bwconncomp(IBW);
    idx = find(.4*BacteriaArea ¡ [S0.Area] & [S0.Area] ¡ 40*BacteriaArea);
    IBW2 = ismember(labelmatrix(cc), idx);

    IFM = immultiply(IBW2,IF);
    ICM = immultiply(IBW2,IC);
    IMABSDIF = imabsdiff(ICM,IFM);
%       IMABSDIF = IFM; % Overwrites when there's no difference.

    % % % Data for each bacteria
    H = str2num(nameFolds1(vid,8:9));
    SN = str2num(nameFolds1(vid,2));
    IN = str2num(nameFolds1(vid,5:6));
    GN = str2num(nameFolds(karp,2:3));
    TN = str2num(nameFolds(karp,5:6));
    Indol = 0;%isletter(nameFolds(karp,10));
    cc1 = bwconncomp(IBW2);
    for m = 1:cc1.NumObjects
        Bacteria = IMABSDIF(cc1.PixelIdxList–m˝);
        AreaB = length(Bacteria);
        Bacterias = round(AreaB/BacteriaArea);
        if ˜isempty(Bacteria(Bacteria¿ThresD & Bacteria ¡ ThresF))
            MDF = mean(Bacteria(Bacteria¿ThresD & Bacteria ¡ ThresF));
        else
            MDF = 0;
        end
        SDF = sum(Bacteria(Bacteria¿ThresD & Bacteria ¡ ThresF));
        PMDF = sum(Bacteria ¿ ThresD & Bacteria ¡ ThresF);
```

```
        if ~isempty(Bacteria(Bacteria ¿ ThresF))
            MFF = mean(Bacteria(Bacteria ¿ ThresF));
        else
            MFF = 0;
        end
        PMFF = sum(Bacteria ¿ ThresF);
        SFF = sum(Bacteria(Bacteria ¿ ThresF));

        DataFile.data(contCnt,1:14)= [H, AreaB, Bacterias, MDF , SDF , PMDF , MFF , SFF , PMFF , SN , IN, GN, TN, Indol];
        contCnt=contCnt+1;
    end

    if 0% Image to corroborate 1 display image
        S1 = regionprops(IMABSDIF ¿ ThresD & IMABSDIF ¡ ThresF, IMABSDIF, 'centroid', 'area', 'perimeter', 'MaxIntensity', 'MeanIntensity');
        S2 = regionprops(IMABSDIF ¿ ThresF, IF, 'centroid', 'area', 'perimeter', 'MaxIntensity', 'MeanIntensity');
        CS1 = cat(1, S1.Centroid);
        CS2 = cat(1, S2.Centroid);
        h1 = figure(1);
        subplot('Position',[0.02, 0.5, 0.96, 0.45]);
        imshowpair(IF,IMABSDIF,'montage');
        title(['Granule(red) and Dispersed F (blue) Image',...
            strtrim(nameFolds(karp,:)),' ',nameFolds1(vid,1:end-6)]);
        hold on;
        if ~isempty(CS1)
            sizecs1 = size(CS1);
            for k = 1:sizecs1(1)
                plot(CS1(k,1), CS1(k,2) , 'b+');
    %           text(S1(k).Centroid(1)+10, S1(k).Centroid(2), num2str(S1(k).Area), 'color', 'b');
            end
        end
        if ~isempty(CS2)
            sizecs2 = size(CS2);
            for j = 1:sizecs2(1)
                plot(CS2(j,1), CS2(j,2) ,'ro');
    %           text(CS2(j,:), 'color', 'r');
            end
        end
        subplot('Position',[0.02, 0.02, 0.96, 0.45]);
        imshowpair(IC,IFM,'montage');
        % Saves resulting image with centroids and brightness
        saveas(gcf,strcat(pwd,res,'“',strtrim(nameFolds(karp,:)),nameFolds1(vid,1:end-5),'R.png'));
        hold off;
        clear h1;
    end
  end
%   csvwrite(strcat(pwd,'“Resultados“DATA',strtrim(nameFolds(karp,:))),M);
  toc
end
```

```
load(dirData);
csvwrite(strcat(pwd,res,'“DATA˙0.csv'),data);

% dlmwrite(strcat(pwd,res,'“DATA˙1.txt'),data,'delimiter','“t','precision',3);
%
% dlmwrite(strcat(pwd,res,'“DATA˙2.csv'),data,'delimiter','“t','precision',3);
%
% header = –'H', 'AreaBac', 'NoBac', 'MeanDisF' , 'SumDisF' , 'PixInMDF' , ...
%     'MeanFocF' , 'SumFocF' , 'PixInMFF' , 'SampleNo' , 'ImageNo' , 'Gluc' , 'Tript' , 'Indol"˝';
% header = strjoin(header, ',');
% fid = fopen(strcat(pwd,res,'“DATA.csv'),'w'); fprintf(fid,'%s“n',header); fclose(fid);
% dlmwrite(strcat(pwd,res,'“DATA.csv'),data,'-append');
```
