Valencia, Enero 2023

Study of spatiotemporal responses of bacterial cells

Author: Roser Montagud Martínez
Thesis advisors:
Guillermo Rodrigo Tárrega
Beatriz Sabater Muñoz
Christina Toft
UPV tutor:
Maria Desamparados Pascual-Ahuir Giner

D. Guillermo J. Rodrigo Tárrega, doctor en Biotecnología por la Universitat Politécnica de Valencia (UPV) y Científico Titular del Consejo Superior de Investigaciones Científicas (CSIC) en el Instituto de Biología Integrativa y de Sistemas (I2SysBio) centro mixto del CSIC y de la Universitat de València Estudio General (UVEG),

D. Beatriz Sabater Muñoz, doctora en Ciencias Biológicas por la UVEG y Científica Titular del CSIC en el Instituto de Biología Molecular y Celular de Plantas (IBMCP) Primo Yufera, centro mixto del CSIC y UPV,

Y,

D. Christina Toft, doctora en Ciencias por el Trinity College de Dublin (TCD) y Científica Titular del CSIC en el I2SysBio,

CERTIFICAN:

Que Dña. Roser Montagud Martínez, graduada en Biotecnología por la Universitat Politécnica de Valencia, ha realizado bajo su supervisión el trabajo titulado "Study of spatiotemporal responses of bacterial cells", que presenta para optar al grado de Doctora en Biotecnología por la Universitat Politécnica de Valencia.

Y para que así conste a los efectos oportunos, firman el presente certificado

Digitally signed by member: E85C35B6-B1BC-44CE-BE8A-766FB3B38483 D5B56ACC-2944-4F9E-B75A-466B535ACC19 Date: 2023.01.13 11:53:09 +01'00'

D. Guillermo J. Rodrigo Tárrega

SABATER MUÑOZ BEATRIZ - 09194430L Digitally signed by SABATER MUÑOZ BEATRIZ - 09194430L Date: 2023.01.13 10:59:10 +01'00'

D. Beatriz Sabater Muñoz

TOFT CHRISTINA - Y2068157B Digitally signed by TOFT CHRISTINA - Y2068157B Date: 2023.01.13 10:54:25 +01'00'

D. Christina Toft

En Valencia, enero 2023.

To Mario Fares,

whose memory will always be inspiring

Resumen

La biotecnología moderna se basa en la aplicación de una mezcla de herramientas experimentales y computacionales para llevar a cabo de forma dirigida la ingeniería genética. El objetivo es obtener células (re)programadas que implementen nuevas funciones o que sirvan como herramientas para el estudio de sistemas biológicos. En este contexto, el uso de bacterias en biotecnología está muy extendido. Sin embargo, la implementación de circuitos genéticos para el aprovechamiento de estos seres vivos puede verse limitada por procesos biológicos naturales; es decir, los circuitos diseñados (o naturales) pueden verse afectados por el transcurso del tiempo o por cambios en el entorno en el que crecen las bacterias. En esta tesis, nos propusimos seguir un enfoque integrador para estudiar cómo las bacterias responden en el tiempo y el espacio a los cambios genéticos y ambientales, que pueden afectar la funcionalidad de los circuitos de interés biotecnológico. Usamos *Escherichia coli* como organismo modelo, explotando una variedad de herramientas experimentales para trabajar con él. En primer lugar, estudiamos cómo los cambios ambientales y genéticos afectan la funcionalidad de un circuito genético sintético que implementa un comportamiento lógico sofisticado. Descubrimos que hay amplios rangos de concentración de entrada que el sistema puede procesar correctamente, que el circuito diseñado es bastante sensible a los efectos de la temperatura, que la expresión de pequeños ARN heterólogos es costosa para la célula y que una reorganización genética adecuada del sistema para reducir la cantidad de ADN heterólogo en la célula puede mejorar su estabilidad evolutiva. En segundo lugar, estudiamos el crecimiento bacteriano en entornos en los que existen materiales nanoestructurados. Descubrimos que las poblaciones bacterianas se

pueden controlar en gran medida mediante el uso de marcos organometálicos, ya que estos materiales nanoestructurados pueden descomponerse lentamente en medios biológicos liberando agentes antimicrobianos (metales y compuestos orgánicos, incluidos los antibióticos). Analizamos la respuesta bacteriana espaciotemporal siguiendo un enfoque experimental y teórico combinado en un entorno tan complejo y desafiante en medios líquidos y sólidos. Además de las variaciones en el rendimiento debido a cambios ambientales, también se debe considerar que esos circuitos genéticos evolucionarán con el tiempo debido a la acumulación estocástica de mutaciones. Estas mutaciones pueden dar lugar a cambios en la funcionalidad de los circuitos reguladores. Por tanto, en tercer lugar, realizamos un experimento de evolución a largo plazo para estudiar la contribución de un sistema de chaperonas de proteínas en la modulación de la estabilidad evolutiva. En los últimos años, se ha demostrado que los sistemas de chaperonas, como GroES/EL, pueden amortiguar o purgar mutaciones. Realizamos la secuenciación del genoma completo en diferentes líneas con diferentes niveles de expresión de GroEL y también medimos la tasa de crecimiento de las células al principio y al final del experimento evolutivo. Sin embargo, nuestros resultados no fueron concluyentes, por lo que se necesita más investigación para comprender completamente el papel de GroES/EL en la evolución y evaluar su utilidad potencial en biotecnología. En conjunto, esta tesis intenta avanzar en nuestro conocimiento sobre cómo las bacterias, y *E. coli* en particular, se comportan como se espera cuando el entorno se altera, la fisiología cambia y pasa mucho tiempo, para posibles aplicaciones industriales o (pre)clínicas.

Resum

La biotecnologia moderna es basa en l'aplicació d'una mescla d'eines experimentals i computacionals per a realitzar de forma dirigida l'enginyeria genètica. L'objectiu és obtindre cèl·lules (re)programades que implementen noves funcions o que servisquen com a eines per a l'estudi de sistemes biològics. En aquest context, l'ús de bacteris en biotecnologia està molt estés. No obstant això, la implementació de circuits genètics per a l'aprofitament d'aquests éssers vius pot veure's limitada per processos biològics naturals; és a dir, els circuits dissenyats (o naturals) poden veure's afectats pel transcurs del temps o per canvis en l'entorn en el qual creixen els bacteris. En aquesta tesi, ens vam proposar seguir un enfocament integrador per a estudiar com els bacteris responen en el temps i l'espai als canvis genètics i ambientals, que poden afectar la funcionalitat dels circuits d'interés biotecnològic. Usem *Escherichia coli* com a organisme model, explotant una varietat d'eines experimentals per a treballar amb ell. En primer lloc, estudiem com els canvis ambientals i genètics afecten la funcionalitat d'un circuit genètic sintètic que implementa un comportament lògic sofisticat. Descobrim que hi ha amplis rangs de concentració d'entrada que el sistema pot processar correctament, que el circuit dissenyat és bastant sensible a l'efecte de la temperatura, que l'expressió de xicotets ARN heteròlegs és costosa per a la cèl·lula i que una reorganització genètica adequada del sistema per a reduir la quantitat d'ADN heteròleg en la cèl·lula pot millorar la seua estabilitat evolutiva. En segon lloc, estudiem el creixement bacterià en entorns en els quals existeixen materials nanoestructurats. Descobrim que les poblacions bacterianes es poden controlar en gran manera mitjançant l'ús de marcs organometàlics, ja que aquests materials nanoestructurats poden descompondre's

lentament en medis biològics alliberant agents antimicrobians (metalls i compostos orgànics, inclosos els antibiòtics). Analitzem la resposta bacteriana espai-temporal seguint un enfocament experimental i teòric integrador en un entorn tan complex i desafiador en mitjans líquids i sòlids. A més de les variacions en el rendiment degut a canvis ambientals, també s'ha de considerar que aqueixos circuits genètics evolucionaran amb el temps degut a l'acumulació estocàstica de mutacions. Aquestes mutacions poden donar lloc a canvis en la funcionalitat dels circuits reguladors. Per tant, en tercer lloc, realitzem un experiment d'evolució a llarg termini per a estudiar la contribució d'un sistema de *chaperones* de proteïnes en la modulació de l'estabilitat evolutiva. En els últims anys, s'ha demostrat que els sistemes de *chaperones*, com GroES/EL, poden esmorteir o purgar mutacions. Realitzem la seqüenciació del genoma complet en diferents línies amb diferents nivells d'expressió de GroEL i també mesurem la taxa de creixement de les cèl·lules al principi i al final de l'experiment evolutiu. No obstant això, els nostres resultats no van ser concloents, per la qual cosa es necessita més investigació per a comprendre completament el paper de GroES/L en l'evolució i avaluar la seua utilitat potencial en biotecnologia. En conjunt, aquesta tesi intenta avançar en el nostre coneixement sobre com els bacteris, i *E. coli* en particular, es comporten com s'espera quan l'entorn s'altera, la fisiologia canvia i passa molt temps, per a possibles aplicacions industrials o (pre)clíniques.

Summary

Modern biotechnology is based on applying a mix of experimental and computational tools to perform in a directed way genetic engineering. The aim is to obtain (re)programmed cells that implement new functions or that serve as tools for the study of biological systems. In this context, the use of bacteria in biotechnology is widespread. However, the implementation of genetic circuits for the use of these living beings may be limited due to natural biological processes; that is, the engineered (or natural) circuits may be affected by the course of time or by changes in the environment in which bacteria grow. In this thesis, we proposed to follow an integrative approach to study how bacteria respond in time and space to genetic and environmental changes, which may affect the functionality of the circuits of biotechnological interest. We used *Escherichia coli* as a model organism, exploiting a variety of experimental tools to work with it. Firstly, we studied how environmental and genetic changes affect the functionality of a synthetic genetic circuit that implements a sophisticated logic behavior. We found that there are wide input concentration ranges that the system can correctly process, that the engineered circuitry is quite sensitive to temperature effects, that the expression of heterologous small RNAs is costly for the cell, and that a proper genetic reorganization of the system to reduce the amount of heterologous DNA in the cell can improve its evolutionary stability. Secondly, we studied of bacterial growth in environments in which there are nanostructured materials. We found that bacterial populations can be greatly controlled through the use of metal-organic frameworks, as these nanostructured materials can slowly decompose in biological media releasing antimicrobials (metals and organic

compounds, including antibiotics). We analyzed the spatiotemporal bacterial response following a combined experimental and theoretical approach in a such a complex and challenging environment in both liquid and solid media. In addition to variations in performance due to environmental changes, it must also be considered that those gene circuits will evolve over time due to the stochastic accumulation of mutations. These mutations can lead to changes in the functionality of the regulatory circuits. Then thirdly, we performed an experiment of long-term evolution to study the contribution of a protein chaperone system in modulating evolutionary stability. In recent years, it has been shown that chaperone systems, such as GroES/EL, can buffer or purge mutations. We performed whole-genome sequencing over different lines with varying expression levels of GroEL, and also measured the growth rate of the cells at the beginning and the end of the evolutionary experiment. However, our results were not conclusive, so further research is needed to fully understand the role of GroES/EL in evolution and to assess its potential utility in biotechnology. Taken together, this thesis tries to advance our knowledge on how bacteria, and *E. coli* in particular, behave as expected when the environment is perturbed, the physiology changes, and long time passes, for potential industrial or (pre)clinical applications.

Table of contents

Introduction

Bacteria are single-celled microorganisms that are found in almost every environment on Earth. They play important roles in various ecological processes and have the ability to adapt and respond to a variety of conditions (*i.e.*, they are extremely plastic)(1). One important aspect relates to the ability of bacteria to detect and respond to various signals in a spatiotemporal manner. Spatiotemporal responses refer to the ability to face signals that change in space (*i.e.*, response in a specific location) and time (*i.e.*, response that can be dynamic and long-term effects). According to the environment, bacteria optimize their growth and survival by exploiting a variety of genetic mechanisms that are able to process external and internal signals. By understanding these mechanisms, scientists can better understand how bacteria function and how they can be controlled or manipulated in various settings.

The use of bacteria, particularly *Escherichia coli* (*E. coli*), has also been important in many foundational studies, making it the default organism to work with at the onset of the molecular biology revolution in the 1950s. As a result, the most fundamental aspects of life, including the genetic code, transcription, translation, and replication, were first studied in this organism (2–4).

Since then, *E. coli* has become not only a model organism but also has been widely used as a workhorse instrument in many circumstances and laboratories. Some of its characteristics have made it the perfect organism to use as a host for recombinant protein production, including the broad knowledge about its genome and the ease of quickly transforming bacterial cells with exogenous deoxyribonucleic acid (DNA). From the first examples

of *in vitro* plasmid construction and transformation into *E. coli* (5) to the ability to transform the bacterial cell's chromosome with a polymerase chain reaction (PCR) product and the phage λ Red recombinase (6), and of course, more recently genomic modifications with the Clustered Regularly Interspaced Short Palindromic Repeats - Cas (CRISPR-Cas) technologies (7, 8) as some examples, virtually any exogenous DNA can be introduced into *E. coli*. Other advantages that the use of *E. coli* presents are its short propagation time to high cell densities, plus the ease of culture in affordable media (9, 10). Moreover, *E. coli* has been largely capable of producing a wide variety of different types of proteins, being insulin a prime example of the production of proteins of interest in *E. coli* (11); moreover, nearly one-third of currently approved recombinant therapeutic proteins are currently produced in *E. coli* (12).

While Systems biology seeks a better understanding of how the molecular components within a biological entity (*e.g.*, a cell, tissue, or organ) connect to its physiological functions and phenotypes through quantitative reasoning, computational models, and high-throughput experimental technologies (13); Synthetic biology aims at designing and engineering genetic regulatory circuits to implement novel functions or discover natural behaviors (14). In this context, the ease of modification and plasticity that *E. coli* offers is essentially unmatched. This bacterium has been engineered with practical applications in bioremediation, biofuel production, and even some potential clinical applications (15–17). But also, it has allowed to quantitively study libraries of promoters that drive reporter proteins, such as fluorescent proteins, allowing for an unbiased measurement of transcriptional activity (18).

E. coli was used in the first large-scale fermentations devoted to manufacturing a new class of high-value-added products like solvents, organic acids, vitamins, enzymes, and other products. In the mid-XX century, processes for the production of penicillin and streptomycin emerged from the development of antibiotic fermentations. Investigations on the scale-up of these two antibiotics implied a new field of biochemical (microbiological) engineering (19).

E. coli also has some limitations when producing a heterologous protein. Factors that might influence the successful production of the protein of interest are structural features of the gene sequence, the stability and translational efficiency of mRNA, the ease of protein folding, or degradation of the protein by host cell proteases. Differences in codon usage between the foreign gene and native *E. coli*, and the potential toxicity of the protein to the host are also important to consider (20–22), as well as *E. coli*'s inability to perform many of the posttranslational modifications found in eukaryotic proteins (9). A variety of techniques have been developed to solve these problems, including the use of different promoters to regulate the level of expression, the use of different host strains, co-expression of chaperones, reduction of culture temperature, and secretion of proteins into the periplasm or culture medium (23, 24).

However, to enhance our ability to use bacteria, and *E. coli* in particular, in the industry or in the clinic, more knowledge is required. An overlooked question is whether the resulting gene cassettes work as expected when the environment is perturbed, the physiology changes, and when a long time passes.

The fine-tuning of gene expression plays a crucial role in maintaining the proper function of cells and allows a fast response to changing environmental conditions or cellular signals, and it is especially important when implementing a synthetic circuit. Gene regulation can occur at different levels. Transcriptional regulation, where ribonucleic acid (RNA) production is controlled, can occur in many diverse manners, for example by exploiting promoters of different strengths (in other words, that produce different quantities of mRNA depending on the affinity of RNA polymerase binding), using transcription factors that bind the DNA and either enhance or inhibit the transcription of a gene(25). In turn, gene expression can further be altered by varying the availability of such transcription factors. This can be achieved, for example, by employing allosteric signals, molecules that can bind to effector molecules (*i.e.*, transcription factors) and alter their activity. Additionally, gene regulation can also take place at the translational level, when the levels of protein synthesized are controlled. Initiation of translation is often regulated by the accessibility of ribosomes to the Shine-Dalgarno sequence(26, 27). This allows for multiple-level signal integration, permitting more complex and intricate circuits to be implemented, enabling more possible phenotypes to present depending on different conditions.

In particular, the circuit design process often assumes a Boolean scenario of input activity (*i.e.*, inductor present or absent in the medium)(28, 29). However, for a comprehensive understanding of the regulation mechanisms, the design circuits must be analyzed in analog terms to recognize the ability of the system to work with ranges of concentrations and therefore damp eventual disturbances on the input signals (30). This is especially relevant for circuits that operate in unsaturated conditions (*e.g.*, as a result of weak

binding constants), as is the case, for example, in riboregulatory systems (31). It is also not fully known to what extent the design principles exploited to design synthetic genetic circuits (14) are not influenced by more complex processes (linked to the physicochemical properties of different biological species and the intricate regulatory circuitry of the cell)(32), although efforts are being made in this direction. For example, RNA interactions (intramolecular or intermolecular) are very sensitive to changes in temperature, pH, osmolarity, or metal ions (33, 34). This makes the RNA circuits eventually sensitive to changes in the environment. On the one hand, because these changes can affect the expression of the RNAs and, on the other hand, because they can affect the way they interact. However, this sensitivity can be used to design biosensors, as is the case of synthetic thermometers based on riboswitches (35).

Moreover, bacteria often live in complex and challenging environments, even for controlled applications, not only those found in natural ecosystems or in animals and plants. In these environments, bacteria are exposed to a variety of inhibitory factors coming from various sources that can affect their growth and survival. These environments can be natural or synthetic. For example, recent work has shown how bacteria cohabit with nanomaterials and other inert but hi-tech structures(36, 37). It is then important to study the response in time and space of the bacteria, either engineered with a synthetic gene circuit or being a natural strain, to gain knowledge on their robustness, susceptibility, and survival.

Some examples of factors that can be manipulated in order to control bacteria proliferation include changing specific conditions in which they have evolved and, therefore, in which they are

adapted to live (38). For example, factors such as temperature, pH, or osmotic pressure directly condition bacterial growth. One can also control bacterial growth by manipulating the availability of resources such as nutrients, oxygen or other gases (*e.g.*, N_2, CO_2). Furthermore, bacterial proliferation can also be inhibited with the addition of substances that become toxic for the cells; for example, the widely-spread use of antibiotics or antiseptics; some heavy metals such as copper, mercury, or lead(39, 40); and other chemicals, can be used to hinder the bacteria's ability to grow (41). Even some peptides could be employed as antibiotic agents as an alternative to conventional antibiotics(42).

In addition to variable performance depending on the environment, designed circuits will also evolve over time (*i.e.*, they are susceptible to acquiring mutations), eventually leading to changes in functionality. These changes can, in some cases, completely abolish their expected function (43), especially when the heterologous expression causes a high burden in the cell (44). For this, it is crucial to evaluate the evolutionary stability of the circuits. Some work has been done to study the stability of circuits based on the regulation of transcription (45, 46), but little is known about how circuits in which RNA plays a fundamental role evolve.

Bacteria can evolve rapidly due to their large population sizes and short generation times. This allows them to adapt to new environments, develop resistance to antibiotics and other drugs, and evolve new traits and capabilities. The study of bacterial evolution can be accomplished through a variety of methods, helping understand the mechanisms of evolution, how adaptation to new environments happens, and the potential impacts of these processes on humans and other aspects of the environment.

Long term evolution experiments (LTEE) have made it feasible to observe the effects of natural selection on mutations that have arisen *de novo* during the course of an experiment. In LTEE a number of individuals evolve in a controlled environment, allowing them to reproduce and evolve over multiple generations. Furthermore, the ability to store bacteria in a non-evolving state and to maintain a strictly clonal system of propagation enables the direct estimation of the biological fitness in a particular environment relative to that of its ancestor (47). Through LTEE, valuable insights into the processes of evolution and the ways in which bacteria adapt and change over time have been established. It has also contributed to our understanding of the mechanisms of antibiotic resistance and the potential consequences of exposing bacteria to selective pressures (48–50). Although extensive efforts have been made to study how evolution works, there are still many gray areas, as most studies have relied on natural evolution, we do not completely understand how an engineered bacterium might evolve with time and change its desired function.

Chaperones, also known as heat shock proteins, could play an important role in keeping circuits stable within cells. Chaperones such as GroES/EL are essential proteins of organisms that are responsible for the correct folding of proteins in the cell, as well as preventing the formation of toxic protein aggregates. This provides cells with mutational robustness by masking mutations that otherwise would have deleterious effects (51–53). They are essential under a wide range of physiological and stress conditions, showing differences in the basal expression level depending on the organism. Likewise, due to their phenotypic modulation capacity, they become enablers of phenotypic variability in a population since they allow a large number of genetic variants to be maintained

(54–56). The overexpression of these chaperones, in particular GroEL, enables the persistence of cells under strong bottleneck dynamics (implying strong gene drift and high accumulation of mutations), such as that observed in certain insect endosymbiont bacteria. It has been shown that the overexpression of GroEL affects the system; its client proteins (those that need it for their correct folding) have a higher mutation rate than those that are not clients (53, 57–59). However, the role that GroES/EL plays in channeling mutations and the possible new functions that may emerge from these nucleotide sequence changes remain ambiguous. Chaperones act by "smoothing" adaptive landscapes, which could favor the maintenance of synthetic circuits, especially those that entail a burden for the cells in the bacterial population.

All in all, by enlarging our knowledge about how bacteria deploy spatiotemporal responses in terms of regulation of gene expression, survival in challenging environments, and evolutionary stability, we expect to boost to the application of these microorganisms as biological tools to fulfill important biotechnological demands in the XXI century.

References

1. Madigan MT, Bender KS, Buckley DH, Sattley WM, Stahl DA. *Brock Biology of Microorganisms.* 15th Edition. Pearson Education, Inc.; 2019.

2. Crick FHC, Barnett L, Brenner S, Watts-Tobin RJ. General Nature of the Genetic Code for Proteins. *Nature.* 1961;192(4809):1227-1232.

3. Nirenberg M, Leder P, Bernfield M, et al. RNA codewords and protein synthesis, VII. On the general nature of the RNA code. *Proc Natl Acad Sci U S A.* 1965;53(5):1161-1168.

4. Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol.* 1961;3(3):318-356.

5. Cohen SN, Chang ACY, Boyer HW, Helling RB. Construction of Biologically Functional Bacterial Plasmids In Vitro. *Proc Natl Acad Sci U S A.* 1973;70(11):3240-3244.

6. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A.* 2000;97(12):6640-6645.

7. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol.* 2013;31(3):233-239.

8. Zhao D, Yuan S, Xiong B, et al. Development of a fast and easy method for Escherichia coli genome editing with CRISPR/ Cas9. *Microb Cell Fact.* 2016;15:205.

9. Makrides SC. Strategies for achieving high-level expression of genes in Escherichia coli. *Microbiol Rev.* 1996;60(3):512-538.

10. Lee SY. High cell-density culture of Escherichia coli. *Trends Biotechnol.* 1996;14(3):98-105.

11. Crea R, Kraszewski A, Hirose T, Itakura K. Chemical synthesis of genes for human insulin. *Proc Natl Acad Sci U S A.* 1978;75(12):5765-5769.

12. Huang CJ, Lin H, Yang X. Industrial production of recombinant therapeutics in Escherichia coli and its recent advancements. *J Ind Microbiol Biotechnol.* 2012;39(3):383-399.

13. Tavassoly I, Goldfarb J, Iyengar R. Systems biology primer: the basic methods and approaches. *Essays Biochem.* 2018;62(4):487-500.

14. Mukherji S, van Oudenaarden A. Synthetic biology: understanding biological design from synthetic circuits. *Nat Rev Genet.* 2009;10(12):859-871.

15. Gilbert ES, Walker AW, Keasling JD. A constructed microbial consortium for biodegradation of the organophosphorus insecticide parathion. Appl Microbiol *Biotechnol.* 2003;61(1):77-81.

16. Liu T, Khosla C. Genetic Engineering of Escherichia coli for Biofuel Production. *Annu Rev Genet.* 2010;44(1):53-69.

17. Anderson JC, Clarke EJ, Arkin AP, Voigt CA. Environmentally controlled invasion of cancer cells by engineered bacteria. *J Mol Biol.* 2006;355(4):619-627.

18. Hammer K, Mijakovic I, Jensen PR. Synthetic promoter libraries - Tuning of gene expression. *Trends Biotechnol.* 2006;24(2):53-55.

19. Buchholz K, Collins J. The roots - A short history of industrial microbiology and biotechnology. *Appl Microbiol Biotechnol.* 2013;97(9):3747-3762.

20. Dong H, Nilsson L, Kurland CG. Gratuitous overexpression of genes in Escherichia coli leads to growth inhibition and ribosome destruction. *J Bacteriol.* 1995;177(6):1497-1504.

21. Doherty AJ, Connolly BA, Worrall AF. Overproduction of the toxic protein, bovine pancreatic DNasel, in Escherichia coli using a tightly controlled T7-promoter-based vector. *Gene.* 1993;136(1-2):337-340.

22. Boël G, Letso R, Neely H, et al. Codon influence on protein expression in E. coli correlates with mRNA levels. *Nature.* 2016;529(7586):358-363.

23. Choi JH, Lee SY. Secretory and extracellular production of recombinant proteins using Escherichia coli. *Appl Microbiol Biotechnol.* 2004;64(5):625-635.

24. de Marco A, Deuerling E, Mogk A, Tomoyasu T, Bukau B. Chaperone-based procedure to increase yields of soluble

recombinant proteins produced in E. coli. *BMC Biotechnol.* 2007;7(1):32.

25. Busby S, Ebright R. Promoter structure, promoter recognition, and transcription activation in prokaryotes. *Cell.* 1994;79(5):743-746.

26. Babitzke P, Baker CS, Romeo T. Regulation of translation initiation by RNA binding proteins. *Annu Rev Microbiol.* 2009;63:27-44.

27. Dolcemascolo R, Goiriz L, Montagud-Martínez R, Rodrigo G. Gene regulation by a protein translation factor at the single-cell level. Umulis DM, ed. *PLoS Comput Biol.* 2022;18(5):e1010087.

28. Regot S, MacIa J, Conde N, et al. Distributed biological computation with multicellular engineered networks. *Nature.* 2011;469(7329):207-211.

29. Nielsen AAK, Der BS, Shin J, et al. Genetic circuit design automation. *Science.* 2016;352(6281).

30. Mayo AE, Setty Y, Shavit S, Zaslaver A, Alon U. Plasticity of the cis-Regulatory Input Function of a Gene. *PLoS Biol.* 2006;4(4):e45.

31. Rodrigo G, Prakash S, Shen S, Majer E, Daròs JA, Jaramillo A. Model-based design of RNA hybridization networks implemented in living cells. *Nucleic Acids Res.* 2017;45(16):9797-9808.

32. Cardinale S, Arkin AP. Contextualizing context for synthetic biology--identifying causes of failure of synthetic biological systems. *Biotechnol J.* 2012;7(7):856-866.

33. Miner JC, Chen AA, García AE. Free-energy landscape of a hyperstable RNA tetraloop. *Proc Natl Acad Sci U S A.* 2016;113(24):6665-6670.

34. Gluick TC, Gerstner RB, Draper DE. Effects of Mg2+, K+, and H+ on an equilibrium between alternative conformations of an RNA pseudoknot. *J Mol Biol.* 1997;270(3):451-463.

35. Neupert J, Karcher D, Bock R. Design of simple synthetic RNA thermometers for temperature-controlled gene expression in Escherichia coli. *Nucleic Acids Res.* 2008;36(19):e124.

36. Aznar E, Oroval M, Pascual L, Murguía JR, Martínez-Máñez R, Sancenón F. Gated Materials for On-Command Release of Guest Molecules. *Chem Rev.* 2016;116(2):561-718.

37. Horcajada P, Gref R, Baati T, et al. Metal-organic frameworks in biomedicine. *Chem Rev.* 2012;112(2):1232-1268.

38. Monod J. The growth of bacterial cultures. *Annu Rev Microbiol.* 1949;3(1):371-394.

39. Dupont CL, Grass G, Rensing C. Copper toxicity and the origin of bacterial resistance—new insights and applications. *Metallomics.* 2011;3(11):1109.

40. Manceau A, Nagy KL, Glatzel P, Bourdineaud JP. Acute Toxicity of Divalent Mercury to Bacteria Explained by the Formation of Dicysteinate and Tetracysteinate Complexes Bound to Proteins in Escherichia coli and Bacillus subtilis. *Environ Sci Technol.* 2021;55(6):3612-3623.

41. Syed A, Zeyad MT, Shahid M, Elgorban AM, Alkhulaifi MM, Ansari IA. Heavy Metals Induced Modulations in Growth, Physiology, Cellular Viability, and Biofilm Formation of an Identified Bacterial Isolate. *ACS Omega.* 2021;6(38):25076-25088.

42. Hancock REW, Chapple DS. Peptide Antibiotics. *Antimicrob Agents Chemother.* 1999;43(6):1317-1323.

43. Canton B, Labno A, Endy D. Refinement and standardization of synthetic biological parts and devices. *Nat Biotechnol.* 2008;26(7):787-793.

44. Ceroni F, Algar R, Stan GB, Ellis T. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat Methods.* 2015;12(5):415-418.

45. Yang S, Sleight SC, Sauro HM. Rationally designed bidirectional promoter improves the evolutionary stability of synthetic genetic circuits. *Nucleic Acids Res.* 2013;41(1):e33.

46. Fernandez-Rodriguez J, Yang L, Gorochowski TE, Gordon DB, Voigt CA. Memory and Combinatorial Logic Based on DNA Inversions: Dynamics and Evolutionary Stability. *ACS Synth Biol.* 2015;4(12):1361-1372.

47. Lenski RE, Rose MR, Simpson SC, Tadler SC. Long-Term Experimental Evolution in Escherichia coli. I. Adaptation and Divergence During 2,000 Generations. *Am Nat.* 1991;138(6):1315-1341.

48. Wiser MJ, Ribeck N, Lenski RE. Long-Term Dynamics of Adaptation in Asexual Populations. *Science.* 2013;342(6164):1364-1367.

49. Lenski RE. Convergence and Divergence in a Long-Term Experiment with Bacteria. *Am Nat.* 2017;190(S1):S57-S68. doi:10.1086/691209

50. Card KJ, Thomas MD, Graves JL, Barrick JE, Lenski RE. Genomic evolution of antibiotic resistance is contingent on genetic background following a long-term experiment with Escherichia coli. *Proc Natl Acad Sci U S A.* 2021;118(5).

51. Hartl FU, Bracher A, Hayer-Hartl M. Molecular chaperones in protein folding and proteostasis. *Nature.* 2011;475(7356):324-332.

52. Tokuriki N, Tawfik DS. Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature.* 2009;459(7247):668-673.

53. Williams TA, Fares MA. The Effect of Chaperonin Buffering on Protein Evolution. *Genome Biol Evol.* 2010;2:609-619.

54. Young JC, Barral JM, Hartl FU. More than folding: Localized functions of cytosolic chaperones. *Trends Biochem Sci.* 2003;28(10):541-547.

55. Queitsch C, Sangster T a, Lindquist S. Hsp90 as a capacitor of phenotypic variation. *Nature.* 2002;417(6889):618-624.

56. Rutherford SL. Between genotype and phenotype: protein chaperones and evolvability. *Nat Rev Genet.* 2003;4(4):263-274.

57. Fares MA, Ruiz-González MX, Moya A, Elena SF, Barrio E. GroEL buffers against deleterious mutations. *Nature.* 2002;417(6887):398.

58. Sabater-Muñoz B, Prats-Escriche M, Montagud-Martínez R, et al. Fitness trade-offs determine the role of the molecular chaperonin GroEL in buffering mutations. *Mol Biol Evol.* 2015;32(10):2681-2693.

59. Bogumil D, Dagan T. Cumulative impact of chaperone-mediated folding on genome evolution. *Biochemistry.* 2012;51(50):9941-9953.

Justification and objectives

Bacteria are extremely plastic and resilient organisms, with a great capacity to adapt to the environment and to the accumulation of mutations. However, unpredictable changes can affect the stability and performance of regulatory circuits implemented on purpose for biotechnological and/or biomedical applications. Overall, by expanding our knowledge on how bacteria implement spatiotemporal responses for regulating gene expression, surviving in challenging environments, and remaining stable upon acquisition of mutations, we address important biotechnology issues. Therefore, this thesis has the following three objectives:

1) Study of how environmental and genetic changes affect the functionality of a synthetic genetic circuit.

2) Study of bacterial growth in environments in which there are nanostructured materials and toxic compounds.

3) Study of the contribution of a protein chaperone, *i.e.*, GroEL, in modulating evolutionary stability through experimental evolution.

Chapter 1: Probing the operability regime of a minimal ribocomputing cell.

This work has been published in a peer-reviewed journal. See the full citation:

Montagud-Martínez R, Ventura J, Ballesteros-Garrido R, Rosado A, Rodrigo G. Probing the operability regime of an engineered ribocomputing unit in terms of dynamic range maintenance with extracellular changes and time. *J Biol Eng.* 2020;14(1).

In this publication, I performed all the experiments except the one corresponding to Figure 8 under the supervision of GR, and performed the preparation of some materials (particularly, the I performed the cloning of the new plasmid for this article pRHA1240). GR and I analyzed the data and prepared the figures. GR designed the research.

1.1. Introduction

The design and subsequent implementation of synthetic gene circuits provides valuable information about the corresponding counterparts found in nature (1). However, the design process irremediably considers a simplified scenario for tractable purposes. In particular, the circuit design process often assumes a Boolean scenario of input activity (*i.e.*, inducer present or absent in the medium)(2,3). Yet, for a comprehensive understanding of the regulatory mechanisms, designer circuits should be analyzed in analog terms in order to recognize the ability of the system to work with ranges of concentrations and hence buffer eventual perturbations in the input signals (4). This is especially relevant for circuits that operate in non-saturated conditions (*e.g.*, as a result of weak binding constants), as it is the case *e.g.* of riboregulatory systems (5).

To what extent the design principles exploited to engineer synthetic gene circuits (1) are not influenced by more complex processes linked to the physicochemical properties of the different biological species and the intricate regulatory circuitry of the cell is also not entirely known (6). For example, intra- or intermolecular RNA interactions are quite sensitive to changes in temperature, pH, osmolarity, or metal ions (7,8). This makes RNA circuits to be eventually sensitive to changes in the medium. On the one hand, because these changes may affect the expression of the RNAs and, on the other hand, because they may affect the interaction mode between them. Notwithstanding, such sensitivity can be exploited to design biosensors, as it is the case of synthetic thermometers based on riboswitches (9).

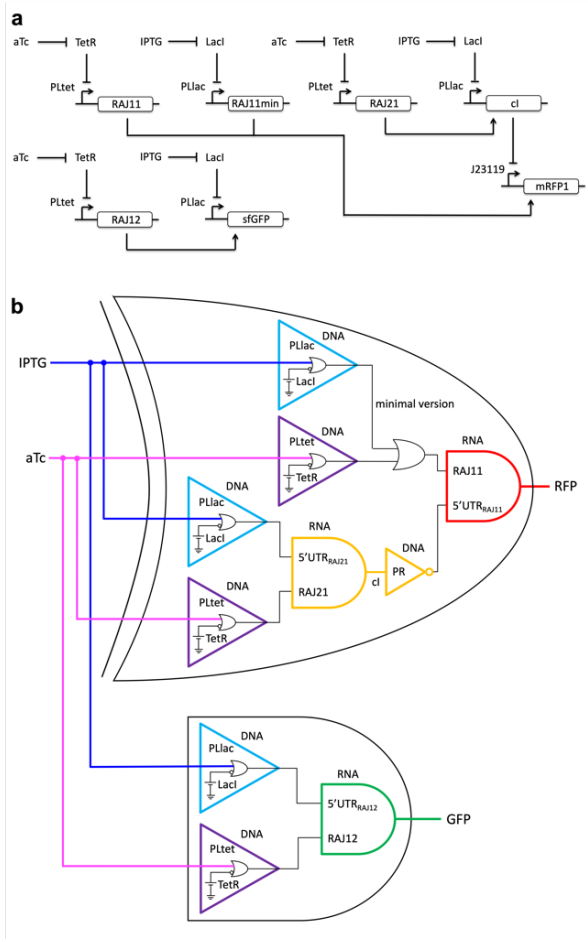

Figure 1.1 Schematics of the engineered gene circuit implementing a half adder in bacteria **a** Symbolic and **b** logic representation of the circuit. IPTG and aTc are the two molecules that work as input signals in the system (note that the PLlac or PLtet promoter implements an IMPLY gate), while the expressions of an RFP (XOR gate; from plasmid pRHA40) and a GFP (AND gate; from plasmid pRHA12) constitute the output responses

In addition to variable performance in different environmental scenarios, the engineered circuits will also evolve with the time course (*i.e.*, they will acquire mutations), eventually leading to

changes in functionality. These changes can in some cases be drastic, abolishing completely the intended function (10), especially when the heterologous expression causes high burden in the cell (11). To this end, it is important to evaluate the evolutionary stability of the circuits. Some works have been carried out to study the stability of circuits based on transcription regulation (12,13), but little is known about how circuits based on regulatory RNAs evolve.

In this work, we considered a recently engineered gene circuit working like a half adder (the combination of XOR and AND gates) (14), which is based on three transcription factors (LacI, TetR, and cI) and four riboregulators (RAJ11, RAJ11min, RAJ12, and RAJ21), illustrated in **Fig. 1.1**. In the following, we present experimental results that served us to critically analyze the operability regime of this synthetic circuit, which is of importance on the light of the considerable growth of RNA synthetic biology in recent years (15).

1.2. Results

1.2.1. Analog behavior of the system

We analyzed the double output response of the RNA circuit for a two-dimensional concentration gradient of its inducers, isopropyl β-D-1-thiogalactopyranoside (IPTG), ranging from 0 to 1000 µM, and anhydrotetracycline (aTc), from 0 to 100 ng/µL. For that, we monitored the expression of the monomeric red fluorescent protein (mRFP1) and the superfolder green fluorescent protein (sfGFP) for each induction condition (**Fig. 1.2**).

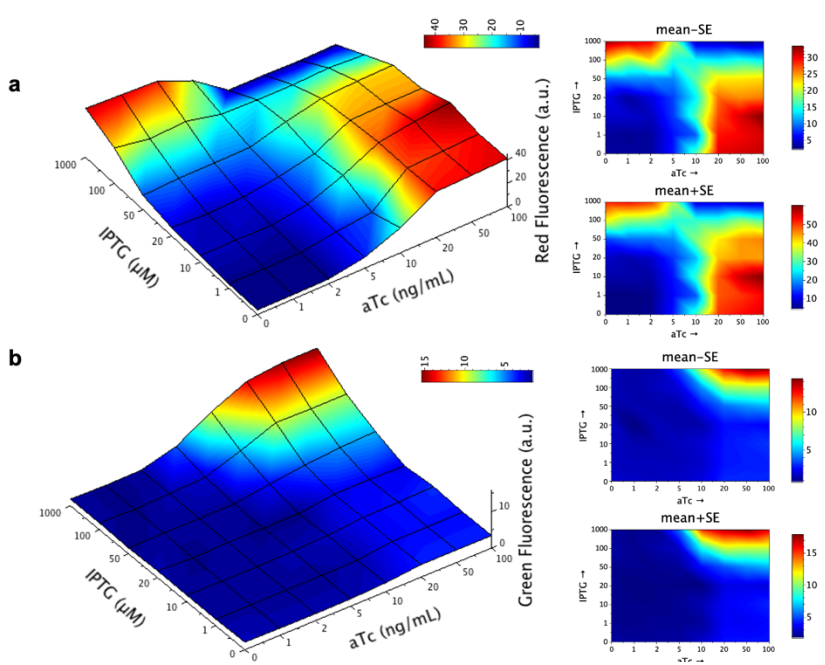

Figure 1.2 Dynamic range of the engineered gene circuit (E. coli co-transformed with plasmids pRHA40 and pRHA12). Fluorescence-based surface response with respect to IPTG and aTc (56 induction conditions tested); mean of four biological replicates. **a** Mean normalized red fluorescence (AND gate). **b** Mean normalized green fluorescence (XOR gate). 3D grid constructed from experimental data, colored by interpolation. On the right, 2D plots showing mean plus/minus standard error

We observed that the state transitions agree in both output channels (at ~100 μM IPTG and at ~10 ng/mL aTc). In essence, this is because synthetic riboregulators act in a linear regime [*i.e.*, protein expression scales with the product between the messenger RNA (mRNA) and small RNA (sRNA) concentrations]. Then, the non-linearity of the response comes from the transcriptional regulation with engineered promoters (16), which are basically the same for the XOR and the AND gates (despite some modifications in the operators). This entails a separation of input variables to describe the output response (17). Indeed, sfGFP expression can

be approached as proportional to the product of the dose-response curves for the *lac* and *tet* promoters, sfGFP ~ $f_{lac} \cdot f_{tet}$, whilst mRFP1 expression as proportional to the sum of these two dose-response curves times the dose-response curve for the λ promoter, mRFP1 ~ $f_{\lambda} \cdot (f_{lac} + f_{tet})$, as shown in **Fig. 1.3**.

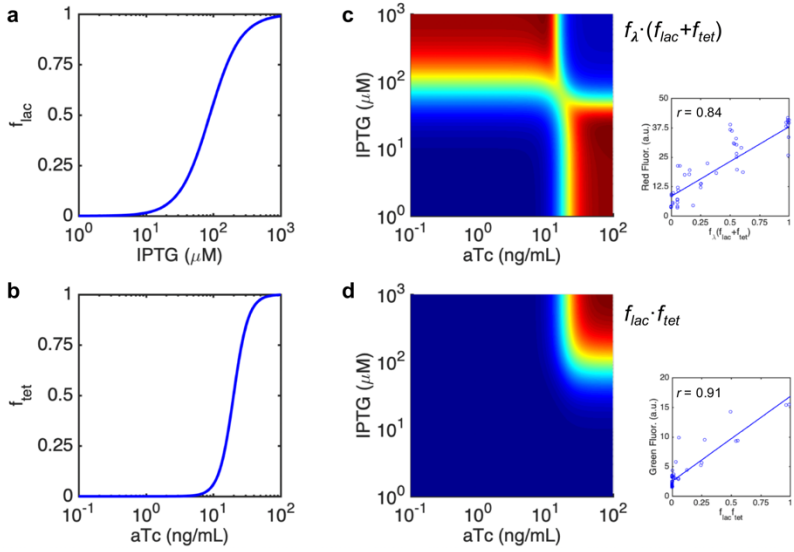

Figure 1.3 Output response prediction from individual promoter characterizations. **a, b** Dose-response curves of the lac and tet promoters with respect to IPTG and aTc. **c, d** Prediction of the surface responses of the engineered gene circuit by considering separation of variables. The subplots show the correlation between the predicted values and mean normalized fluorescence data.

Moreover, this scheme entails that the system is able to correctly process two concentration ranges instead of two concentration values. For instance, values of IPTG up to 20 μM would correspond to zero in binary code or values of aTc from 20 to 100 ng/mL to one in binary code too. Basically, this allows being tolerant to fluctuations occurring at the input level from uncontrollable upstream processes.

1.2.2. Environmental robustness of the system

To study how different environmental conditions affect the performance of the RNA circuit, we decided to vary temperature and pH. We also studied the effect of high amounts of Fe^{3+} and urea in the medium. Upon characterizing the circuit in all the environments for each induction condition (**Fig. 1.4**), we first observed certain robustness to fluctuations in pH in both output channels. We only noted certain impact on the ability of cI to transcriptionally repress mRFP1. The addition of urea neither left a significant imprint, but Fe^{3+} caused a reduction in red

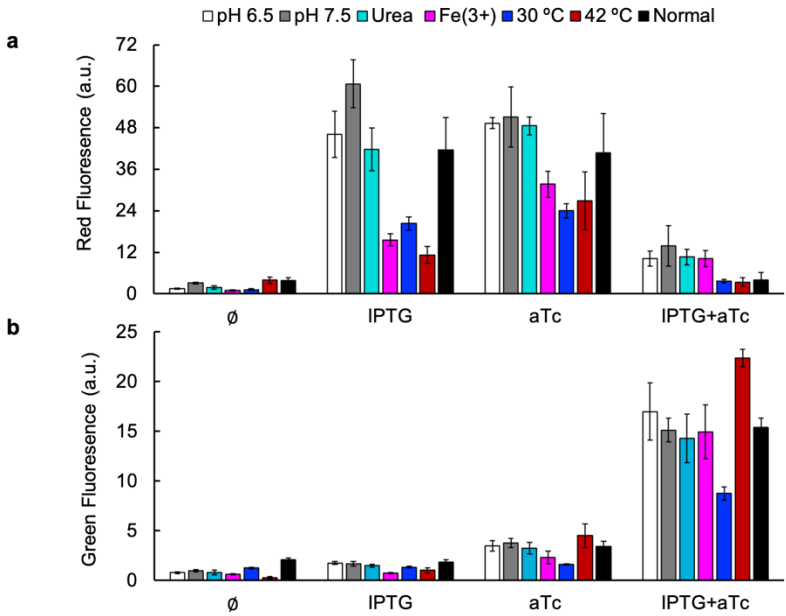

Figure 1.4 Environmental robustness assessment of the engineered gene circuit (E. coli co-transformed with plasmids pRHA40 and pRHA12). Fluorescence monitoring for each induction condition (IPTG, aTc) when the culture medium changes (pH variation, urea addition, Fe3+ addition, and temperature variation). **a** Normalized red fluorescence. **b** Normalized green fluorescence. Error bars correspond to standard errors over four biological replicates

fluorescence upon induction with IPTG or aTc, although not in green fluorescence. Arguably, high amounts of Fe^{3+} might quench the fluorescence from mRFP1 (18). Conversely, we found both output channels sensitive to changes in temperature. We hypothesized that the higher temperature, the lower the ability of the structured 5' untranslated region (UTR) to exert the *cis*-repression (9). For mRFP1, nonetheless, the effect is not as straightforward as for sfGFP. On the one hand, the destabilization of the 5' UTR of mRFP1 contributes to increase the output signal; but, on the other hand, the destabilization of the 5' UTR of cI to reduce it. On the light of our results, a leakage cI expression at high temperature can explain the reduction in mRFP1 expression upon induction with IPTG or aTc.

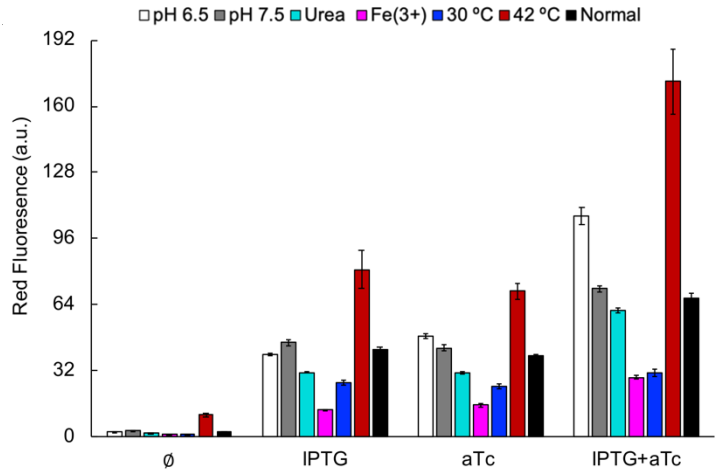

Figure 1.5 Environmental robustness assessment of a control system implementing an OR gate (E. coli transformed with plasmid pRHA37; this system lacks the cI repressor so that mRFP1 is only regulated post-transcriptionally). Fluorescence monitoring (normalized red fluorescence) for each induction condition (IPTG, aTc) when the culture medium changes (pH variation, urea addition, Fe^{3+} addition, and temperature variation; same legend as in Fig. 4). Error bars correspond to standard errors over four biological replicates.

To confirm the proposed mode of action of temperature, we considered a system in which there is no cI-mediated repression of mRFP1 (note that this circuit works like an OR gate) (14). Fluorescence readouts now revealed the monotonous increase of mRFP1 expression with temperature in all induction conditions (**Fig. 1.5**). We also observed a fluorescence boost with both inducers at low pH, but we attributed this to a cell growth defect of this system in this condition.

1.2.3. Genetic burden of the system

To assess the impact of the synthetic circuit on the chassis cell, we quantified the growth rate in exponential phase for each induction condition (19). As expected, we found that as long as the heterologous species were expressed (sRNAs and proteins), cells grew slower (**Fig. 1.6a**). In fact, cells grew on average a 34% slower with both inducers than without them. But we know that *i*) in absence of inducers, no sRNAs or proteins are expressed from the system; *ii*) when IPTG is present, the system only expresses one sRNA (RAJ11min) and one protein (mRFP1); *iii*) when aTc is in the medium, the system expresses three sRNAs (RAJ11, RAJ12, and RAJ21) and one protein (mRFP1); and *iv*) in the presence of both inducers, there are four sRNAs (RAJ11, RAJ11min, RAJ12, and RAJ21) and two proteins (sfGFP and cI) expressed. A statistical analysis by grouping the conditions as non-induced, induction with IPTG alone, and induction with aTc (either alone or in conjunction with IPTG) resulted in the more significant difference (one-way ANOVA, df = 2, $P = 3 \cdot 10^{-5}$); more than controlling for the number of expressed proteins.

In addition, we quantified the time required for the cells to enter in exponential phase from diluted cultures (lag time) for each induction condition (19). We found that the lag time increased on average about 1 h when there was only one inducer in the medium, whilst it was about 1.5 h when the two inducers were present (**Fig. 1.6b**). Yet, statistical analyses only revealed a significant difference when controlling for induction *vs.* non-induction (one-way ANOVA, df = 1, $P = 0.026$). Taken together, our results revealed that the expression of heterologous sRNAs can be as costly for the

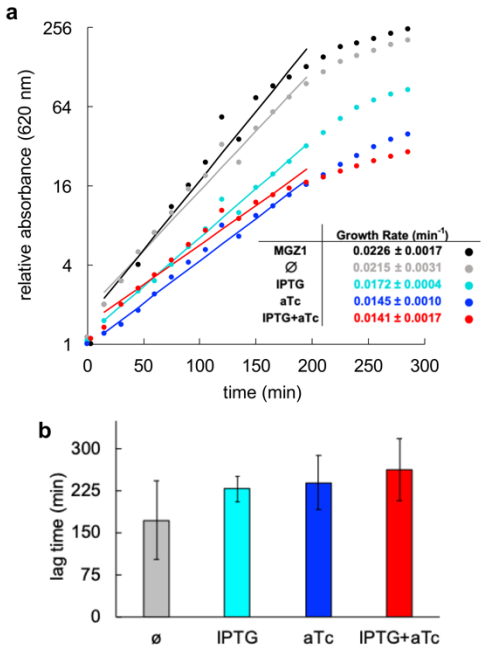

Figure 1.6 Genetic burden of the engineered gene circuit (E. coli co-transformed with plasmids pRHA40 and pRHA12). **a** Relative absorbance with time for each induction condition (IPTG, aTc; data for a representative replicate). Lines correspond to exponential fittings. The inset shows the growth rate values (mean and standard deviation) from four biological replicates. **b** Lag time before exponential growth for each induction condition (IPTG, aTc) from diluted overnight cultures. Error bars correspond to standard deviations from four biological replicates

cell as the expression of heterologous proteins. This challenges a model in which ribosome allocation mainly limits cell growth rate (11).

1.2.4. Evolutionary stability of the system

To evaluate the evolutionary stability of the system (*i.e.*, its ability to behave as designed with time), we set up an experiment of serial dilutions by carrying in parallel different bacterial populations (20). Every day, circuit functionality was assessed by fluorometry. Surprisingly, we found that mRFP1 expression was lost in all lines in just one passage (corresponding to 6.64 generations), whilst sfGFP was able to hold for ten passages (**Fig. 1.7**). We attributed these courses to the fact that mRFP1 was expressed from a very high-copy number plasmid (pUC ori), whilst the plasmid from which sfGFP was expressed presents a more moderate copy number (mutated pSC101 ori). Sequencing results after the first passage revealed the 5' UTR controlling the expression of cI was deleted in all lines, leading to its constitutive expression and in turn the repression of mRFP1. Moreover, in line A, the quality of the chromatogram was poor in the region coding for the sRNA RAJ11, and the mRFP1 coding region was correct in all lines. By contrast, we found mutations in the sfGFP coding region in all lines, but not in the regulatory region, by sequencing at last day of the evolutionary experiment.

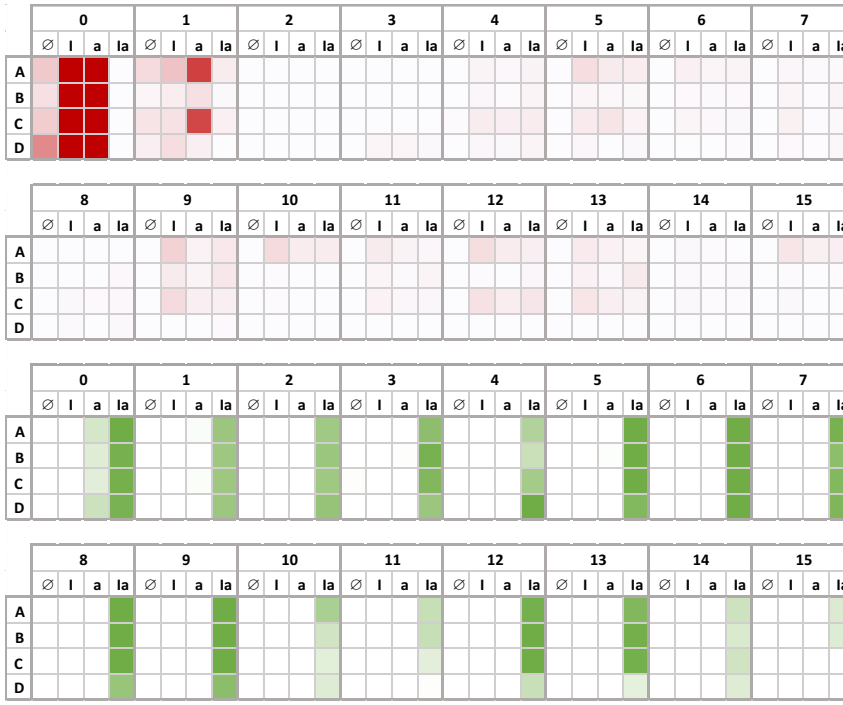

Figure 1.7 Evolutionary stability of the engineered gene circuit (E. coli co-transformed with plasmids pRHA40 and pRHA12). Continuous fluorescence monitoring of four biological replicates (A-D) over two weeks for each induction condition (I, IPTG; a, aTc). **a** Normalized red fluorescence. **b** Normalized green fluorescence

With the aim of increasing the evolutionary stability of the system, we placed the genetic cassette corresponding to the XOR gate into the plasmid carrying the AND gate. Quantitatively, we observed lower red fluorescence upon induction with IPTG or aTc and a marked leakage with the two inducers, suggesting that now the repressor cI was not sufficiently expressed as a result of a worse performance of the riboregulatory system RAJ21. Repeating again the experimental evolution, we found that this new system was able to maintain its arithmetic functionality much more time (**Fig. 1.8**).

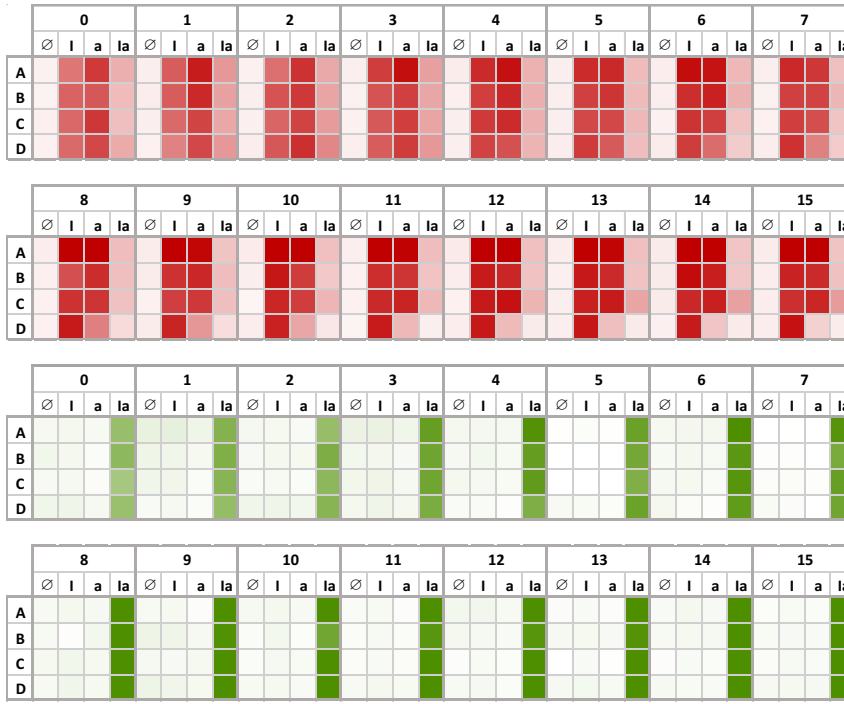

Figure 1.8 Evolutionary stability of the engineered gene circuit when it is deployed from a single plasmid (E. coli transformed with plasmid pRHA1240). Continuous fluorescence monitoring of four biological replicates (A-D) over 2 weeks for each induction condition (I, IPTG; a, aTc). **a** Normalized red fluorescence. **b** Normalized green fluorescence

In particular, mRFP1 was correctly expressed over the two weeks in three lines. Only in line D, the expression dropped after one week when inducing with aTc. Sequencing results at last day revealed a 25 bp deletion in the *tet* promoter controlling the expression of the sRNA RAJ11, which is explicative of the observed behavior. In line C, nonetheless, we could not identify a unique sequence from the chromatogram in the region coding for the sRNAs RAJ11 and RAJ21, suggesting that neutral mutations might be accumulated there. Furthermore, sfGFP was correctly expressed over the two weeks in all lines, better than before, and

sequencing results at last day confirmed the no acquisition of mutations in that cassette.

1.3. Discussion

This work contributes to better understand the operability regime of engineered circuitries based on regulatory RNAs. Notably, our results allow disclosing a series of principles in order to predict the effect of a variety of changes, and in turn develop strategies to mitigate them. First, by exploiting a pre-characterized transcriptional layer for RNA circuit engineering (3), and noting that riboregulatory performance can be correctly predicted from energetic and structural calculations through the use of RNA folding routines (21,22), the predictability of the output responses for combinatorial induction conditions is possible, as our results point out. Subsequently, we identified temperature effects as environmental variations with clear impact on circuit performance; effects that mainly appear as a consequence of using structured 5' UTRs. In this regard, a circuit implementation with sRNAs acting as repressors on non-structured 5' UTRs (21) might be an option to buffer changes in temperature. In addition, our growth curves suggest that the expression of heterologous sRNAs is costly for the cell, as it can be the expression of heterologous proteins (11). We hypothesized that the overexpression of sRNAs from high-copy number plasmids creates a high demand for the different ribonucleases that is translated into cell growth defects, atop of genome-wide transcriptomic perturbations (23). Finally, our evolutionary experiments revealed a trade-off for synthetic RNA circuits. On the one hand, these circuits require the expression from high-copy number plasmids (or from T7 promoters [24]) to deal with binding constants at the micromolar scale (5). On the other

hand, these plasmids introduce an elevated cost for the cell and then they are prone to the accumulation of deleterious mutations. Perhaps, circuits involving synthetic sRNAs with domains able to recruit endogenous RNA chaperones (25) and integrated into the chromosome would be the best strategy to enhance substantially the evolutionary stability.

1.4. Materials and Methods

1.4.1. Strain, plasmids, and reagents

E. coli Dh5α was used for cloning purposes, while *E. coli* MG1655-Z1 (*lacI*$^+$, *tetR*$^+$) was used as a chassis cell to express our genetic systems (gifted by M.B. Elowitz)(26). This was co-transformed with plasmids pRHA12 and pRHA40 (14) by electroporation, which implement a genetic half adder. This chassis cell was also transformed with plasmid pRHA37 (14) or plasmid pRHA1240. This latter plasmid was obtained by digesting pRHA40 with EcoRI and BamHI and inserting the resulting cassette into pRHA12 by digesting with EcoRI and ligating with T4 DNA ligase.

Luria-Bertani (LB) medium (L3522, Merck) was used for overnight cultures, while M9 minimal medium (1x M9 salts, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.4% glucose, 0.05% casamino acids, and 0.05% thiamine) for characterization cultures. Ampicillin and kanamycin were used as antibiotics at the concentration of 50 µg/mL. Isopropyl β-D-1-thiogalactopyranoside (IPTG) and anhydrotetracycline (aTc) were used as inducers at the concentrations of 1000 µM and 100 ng/mL, respectively, for maximal induction of the synthetic PL-based promoters (16). Concentration gradients of inducers were also considered: 0, 1, 10,

20, 50, 100, and 1000 µM IPTG and 0, 1, 2, 5, 10, 20, 50, and 100 ng/µL aTc. Compounds provided by Sigma-Aldrich.

1.4.2. Preparation of cultures for characterization

Cultures (2 mL) inoculated from single colonies (four replicates) were grown overnight in LB medium at 37 ºC and 200 rpm. Cultures were then diluted 1:100 in M9 minimal medium (2 mL) and were grown for 5-7 h at 37 ºC and 200 rpm to reach exponential phase (OD_{600} around 0.5). In the case of cultures from the experimental evolution, glycerol stocks were diluted 1:50 in M9 minimal medium in a microplate (200 µL) and were grown for 5-7 h at 37 ºC and 1000 rpm to reach exponential phase (OD_{600} around 0.5) in a plate shaker (PST-60HL, Biosan). Cultures were then diluted, in both cases, 1:50 in M9 minimal medium with appropriate inducers (IPTG, aTc).

1.4.3. Fluorescence quantification

From the prepared cultures for characterization, the microplate (96 wells, black, clear bottom; Corning) was loaded (200 µL/well). The microplate was then incubated in the plate shaker at 37 ºC (unless otherwise specified) and 1000 rpm up to 7 h (to reach an OD_{600} around 0.5-0.7). Subsequently, the microplate was assayed in a fluorometer (Varioskan Lux, Thermo Sci.) to measure absorbance (600 nm), green fluorescence (excitation: 485 nm, emission: 535 nm), and red fluorescence (excitation: 570 nm, emission: 610 nm). Mean background values of absorbance and fluorescence, corresponding to M9 minimal medium, were subtracted to correct the signals. Normalized fluorescence was calculated as the ratio of

fluorescence and absorbance. The mean value of normalized fluorescence corresponding to non-transformed cells was then subtracted to obtain a final estimate of expression.

1.4.4. Empirical model to predict the response

The individual dose-response curves of the promoters used to implement our circuit were already characterized. Promoter activity follows the sigmoidal function $f = \frac{1}{1+\left(\frac{X}{K}\right)^n}$, where X is the concentration of the regulator, K the effective dissociation constant, and n the Hill coefficient. Then, we have *i*) K = 89 µM (relative to IPTG) and n = -1.9 for the *lac* promoter (16), *ii*) K = 20 ng/mL (relative to aTc) and n = -4 for the *tet* promoter (16), and *iii*) K = 55 nM (relative to cI) and n = 2.4 for the λ promoter (27). Consequently, we can write $\text{mRNA}_{\text{mRFP1}} \sim f_{\lambda}$, RAJ11min ~ f_{lac}, RAJ11 ~ f_{tet}, $\text{mRNA}_{\text{cI}} \sim f_{lac}$, RAJ21 ~ f_{tet}, $\text{mRNA}_{\text{sfGFP}} \sim f_{lac}$, and RAJ12 ~ f_{tet}. Of note, f_{λ} also depends on IPTG and aTc, as cI ~ $f_{lac} \cdot f_{tet}$ (for this we assumed a proportionality factor of 250 nM). Finally, it turns out that mRFP1 ~ $\text{mRNA}_{\text{mRFP1}} \cdot$(RAJ11min + RAJ11) ~ $f_{\lambda} \cdot (f_{lac} + f_{tet})$ and sfGFP ~ $\text{mRNA}_{\text{sfGFP}} \cdot$RAJ12 ~ $f_{lac} \cdot f_{tet}$.

1.4.5. Different environmental conditions

To assess the effect of temperature on the genetic half adder, the microplate was also incubated in the plate shaker at 30 or 42 ºC. To assess the effect of pH, the culture medium in the microplate was prepared to be more acid (6.5, with HCl) or more alkaline (7.5, with NaOH). Finally, to assess the effect of compounds present in the human body, that medium was supplemented with 157 ng/µL $FeCl_3$ (Fe^{3+}) or 635 ng/µL $CO(NH_2)_2$ (urea), leading to ~10 mM in both cases.

1.4.6. Growth rate curves

From the prepared cultures for characterization, the microplate was loaded (200 µL/well). The microplate was then incubated for 24 h at 37 ºC with shaking in a spectrophotometer (Multiskan FC, Thermo Sci.), measuring absorbance (620 nm) every 15 min.

1.4.7. Experimental evolution

Cultures (2 mL) inoculated from single colonies (four replicates) were grown overnight in LB medium at 37 ºC and 200 rpm. Propagation of the lines during 15 d was carried out by imposing a daily 1% bottleneck (dilution 1:100), also in LB at 37 ºC and 200 rpm. Each day, the induction condition was varied (from none, to IPTG, to aTc, to IPTG plus aTc) to simulate a changing environment. A daily fossil record of each line was generated by adding glycerol 20% to an equal volume of culture.

1.4.8. Sequencing of the genetic circuit

To sequence the cassette corresponding to the XOR gate, we used the following primers: CTTTGATAACGTCTTCGGAGGAAGC (forward, landing in the mRFP1 coding region), GCTTTTTCTTGGTTGAGGGCC (reverse, landing in the cI coding region). To sequence the cassette corresponding to the AND gate, we used the following primers: GCTTCTTCCGCTAGCATCGA (forward, landing in the backbone upstream of the insert) and ACAGCTCTTCGCCTTTACGG (reverse, landing in the sfGFP coding region). Sequencing carried out by Eurofins Genomics.

1.4.9. Statistical analysis

To compare the growth rates and lag times of the cell cultures at different induction conditions, one way analysis of variance (ANOVA) tests were performed. Different groups were generated to perform the comparisons. P values were used to assess the statistical significance.

1.5. References

1. Mukherji S, van Oudenaarden A. Synthetic biology: understanding biological design from synthetic circuits. *Nat Rev Genet*. 2009;10(12):859-871.

2. Regot S, MacIa J, Conde N, et al. Distributed biological computation with multicellular engineered networks. *Nature*. 2011;469(7329):207-211.

3. Nielsen AAK, Der BS, Shin J, et al. Genetic circuit design automation. *Science*. 2016;352(6281).

4. Mayo AE, Setty Y, Shavit S, Zaslaver A, Alon U. Plasticity of the cis-Regulatory Input Function of a Gene. *PLoS Biol*. 2006;4(4):e45.

5. Rodrigo G, Prakash S, Shen S, Majer E, Daròs JA, Jaramillo A. Model-based design of RNA hybridization networks implemented in living cells. *Nucleic Acids Res*. 2017;45(16):9797-9808.

6. Cardinale S, Arkin AP. Contextualizing context for synthetic biology--identifying causes of failure of synthetic biological systems. *Biotechnol J*. 2012;7(7):856-866.

7. Miner JC, Chen AA, García AE. Free-energy landscape of a hyperstable RNA tetraloop. *Proc Natl Acad Sci U S A*. 2016;113(24):6665-6670.

8. Gluick TC, Gerstner RB, Draper DE. Effects of Mg2+, K+, and H+ on an equilibrium between alternative conformations of an RNA pseudoknot. *J Mol Biol*. 1997;270(3):451-463.

9. Neupert J, Karcher D, Bock R. Design of simple synthetic RNA thermometers for temperature-controlled gene expression in Escherichia coli. *Nucleic Acids Res*. 2008;36(19):e124.

10. Canton B, Labno A, Endy D. Refinement and standardization of synthetic biological parts and devices. *Nat Biotechnol*. 2008;26(7):787-793.

11. Ceroni F, Algar R, Stan GB, Ellis T. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat Methods*. 2015;12(5):415-418.

12. Yang S, Sleight SC, Sauro HM. Rationally designed bidirectional promoter improves the evolutionary stability of synthetic genetic circuits. *Nucleic Acids Res*. 2013;41(1):e33.

13. Fernandez-Rodriguez J, Yang L, Gorochowski TE, Gordon DB, Voigt CA. Memory and Combinatorial Logic Based on DNA Inversions: Dynamics and Evolutionary Stability. *ACS Synth Biol*. 2015;4(12):1361-1372.

14. Rosado A, Cordero T, Rodrigo G. Binary addition in a living cell based on riboregulation. *PLoS Genet*. 2018;14(7): e1007548.

15. Qi LS, Arkin AP. A versatile framework for microbial engineering using synthetic non-coding RNAs. *Nat Rev Microbiol*. 2014;12(5):341-354.

16. Lutz R, Bujard H. Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res*. 1997;25(6):1203-1210.

17. Kaplan S, Bren A, Zaslaver A, Dekel E, Alon U. Diverse Two-Dimensional Input-Functions Control Bacterial Sugar Genes. *Mol Cell*. 2008;29(6):786.

18. Eli P, Chakrabartty A. Variants of DsRed fluorescent protein: Development of a copper sensor. *Protein Sci*. 2006;15(10):2442-2447.

19. Hall BG, Acar H, Nandipati A, Barlow M. Growth rates made easy. *Mol Biol Evol*. 2014;31(1):232-238.

20. Elena SF, Lenski RE. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat Rev Genet*. 2003;4(6):457-469.

21. Mutalik VK, Qi L, Guimaraes JC, Lucks JB, Arkin AP. Rationally designed families of orthogonal RNA regulators of translation. *Nat Chem*. 2012;8(5):447-454.

22. Rodrigo G, Landrain TE, Jaramillo A. De novo automated design of small RNA circuits for engineering synthetic riboregulation in living cells. *Proc Natl Acad Sci U S A*. 2012;109(38):15271-15276.

23. Pobre V, Arraiano CM. Next generation sequencing analysis reveals that the ribonucleases RNase II, RNase R and PNPase affect bacterial motility and biofilm formation in *E. coli*. *BMC Genomics*. 2015;16(1):72.

24. Green AA, Kim J, Ma D, Silver PA, Collins JJ, Yin P. Complex cellular logic computation using ribocomputing devices. *Nature 2017 548:7665*. 2017;548(7665):117-121.

25. Ghodasara A, Voigt CA. Balancing gene expression without library construction via a reusable sRNA pool. *Nucleic Acids Res*. 2017;45(13):8116-8127.

26. Cox RS, Surette MG, Elowitz MB. Programming gene expression with combinatorial promoters. *Mol Syst Biol*. 2007;3(1):145.

27. Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB. Gene regulation at the single-cell level. *Science*. 2005;307(5717):1962-1965.

Chapter 2: Bacterial population control with macroscopic HKUST crystals.

This work has been published in a peer-reviewed journal. See full citation:

*Ballesteros-Garrido R, *Montagud-Martínez R, Rodrigo G. Bacterial Population Control with Macroscopic HKUST Crystals. CS *ACS Appl Mater Interfaces*. 2019;11(22):19878-19883.

*RBG and RMM contributed equally to this work. RBG, who was already a PhD in chemistry when preparing this manuscript, prepared the HKUST crystals and I did all the experiments involving *E. coli*. All authors contributed equally to create the figures.

2.1. Introduction

Bacterial infections represent one of the main causes of death at a global scale, especially in children and in low-income countries (1). Thus, the ability to control bacterial populations is an important challenge that today can involve different disciplines; for example, to warrant a biosafe food and water supply or to avoid the spread within the body once an infection focus has appeared. In a modern era in which the continuous emergence of antibiotic resistances is a threat (2) and in which the progress of effective treatments (in biomedicine and beyond) demands increasing precision (3), novel repression ways are certainly needed to fight bacteria.

Consequently, the possibility of controlling bacterial proliferation by using a suitable nanostructured material is appealing (4). Indeed, self-assembled nanostructures have emerged in recent years as new efficient and programmable elements for the distribution of active biochemical compounds (drugs) in different biological contexts (5,6), going beyond the traditional applications of catalysis and gas storage (7). In part, this has allowed addressing issues associated with efficacy, toxicity, and stability of such drugs. Moreover, some nanostructures have been shown to have antibacterial activity, especially those involving metals (8,9). However, most of the developments are focused on the nanoscale (particles of nm). This may limit the applicability of the technology because those materials are difficult to control once disseminated (10).

In this work, we decided to investigate the growth of bacterial populations in environments with nanostructured materials of macroscopic size (particles of mm). This type of studies in the

nanobiological interface may offer new horizons (11), but they remain largely unexplored. Potential advantages of these materials to control microorganisms are *i*) a slow release kinetics thanks to a smaller surface per volume ratio, following the application of the Fick's laws to such systems (12); *ii*) the ability to interact with the cells of the medium and experience phenomena such as gravity and orientation; or *iii*) the ability to manually locate materials in a controlled way to then remove them (and even reuse them with new loads) (13).

Among the different materials, metal-organic frameworks (MOFs)(14) have achieved great relevance in biomedicine over the last years (6). In particular, we aimed our focus at controlling a population of *Escherichia coli* cells in lab conditions by employing Hong Kong University of Science and Technology-1 (HKUST-1) materials, or simply HKUST, which are Cu_3BTC_2 crystals (15) (**Fig. 2.1**). We chose this material because it can be prepared to

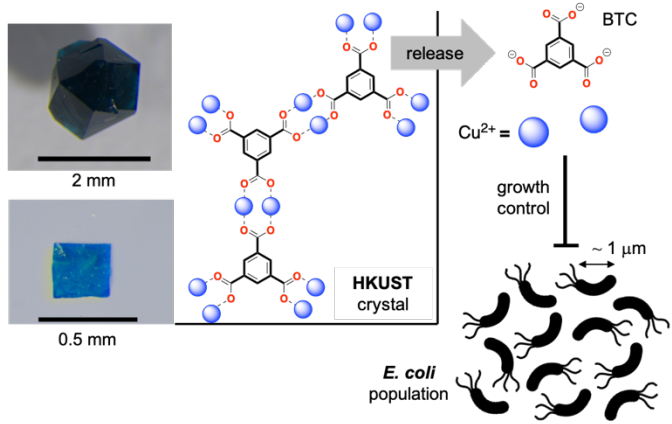

Figure 2.1 Macroscopic HKUST crystals can control the growth of a bacterial population (*e.g.*, of *E. coli*). Copper release (in the form of Cu^{2+}) due to partial crystal degradation is instrumental. Depending on the reaction conditions, the size of the observable crystal can greatly vary (from 0.2 to 2 mm edge). BTC stands for benzene-1,3,5-tricarboxylic acid.

create mm-sized crystals (16), and we considered *E. coli* as a model of Gram-negative bacteria (note that many pathogenic bacteria are of this type [17]). Copper, the metal in HKUST-1, is well-known to have antimicrobial activity (what is called oligodynamic effect)(18). MOFs have been reported to experience water and moisture sensitivity (19,20), but to what extent can this be profitable? In the following, we present results that show that HKUST-1 crystals of observable size are suitable elements for controlled copper release to the medium (in the form of Cu^{2+}), at the same time that they can serve as containers for other deliverable molecules (*e.g.*, antibiotics), which effectively leads to bacterial repression.

2.2. Results

We observed that macroscopic HKUST-1 crystals release copper in saline medium (in particular, LB medium), but not in water (**Fig. 2.2a**); arguably, due to electrostatic forces during the competition

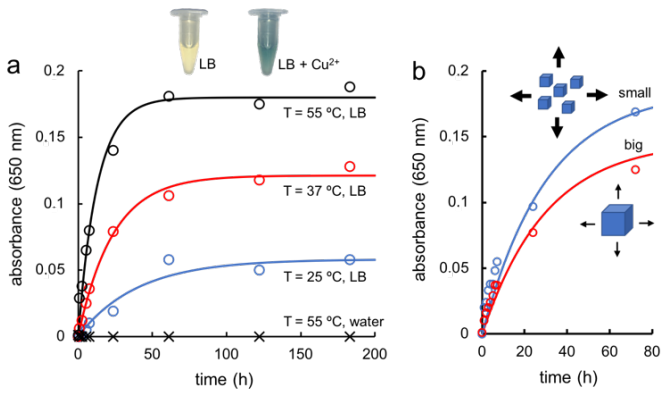

Figure 2.2 Copper release from macroscopic HKUST crystals. **a** Release kinetics in LB medium (or water) at different temperatures (25, 37, and 55 ºC). **b** Release kinetics from big (1-2 mm) or small (~0.2 mm) crystals at 37 ºC. Circles, experimental data; lines, theoretical models.

between lattice energy, solvation, and coordination with nitrogen-based compounds. This was clearly observed by a bluish coloration of the LB medium. The kinetics of the copper release was well explained by an exponential decay model in increasing form. That is, $x(t) = x_{max}\ (1 - e^{-kt})$, where $x(t)$ represents the amount of Cu^{2+} in the medium at time t, x_{max} the maximal amount of this metal observed in the medium, and k the release rate. This corresponds to the differential equation $dx/dt = k\ (x_{max} - x)$. At initial times, the release flux is roughly constant ($k \cdot x_{max}$). But, as the amount of Cu^{2+} in the medium increases, the flux is lower (as the electrostatic forces equilibrate). The absorbance at 650 nm was a good proxy to monitor the release. Due to the direct correlation between absorbance and Cu^{2+} concentration, obtained with control solutions, in **Fig. 2.3** we present the kinetics of the release in terms of concentration. Despite immersion in water was already shown to induce structural changes that can be observed by powder X-ray diffraction (PXRD), scanning electron microscopy, and even visually (20), we show here that copper cations are not released in this pure medium.

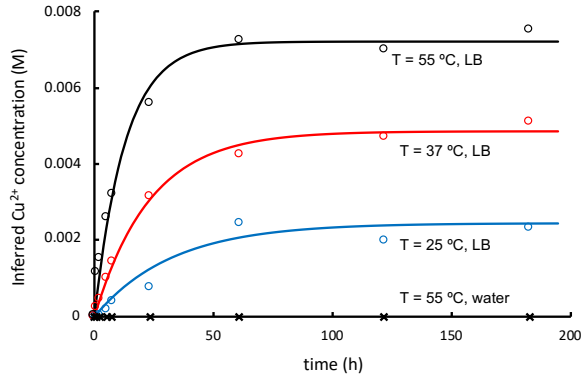

Figure 2.3 Copper release from macroscopic HKUST crystals in terms of concentration. The values of k (kinetic rate of release) are preserved

At 37 ºC, we fitted $k = 0.0428 \pm 0.0032$ h^{-1} and $x_{max} = 0.1213 \pm 0.0026$ ($R^2 = 0.9942$), for 10 mg HKUST-1 (particles of ~0.2 mm size) immersed in 2 mL LB. That is, the system takes $3/k = 70$ h

in reaching the equilibrium (95% of x_{max}). In a colder environment (25 ºC), the kinetics is slower (almost the half) and the maximal amount released lower, obtaining $k = 0.0251 \pm 0.0006$ h^{-1} and $x_{max} = 0.0581 \pm 0.0047$ ($R^2 = 0.9549$); whilst in a hotter environment (55 ºC), the kinetics is faster (three-fold) and the maximal amount released higher, giving $k = 0.0774 \pm 0.0077$ h^{-1} and $x_{max} = 0.1801 \pm 0.0049$ ($R^2 = 0.9863$). Qualitatively, this is in tune with a thermodynamic view of this process of partial crystal degradation.

Because of diffusion and degradation are surface-driven phenomena, we hypothesized that the bigger the HKUST-1 crystal, the higher the ability and slower the kinetics of copper release. To prove this, we prepared big and small crystals of observable size (particles of 1-2 mm edge and ~0.2 mm edge, respectively) to be assayed in the same conditions (maintaining constant the total material mass). Results shown in **Fig. 2.2b** indeed confirmed our hypothesis (with $k = 0.0333 \pm 0.0048$ h^{-1} and $x_{max} = 0.1472 \pm 0.0051$, $R^2 = 0.9734$ in the case of big material; and $k = 0.0399 \pm 0.0054$ h^{-1} and $x_{max} = 0.1853 \pm 0.0060$, $R^2 = 0.9733$ for small material). We found statistically significant the difference between the two fittings, taking the mean and standard deviation of k and

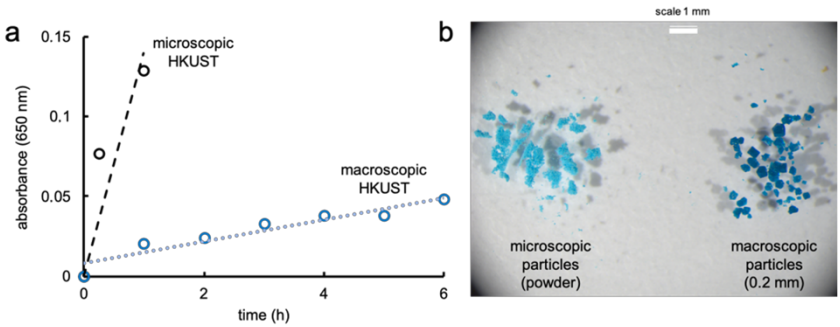

Figure 2.4 **a** Kinetics of copper release from macroscopic or microscopic HKUST particles (in LB medium at 37 ºC). **b** Photo of materials used taken with a Leica optical microscope.

x_{max} (Welch's *t*-tests, two-tailed $P < 0.05$ in both cases). Moreover, we found that the copper release from a microscopic HKUST-1 material (grinding mm-sized particles) is substantially faster (see **Fig. 2.4**).

Once we characterized the ability of HKUST-1 crystals to release copper in LB medium, and knowing the oligodynamic effect of this metal (18), we proceeded to test if a bacterial population could be maintained under control with our materials of observable size. To tackle this important question, we took diluted *E. coli* cultures (1:100 from overnight, strain DH5α) to then grow them in liquid medium (LB) with and without crystals (at 37 ºC with shaking). The absorbance at 500 nm was now a good proxy of bacterial amount, avoiding potential interfering signals from copper at > 550 nm (**Fig. 2.5**).

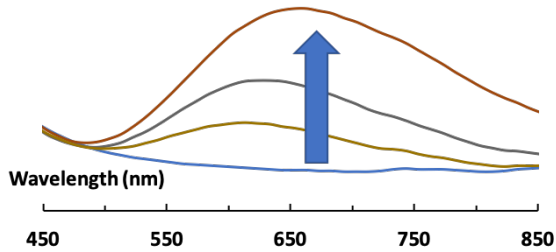

Figure 2.5 Representative UV/Vis spectra of LB upon the addition of Cu^{2+} at different concentrations (arrow).

As this metal acts as a bacteriostatic agent on *E. coli*, the use of absorbance is justified to monitor the repression. We found that these observable crystals, both big and small, inhibit the growth of *E. coli* (**Fig. 2.6a**).

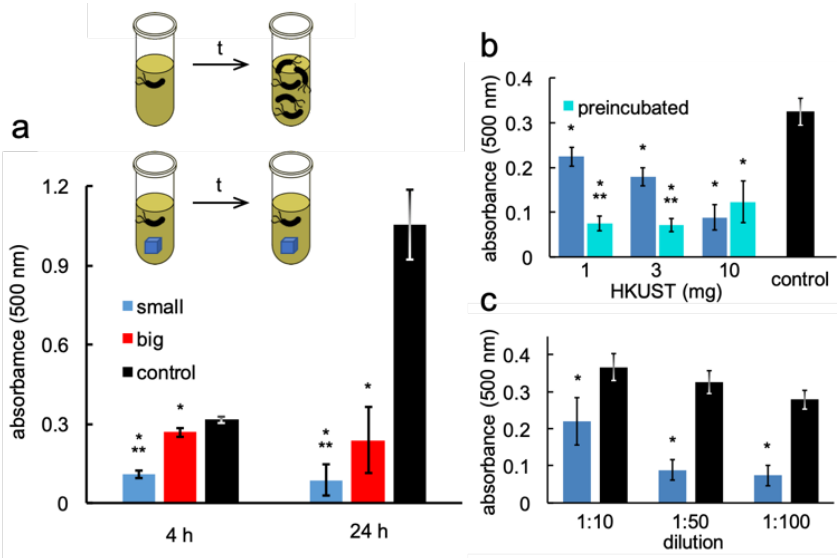

Figure 2.6 Bacterial population control with macroscopic HKUST crystals. **a** Absorbance of bacterial cultures (LB) in presence of HKUST (either small or big) at 4 and 24 h. **b, c** Absorbance of bacterial cultures growing in exponential phase with and without HKUST by varying the amount of material or the dilution factor. Error bars represent standard deviations over biological replicates ($n = 3$, except $n = 6$ for big crystals). *Statistical significance with respect to control. **Statistical significance of small *vs.* big (a, although marginal at 24 h) or preincubated *vs.* non-preincubated (b).

The differences reported (at 4 and 24 h after material addition, 10 mg) were indeed significant (Welch's *t*-tests, two-tailed $P < 0.05$) compared to populations grown in absence of materials. However, the inhibition achieved with big crystals at 4 h was minimal. The different cultures almost did not grow from 4 to 24 h with material, pointing out that a threshold level in the Cu^{2+} amount in the medium needs to be reached to ensure growth control. Moreover, the inhibition was significantly higher with smaller crystals (Welch's *t*-tests, two-tailed $P < 0.05$), in tune with the differential copper release previously shown (**Fig. 2.2b**).

It is important to note that the efficacy of the system depends on the initial bacterial density and the material amount in the medium. To appreciate these issues, we then performed a series of complementary assays with diluted *E. coli* cultures already in exponential phase (*i.e.*, when bacteria replicate at their maximal speed). We first measured how different amounts of small HKUST-1 crystals (1, 3, and 10 mg) repress bacterial growth (Welch's *t*-tests, two-tailed $P < 0.05$), revealing a progressive effect (**Fig. 2.6b**). We repeated these measurements when crystals had the opportunity to release copper before bacteria appeared (material preincubation), observing now that even 1 mg was enough to keep under control the population (inhibition significantly higher with respect to non-preincubation; Welch's *t*-test, two-tailed $P < 0.05$). Second, we studied how different dilutions (1:10, 1:50, and 1:100) of bacterial cultures impact the system, finding repression in all cases (Welch's *t*-tests, two-tailed $P < 0.05$) and again a progressive effect (**Fig. 2.6c**). Collectively, these data promise to be important to develop accurate quantitative models of real (preclinical) scenarios (21). In addition, HKUST-1 crystals were also able to inhibit bacteria in solid medium (LB-agar; **Fig. 2.7a**). For that, we placed big and small HKUST-1 crystals in plates seeded homogenously with bacteria, finding the day-after halos around the crystals. Halos even overlapped when particles were close.

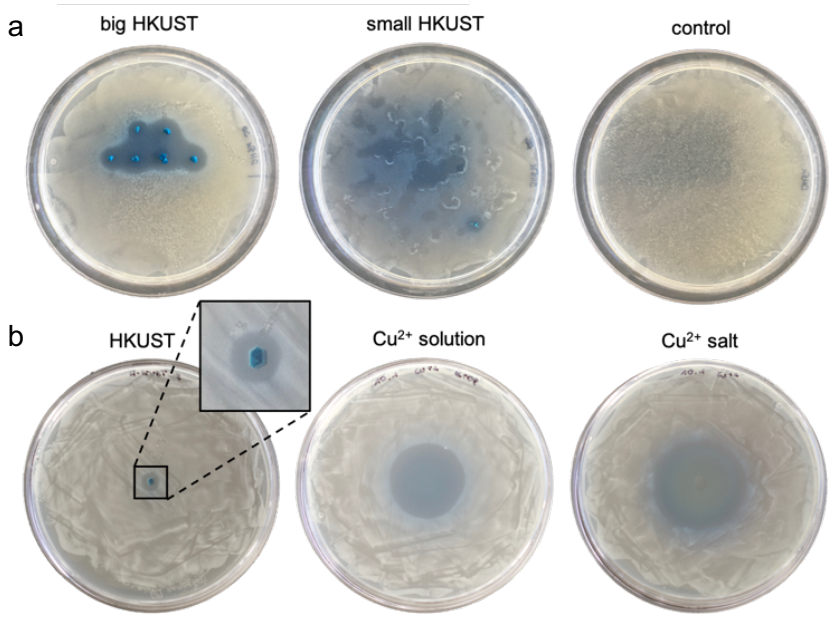

Figure 2.7 Bacterial population control with macroscopic HKUST crystals. **a** Petri dishes (LB-agar) seeded with bacteria and set with big HKUST crystals (left), small ones (middle), or without (right). **b** Petri dishes seeded with bacteria and set with a big HKUST crystal (left), a solution of $Cu(NO_3)_2$ (middle), or a powder of $Cu(NO_3)_2$ (right). Images taken after overnight incubation.

Notably, a slow release of antimicrobial agents by autonomous structures is key to achieve a more localized effect. This can allow obtaining higher levels of these agents at a given point minimizing systemic toxicity, which can enhance the effectiveness of a treatment (22). To illustrate that macroscopic HKUST-1 crystals are suitable elements to achieve this spatiotemporal behavior, we placed in plates seeded homogenously with bacteria, on the one hand, a big crystal (8.1 mg) and, on the other hand, either a microdrop of a Cu^{2+} solution (5 µL, 9.7 mg) or a microspoon of a Cu^{2+} salt (10.1 mg), ensuring the same metal amount. Suggestively, we found that the bacterial repression halo produced

by the crystal (~7 mm diameter) was substantially smaller than the halo produced by the microdrop (~26 mm diameter) or the microspoon (~34 mm diameter; **Fig. 2.7b**).

In addition to copper, HKUST-1 crystals release to the medium BTC (**Fig. 2.1**), the organic ligand in this MOF (15). Thus, we decided to explore the corresponding dose-response surface (with separate compounds). From diluted *E. coli* cultures, we observed that Cu^{2+} inhibits bacterial growth when its concentration is above 2.5 mM (see **Fig. 2.8**).

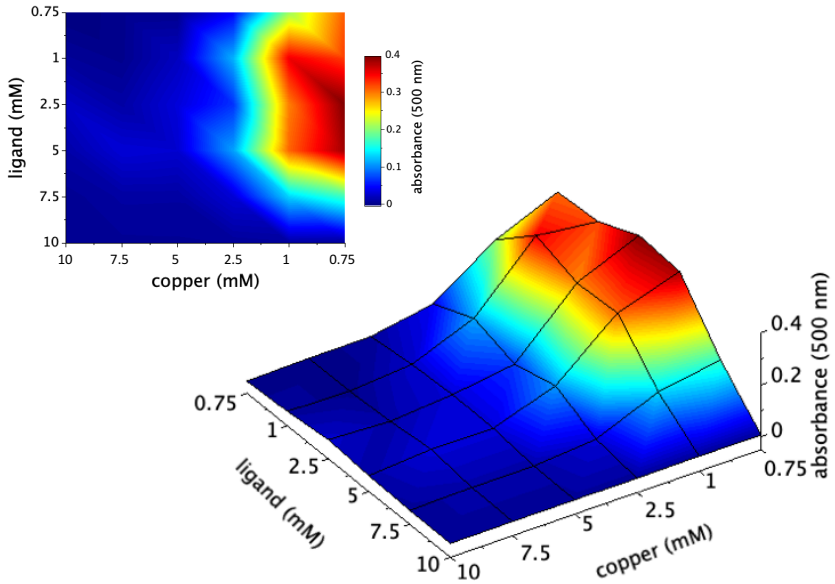

Figure 2.8 Two-dimensional analysis of toxicity over a bacterial population of Cu^{2+} (metal) and benzene-1,3,5-tricarboxylic acid (BTC, organic ligand in HKUST).

Following **Fig. 2.3**, this concentration is reached after ~16.7 h of material incubation at 37 ºC without shaking. However, when shaking is applied, a faster profile was observed, reaching that

threshold after just ~2.2 h (see **Fig. 2.9**; likely due to mechanical effects). In this case, we fitted $k = 0.2711 \pm 0.0471$ h^{-1} and $x_{max} = 0.1413 \pm 0.0089$ ($R^2 = 0.9444$). While, for BTC the threshold is at 7.5 mM (**Fig. 2.8**). Interestingly, BTC alone can repress bacteria, which agrees with the use of phenolic acids as antiseptic agents (23). Yet, because the inhibition threshold is 3-fold lower for Cu^{2+} and there is more metal than ligand molecules in this MOF (3:2 ratio), we can argue that BTC contributes less substantially than Cu^{2+} to bacterial repression when using the material for this purpose.

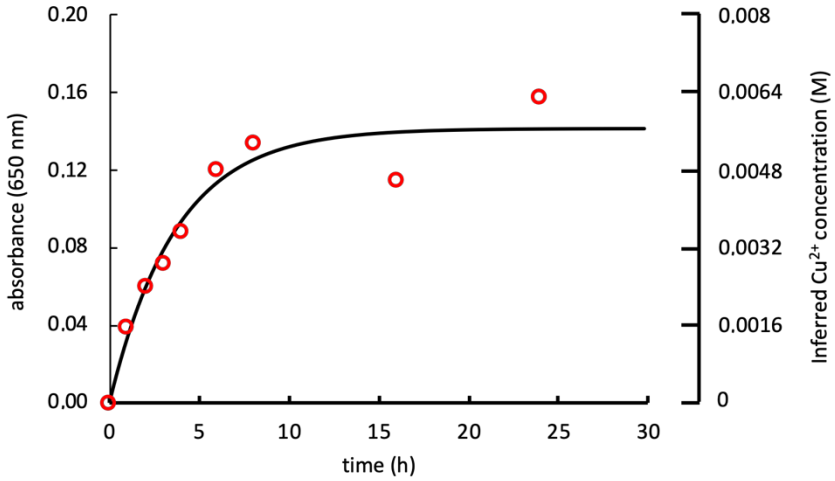

Figure 2.9 Copper release from macroscopic HKUST crystals in terms of absorbance (left y axis) and concentration (right y axis). Circles, experimental data; line, theoretical model. The difference between this and Fig. 3 is that here shaking was applied to the cultures, which inevitably contributes in a mechanical way to crystal damage, accelerating the release process.

Subsequently, we studied to what extent HKUST-1 crystals can be reused. For that, we cultured bacteria in the presence of crystals (10 mg) for long time (> 10 h), showing repression, and then we recycled those crystals to constrain a new bacterial culture. We did

so in liquid and solid media, obtaining repression again (see **Fig. 2.10**).

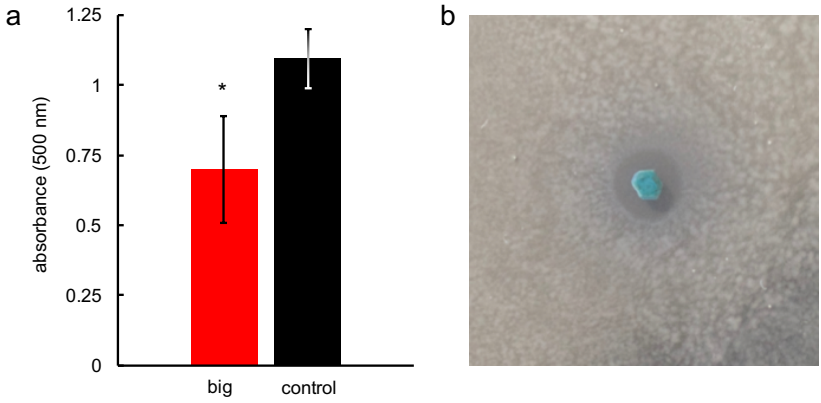

Figure 2.10 Reusability of HKUST crystals. **a** Control of a bacterial population in liquid medium with reused crystals (10 mg initially). Error bars are standard deviations over 3 replicates.* Welch's *t*-test, two-tailed $P < 0.05$. **b** Control of a bacterial population in solid medium with a reused crystal (8.1 mg initially). Results obtained with big crystals at 24 h (second use).

Certainly, the macroscopic size of the material was instrumental to achieve such reusability, allowing the manual collection and disposition of individual MOF particles. Our experiments further revealed that ~3 mg material was decomposed into metal and ligand after 24 h in LB with shaking, leading to a final Cu^{2+} concentration of ~5.6 mM (**Fig. 2.9**). This indeed agrees with our results with preincubated materials. In terms of activity, reused big crystals were shown less efficient (36% *vs.* 77% when fresh; likely due to 30% mass loss). In addition, ~3.2 mg material was decomposed after 5 d in LB-agar (note that we observed fully decomposition at very long times), which is a much slower process. Accordingly, consistent activity was observed after different cycles in solid medium (halos of equal diameter; **Fig. 2.10**).

Finally, we aimed at maximizing the utility of this material to control bacterial populations. We investigated the ability to exploit macroscopic HKUST-1 crystals as molecular containers by loading crystals with a model antibiotic. We considered chloramphenicol, which is a broad-spectrum antibiotic with mainly bacteriostatic mode of action (24), as well as copper (see **Fig. 2.11**). For that, already prepared HKUST-1 crystals of ~0.2 mm edge were immersed in a saturated solution of chloramphenicol. We quantified in ~7% (mass) the presence of the guest molecule into the material (through ethanol washes, absorbance at 280 nm), meaning that only the pores more in the periphery of the material were occupied.

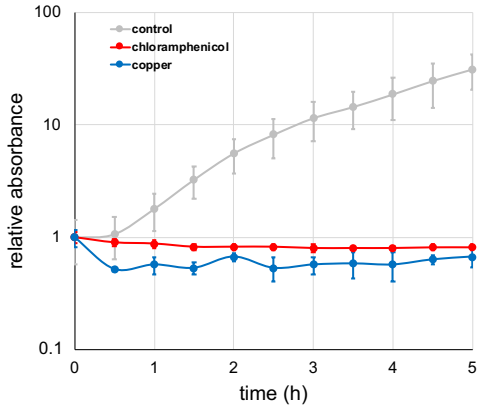

Figure 2.11 Time-dependent relative absorbance (absorbance after 4 h of incubation over the initial absorbance with a dilution 1:10, 500 nm), showing the bacteriostatic effect of copper [6 mg/mL $Cu(NO_3)_2{\cdot}(H_2O)_3$] and chloramphenicol (34 µg/mL). Control cultures grown normally. Error bars are standard deviations over 3 replicates (variance is really minimal in the case of chloramphenicol).

From diluted *E. coli* cultures in liquid medium (1:100 from overnight), we found significant repression with these new crystals (at 4 h, with 1 and 3 mg; Welch's *t*-tests, two-tailed $P < 0.05$; **Fig.**

2.12a). Even with 1 mg material and without preincubation, the population was tightly controlled. Indeed, 1 mg material leads to a maximal concentration of released chloramphenicol of ~35 µg/mL (in 2 mL culture), which is the standard concentration for resistance selection. Thus, a low loading percentage was not an obstacle here. Note also that bacteria grew about 10 times more with mere HKUST-1 crystals (**Fig. 12a**; Welch's *t*-tests, two-tailed $P < 0.05$). Consequently, we inferred that guest molecules release much faster than copper (perhaps, $k > 0.4$ h^{-1}) in these MOF particles.

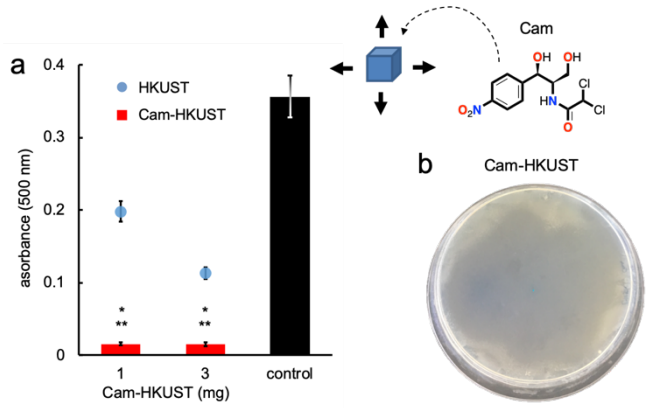

Figure 2.12 Bacterial population control with macroscopic HKUST crystals loaded with chloramphenicol (Cam-HKUST). **a** Absorbance of bacterial cultures (LB) in presence of Cam-HKUST at 4 h. Error bars represent standard deviations over biological replicates ($n = 3$). *Statistical significance with respect to control. **Statistical significance of the enhancement over mere HKUST crystals. **b** Petri dish (LB-agar) seeded with bacteria and set with Cam-HKUST crystals. Image taken after overnight incubation.

We further assessed the contribution of chloramphenicol to the joint antibiotic-MOF action by playing with a concentration gradient, with fixed 3 mg material, pointing out how sublethal chloramphenicol levels (around 1 µg/mL) can work relatively well in the presence of MOF-derived Cu^{2+} ions (see **Fig. 2.13**).

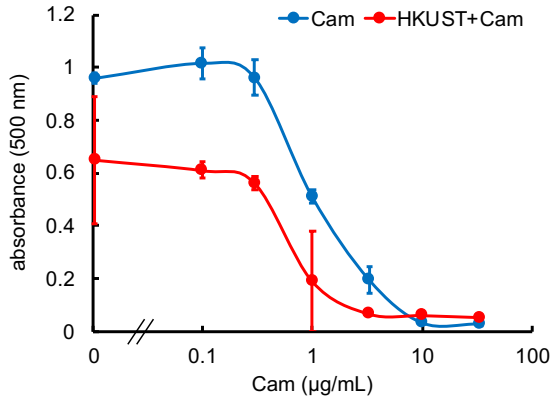

Figure 2.13 Absorbance of bacterial cultures for a gradient of chloramphenicol (0 - 34 µg/mL), with and without HKUST crystals (3 mg) at 24 h. Error bars are standard deviations over 3 replicates.

Expectedly, chloramphenicol-loaded HKUST-1 crystals were also able to inhibit bacteria in solid medium (**Fig. 2.12b**). As before, we placed crystals in plates seeded homogenously with bacteria, but this time we did not observe a set of halos (compare with **Fig. 2.7b**, middle). Instead, we found a large region in which bacteria did not grow, suggesting a fast release and spread of the guest molecule.

2.3. Discussion

In conclusion, this work shows that bacterial population control can be achieved by repurposing a given MOF prepared to be observable (size $>$ 0.2 mm). We found that HKUST-1 crystals are slowly decomposed into their constituents (Cu^{2+} and BTC)(15) in biological media, which function as antimicrobial agents (mainly bacteriostatic). Note that this is not the case in pure water. Because of the size of the material, such a decomposition leads to a continuous, but slow release of antimicrobials, even for several days (at 37 ^{o}C, we quantified it in ~3 d, without affecting the material-medium equilibrium). In addition, we proved that this

material can be loaded with a suitable antibiotic (chloramphenicol, which fits inside the pore) to enhance the antimicrobial action.

This allowed the control of bacterial populations (*E. coli*) in liquid (LB) or solid (LB-agar) media; a complex spatiotemporal process that depends on the initial bacterial density and the amount of material disclosed. In solid media, the slow release ensured a more localized action. Moreover, thanks to have mm-sized materials, we were able to precisely reuse them in different days against different populations, something that certainly aligns with a sustainable development.

In the literature, there are several game-changing examples of porous nanomaterials (typical size < 200 nm) that go towards the miniaturization or substitution of conventional systems for the administration of biochemical compounds (5,6). For example, Fe^{3+}-based MOFs (MIL-53 and variants)(25). Our results contribute to shift the paradigm, as they rely on MOFs of macroscopic size (jumping 3-4 orders of magnitude) that can be at the same time containers of guest molecules and self-degrading reservoirs of metals. We can imagine several scenarios in which a programmed supply of different bioactive ions/molecules can be highly beneficial. For example, the antibiotic dose might be reduced if metals were at play. They might also be used to enhance the activity of established drugs in preclinical assays (26). Moreover, different guest molecules might be released according to a predefined temporal pattern if we filled such particles of big size in consecutive steps. From a systems biology perspective (27), a slow release might be interesting to study the response of biological systems to ramp-like instead of step-like stimuli, trying to unveil overlooked features in stress response (28).

Future work should be focused on testing whether it is possible to repress pathogenic bacteria using macroscopic HKUST-1 crystals (loaded with antibiotic or not)(17), both *in vitro* and *in vivo* (with animal models). However, a limitation of this technology is the fact that BTC is a relatively small ligand, which constrains the palette of guest molecules (at the same time that hinders the loading). Hence, future work should also investigate the possibility of preparing other mm-sized MOFs, with wider pores and based on different metals (14) (*e.g.*, silver or zinc, which also present antibacterial properties [9]). Moreover, as these materials admit multiple post-synthetic modifications (metal exchange, ligand exchange, surface modification, *etc.*)(29), there is a great room for maneuver.

2.4. Materials and Methods

2.4.1. Material preparation.

HKUST crystals were prepared following the conditions reported by Zhu and colleagues with some modifications. Briefly, $Cu(NO_3)_2{\cdot}(H_2O)_3$ (622 mg, 2,5 mmol) and benzene-1,3,5-tricarboxylic acid (422 mg, 2 mmol) were dissolved in dimethylformamide (DMF) (25 mL) with ultrasonication at 25 ^{o}C. Then, 100 mL of a 2 M nitric acid aqueous solution were added. The resulting mixture was stirred at 5 ^{o}C for 48 h. On the one hand, to prepare big crystals (1-2 mm), 5 x 20 mL of this mother solution were transferred to five Erlenmeyers (of 100 mL capacity) and heated at 81 ^{o}C for 7 d in a preheated sand bath. A glass top was placed above each Erlenmeyer allowing gas to scape. The resulting HKUST crystals were washed three times with DMF (an ideal solvent for copper nitrate and BTC) over a period of 72 h and three additional times with ethanol (48 h), and they were

stored in ethanol. Activation was performed by heating crystals at 55 ºC for 4 h in order to prevent crystal damage. Big crystals were collected manually among the resulting MOF particles.

On the other hand, to prepare smaller crystals (0.2-0.5 mm), the mother solution was directly heated in a round-bottom flask (of 250 mL capacity) at 81 ºC without stirring. After 7 d, crystals were recovered. As before, the resulting crystals were washed three times with DMF and three times with ethanol. Furthermore, to obtain microscopic HKUST crystals, we ground with a mortar a given quantity of small crystals to obtain a fine powder (this way, we ensured the same nanostructure). The structures of all materials were confirmed by PXRD. Of note, Zhu and colleagues performed these experiments in closed flasks (16). However, we could not perform as such for security reasons due to substantial overpressure experienced.

2.4.2. Chloramphenicol loading into material.

Chloramphenicol@HKUST was prepared by immersion of HKUST crystals (200 mg, small crystals, *i.e.*, 0.2-0.5 mm) in a saturated solution of chloramphenicol in ethanol (2 mL, 250 mg/mL) at 55 ºC for 3 d. Afterwards, the material was washed three times with cold ethanol. IR spectra indicated the presence of the guest molecule and PXRD ensured the crystallinity of the materials. Antibiotic content was calculated to be 7% (of total material mass) after complete release in ethanol (10 mg in 2+2+2 mL). UV/Vis was employed to measure the amount of chloramphenicol (absorbance at 280 nm, using a calibration curve obtained by different dilutions).

2.4.3. Measurement of copper release from material.

HKUST crystals (10 mg) were introduced in Eppendorf tubes with 2 mL LB (Sigma) or milli-Q water, and they were incubated without shaking at different temperatures (25, 37, and 55 ºC) for long time (> 3 d). At each time point, 40 µL were collected (without touching the material) to make a measurement of absorbance (650 nm) in a fluorometer (Varioskan LUX, Thermo). A 384-well plate (black, clear bottom; Corning) was used. 40 µL fresh medium was reintroduced in the Eppendorf tube at each time point to continue with the incubation when the experiment was for > 1 d.

In addition, to compare the release of small *vs.* big materials we weighed 8.6 mg of small HKUST (> 100 crystals, introduced in 2 mL LB) and 8.2 mg of big HKUST (two single crystals, introduced in 1.9 mL LB).

To estimate the amount of decomposed HKUST in liquid media, we introduced materials in culture tubes with LB for 24 h (at 37 ºC with shaking), collecting 40 µL at different time points to measure absorbance (650 nm), from which to infer copper amount and then material amount. In solid media, we placed materials in LB-agar plates and were weighed before and after incubation (5 d at 37 ºC). Note that HKUST crystals can initially contain ethanol and air, while they can contain water and other small molecules present in the medium after incubation. This may introduce a marginal error in the estimations of material amount.

As controls, different solutions were prepared with $Cu(NO_3)_2{\cdot}(H_2O)_3$ to study the relationship between absorbance at

650 nm and Cu^{2+} concentration (in LB and water). Calibration curves and UV/Vis spectra were obtained.

2.4.4. Bacterial strain and precultures.

In this work, we considered *E. coli* as a model bacterium, using the strain Dh5α. Features of this strain are: relaxed phenotype, lactose non-utilization, recombination deficient, and endonuclease A deficient.

Three different colonies were picked from plates and were grown overnight at 37 ºC with shaking (200 rpm) in 2 mL LB (in tubes of 13 mL capacity). These precultures were then diluted for subsequent use.

2.4.5. Measurement of bacterial growth in liquid medium.

In a first experience, the precultures were diluted 1:100 and then 10 mg HKUST (small or big crystals) were introduced in each of the new cultures (2 mL LB). These cultures were grown for 4 or 24 h at 37 ºC with shaking (200 rpm). At each time point, 200 µL were collected (without touching the material) to make a measurement of absorbance (500 nm) in a fluorometer (Varioskan LUX, Thermo). A 96-well plate (black, clear bottom; Corning) was used.

In a second experience, the precultures were diluted 1:100 and were grown for 4 h at 37 ºC with shaking (200 rpm), ensuring that the resulting cultures reach exponential phase. Then, the cultures were diluted 1:50 (unless otherwise specified) and 1, 3, or 10 mg HKUST (small crystals) were introduced in each of the new cultures (2 mL

LB). These cultures were grown for 4 h at the same conditions and then 200 µL were collected to perform the measurements of absorbance as stated before. Moreover, we performed an incubation of 18 h of small HKUST crystals (1, 3, or 10 mg) in 2 mL LB prior to inoculation with bacteria.

In a third experience, the precultures were diluted 1:100 and then 1 or 3 mg chloramphenicol@HKUST (small crystals) were introduced in each of the new cultures (2 mL LB). These cultures were grown for 4 h at 37 ºC with shaking (200 rpm) and then 200 µL were collected to perform the measurements of absorbance as stated before.

2.4.6. Measurement of bacterial growth in solid medium.

In a Petri dish (90 mm diameter) of LB medium jellified with agar (15 g/L), 100 µL of a saturated bacterial culture were seeded homogenously. On the one hand, big HKUST crystals were manually placed in the center of the plate. On the other hand, small HKUST crystals (3 mg) were spread on the plate randomly. In the case of chloramphenicol@HKUST, the crystals (1 mg) were spread on the plate randomly. The different plates were incubated overnight at 37 ºC. Photos taken with a Canon EOS 600D (obj. Canon EF 50 mm, F/1.8 STM). The resulting halos were then quantified (diameter) with a ruler.

As controls in the evaluation of the local effect of the material, a single microdrop of a solution prepared with $Cu(NO_3)_2{\cdot}(H_2O)_3$ was used (and also a microspoon of that salt). We calculated the amount of copper cations in the HKUST crystal, the microdrop, and the microspoon to be equal.

To evaluate the reusability of the material (big crystals), new plates were seeded homogenously with bacteria the day after seeding the first plates and the same crystals were used again (removed from one plate and placed in another).

2.4.7. Statistical analysis

To compare the effect of two different conditions, two-tailed Welch's *t*-test tests were performed. These tests are used when the two population variances are not assumed to be equal (the two sample sizes may or may not be equal) and hence must be estimated separately. *P* values were used to assess the statistical significance.

2.5. References

1. World Health Organization. (2018) World health statistics: monitoring health for the sustainable development goals. Geneva.

2. Laxminarayan R, Duse A, Wattal C, et al. Antibiotic resistance-the need for global solutions. *Lancet Infect Dis.* 2013;13(12):1057-1098.

3. Hawgood S, Hook-Barnard IG, O'Brien TC, Yamamoto KR. Precision medicine: Beyond the inflection point. *Sci Transl Med.* 2015;7(300):300ps17.

4. Huh AJ, Kwon YJ. "Nanoantibiotics": a new paradigm for treating infectious diseases using nanomaterials in the antibiotics resistant era. *J Control Release.* 2011;156(2):128-145.

5. Aznar E, Oroval M, Pascual L, Murguía JR, Martínez-Máñez R, Sancenón F. Gated Materials for On-Command Release of Guest Molecules. *Chem Rev.* 2016;116(2):561-718.

6. Horcajada P, Gref R, Baati T, et al. Metal-organic frameworks in biomedicine. *Chem Rev.* 2012;112(2):1232-1268.

7. Corma A. From microporous to mesoporous molecular sieve materials and their use in catalysis. *Chem Rev.* 1997;97(6):2373-2419.

8. Rodríguez HS, Hinestroza JP, Ochoa-Puentes C, Sierra CA, Soto CY. Antibacterial activity against Escherichia coli of Cu-BTC (MOF-199) metal-organic framework immobilized onto cellulosic fibers. *J Appl Polym Sci.* 2014;131(19):40815.

9. Demirci S, Ustaoğlu Z, Yilmazer GA, Sahin F, Baç N. Antimicrobial properties of zeolite-X and zeolite-A ion-exchanged with silver, copper, and zinc against a broad range of microorganisms. *Appl Biochem Biotechnol.* 2014;172(3):1652-1662.

10. Dobrovolskaia MA, Aggarwal P, Hall JB, McNeil SE. Preclinical studies to understand nanoparticle interaction with the immune system and its potential effects on nanoparticle biodistribution. *Mol Pharm.* 2008;5(4):487-495.

11. Whitesides GM. The "right" size in nanobiotechnology. *Nature Biotechnology.* 2003;21(10):1161-1165.

12. Titze T, Lauerer A, Heinke L, et al. Transport in Nanoporous Materials Including MOFs: The Applicability of Fick's Laws. *Angew Chem, Int E. Engl.* 2015;54(48):14580-14583.

13. Liu Y, Feng Y, Yao J. Recent advances in the direct fabrication of millimeter-sized hierarchical porous materials. *RSC Adv.* 2016;6(84):80840-80846.

14. Furukawa H, Cordova KE, O'Keeffe M, Yaghi OM. The chemistry and applications of metal-organic frameworks. *Science.* 2013;341(6149):974.

15. Chui SSY, Lo SMF, Charmant JPH, Orpen AG, Williams ID. A chemically functionalizable nanoporous material [Cu3(TMA)2 (H2O)3](n). *Science.* 1999;283(5405):1148-1150.

16. Li L, Sun F, Jia J, Borjigin T, Zhu G. Growth of large single MOF crystals and effective separation of organic dyes. *Cryst Eng Comm.* 2013;15(20):4094-4098.

17. Peleg AY, Hooper DC. Hospital-Acquired Infections Due to Gram-Negative Bacteria. *N Engl J Med.* 2010;362(19):1804.

18. Dupont CL, Grass G, Rensing C. Copper toxicity and the origin of bacterial resistance—new insights and applications. *Metallomics.* 2011;3(11):1109.

19. Burtch NC, Jasuja H, Walton KS. Water stability and adsorption in metal-organic frameworks. *Chem Rev.* 2014;114(20):10575-10612.

20. Álvarez JR, Sánchez-González E, Pérez E, et al. Structure stability of HKUST-1 towards water and ethanol and their effect on its CO2 capture properties. *Dalton Transactions.* 2017;46(28):9192-9200.

21. Nielsen EI, Friberg LE. Pharmacokinetic-pharmacodynamic modeling of antibacterial drugs. *Pharmacol Rev.* 2013;65(3):1053-1090.

22. Nandi SK, Mukherjee P, Roy S, Kundu B, De DK, Basu D. Local antibiotic delivery systems for the treatment of

osteomyelitis – A review. *Materials Science and Engineering: C*. 2009;29(8):2478-2485.

23. Cueva C, Moreno-Arribas MV, Martín-Álvarez PJ, et al. Antimicrobial activity of phenolic acids against commensal, probiotic and pathogenic bacteria. *Res Microbiol*. 2010;161(5):372-382.

24. Malik VS. Chloramphenicol. *Adv Appl Microbiol*. 1972;15(C):297-336.

25. Lun X, Wells JC, Grinshtein N, et al. Disulfiram when combined with copper enhances the therapeutic effects of temozolomide for the treatment of glioblastoma. *Clin Cancer Res*. 2016;22(15):3860-3875.

26. Horcajada P, Chalati T, Serre C, et al. Porous metal–organic-framework nanoscale carriers as a potential platform for drug delivery and imaging. *Nat Mater*. 2009;9(2):172-178.

27. Alon U. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. 1st ed. Chapman and Hall/CRC; 2006.

28. Young JW, Locke JCW, Elowitz MB. Rate of environmental change determines stress response specificity. *Proc Natl Acad Sci U S A*. 2013;110(10):4140-4145.

29. Yin Z, Wan S, Yang J, Kurmoo M, Zeng MH. Recent advances in post-synthetic modification of metal–organic frameworks: New types and tandem reactions. *Coord Chem Rev*. 2019;378:500-512.

Chapter 3: Evolutionary canalization mediated by altered expression of the GroEL chaperone.

This work has not been published in a journal yet.

Montagud-Martínez R, Toft C, Sabater-Muñoz B, Fares MA. In preparation.

RMM fulfilled the experimental work, prepared the draft, and the figures under the supervision of BSM and CT. CT performed the bioinformatic analysis. MAF designed the research.

Further reading about this topic in which I have contributed:

Aguilar-Rodríguez J, Sabater-Muñoz B, Montagud-Martínez R, Berlanga V, Alvarez-Ponce D, Wagner A, Fares MA. The Molecular Chaperone DnaK Is a Source of Mutational Robustness. *Genome Biol Evol.* 2016;8(9):2979-2991.

Sabater-Muñoz B, Prats-Escriche M, Montagud-Martínez R, López-Cerdán A, Toft C, Aguilar-Rodríguez J, Wagner A, Fares MA. Fitness Trade-Offs Determine the Role of the Molecular Chaperonin GroEL in Buffering Mutations. *Mol Biol Evol.* 2015;32(10):2681-2693.

3.1. Introduction

Throughout evolution, mutations and the balance between natural selection and genetic drift have mostly determined the origin of biological variability. The mutations, which occur stochastically, have deleterious, neutral, or beneficial effects on biological fitness. Some neutral variants do not have any phenotypic manifestation and are referred to as phenotypically silent or cryptic variants. A number of experiments have demonstrated that these variants could bear pre-adaptive mutations that can spark innovations in new environments (1–6). Many of these variants are silent thanks to mechanisms of mutational robustness that buffer their slightly deleterious effects on fitness and facilitate their persistence in a population.

Even though characterizing mechanisms of mutational buffering is a fundamental aim of evolutionary biology (7), it remains unclear how these mechanisms exactly act. Moreover, despite the fact that beneficial mutations are the "building blocks" of adaptive changes, little is known about the distribution of the effects of mutations (8).

Focusing on non-synonymous mutations, protein evolution is constrained by mutations that destabilize the proteins' structure (9–11). In most proteins, a few amino acids are directly responsible for the protein activities, while the majority of amino acids help provide a stable structural scaffold. Mutations can reduce protein activity and biological fitness, for example, by reducing the amount of correctly folded and thus active protein. They can also increase a protein's propensity to form toxic aggregates of misfolded

proteins (12–14). Mutations that create a new protein activity are especially often destabilizing (15–18).

Chaperones, or heat shock proteins, could play a fundamental role in helping protein evolution by buffering destabilizing mutations that otherwise would be detrimental (19–21). They are essential among the regulatory mechanisms of evolution since they are responsible for the correct folding of proteins in the cell, as well as, avoiding the formation of toxic protein aggregates. They also are indispensable under a wide range of conditions, both physiological and under stress, and are ubiquitous in nature, although expression levels may vary. Furthermore, as a consequence of their modulation capacity, they become phenotypic variability enablers since they allow a large number of genetic variants to remain in the population (4, 22, 23). The overexpression of these chaperones, in particular GroEL, allows the persistence of cells under strong bottleneck dynamics, and therefore also under the high load of deleterious mutations that gene drift entails. In addition, as a consequence, their client proteins (those that need GroEL for their correct folding) have a higher mutation rate than those that are not (21, 24–26). The overexpression of groE also occurs naturally in the endosymbiotic bacteria of insects even in the absence of stress, since when propagation between generations occurs, a large bottleneck occurs; the overexpression of these chaperones becomes essential for the maintenance of the already reduced proteome of endosymbiotic bacteria compared to closely related free-living (non-symbiotic) bacteria (27). In addition, the essentiality of the groEL function is also portrayed in the fact that the number of mutations observed in groEL is much lower than that observed in other genes of these same endosymbionts (27, 28).

However, more recently it has been found that chaperones, by facilitating protein folding, can both suppress the effect of deleterious mutations that affect protein folding and enhance the effect of phenotype-altering mutations. More importantly, chaperones can simultaneously enhance a phenotype while suppressing another (29). GroE overexpression can reduce genetic diversity during experimental evolution, implying that it helps purge deleterious mutations (30). This suggests that chaperones can possibly affect adaptive evolution in more than one way.

Consequently, the role that GroES/EL plays in channeling mutations and the possible new functions that may emerge from these changes in nucleotide sequence remains ambiguous. Chaperones act as potentiators of evolvability by 'smoothing' adaptive landscapes, making possible 'jumps' to remote peaks that would otherwise be inaccessible due to low biological fitness valleys, acting as capacitors of adaptation. But also, chaperones can have buffering effects by “hiding” the deleterious effect of some destabilizing mutations, entailing as a result less phenotypic variability. Hence, the terms buffering and potentiation are highly contextual (30).

Long-term evolution experiments (LTEE) in bacteria involve studying the changes that occur in a bacterial population over many generations, and allow the direct study of various traits, such as antibiotic resistance, metabolism, etc (31–33). Undoubtedly, experimental evolution has great potential for testing hypotheses about the mechanisms underlying certain evolutionary events. Such experiments make it possible to observe evolution as it is happening. Moreover, whole genome sequencing has revolutionized studies using the experimental evolution in bacteria because it

readily provides a comprehensive insight into the genetic bases of adaptation (34).

We decided to carry out a LTEE to study how bacteria are able to maintain a given function (or phenotype) at long times, *i.e.*, when the accumulation of mutations may present a challenge. Especially, when a key factor in the cell, such as GroEL, is present in a range of levels of expression, specifically where the difference in expression is not provided by an inducible *groES/L* harbored in a plasmid, but a stable change in the bacterial chromosome.

To study the role of the protein chaperone GroEL on the evolution (evolvability and stability) of *E. coli*, we engineered a series of clones that differed from the MG1655 strain in the promoter controlling the groE operon, thus providing a range of different expression of the chaperone. It would be expected that different intracellular concentrations of GroEl modulated the evolutionary pathway. Then, we subjected these lines to an LTEE over 2.2 k generations under strong genetic drift to study their changes in biological fitness and mutation accumulation over time.

3.2. Results

3.2.1. Experimental evolution under strong bottlenecks: clonal evolution.

From the strain of *E. coli* str. K-12 substr. MG1655 nine isogenic lines were created. These lines only differed from each other in their promoter regions (**Fig. 3.1a**). Eight of them were obtained by creating a synthetic promoter library (SPL) based on randomized sequences flanking the consensus -10 and -35 promoter regions preceding the operon. The ninth line harbored the PBla promoter

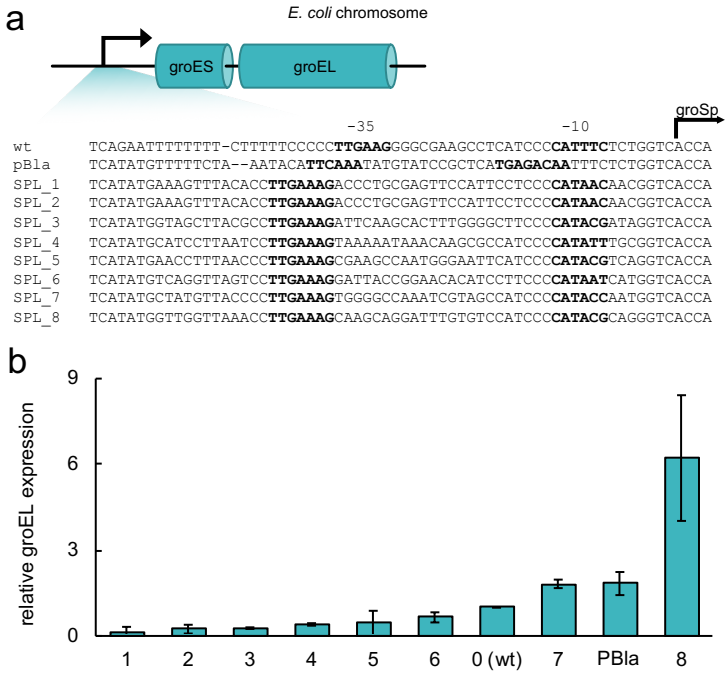

Figure 3.1 a Schematic representation of the genomic region of the *E.* coli chromosome harboring the groE operon and sequences of the SPL promoters generated for the evolution experiment. The clones only differ from each other in the promoter sequences presented in the figure. **b** Promoter characterization showing the mean relative groEL expression determined by qPCR at the initiation of the evolution experiment. Error bars correspond to standard deviation over three replicates (n=3). An ANOVA test shows that there is an overall statistically significant difference among the expression levels $P<0.0001$.

instead of the original promoter in the chromosome. This resulted in the fine-tuning of bacterial gene expression as a consequence of the σ factor interacting with this region and facilitating RNA polymerase-binding. As a result, these different promoter regions prodcuced different levels of mRNA, and consequently the expression of the GroES/L chaperonin changed compared to the MG1655 strain. The modification in expression was quantified by qPCR, showing a range from 0.24 times less to 6.22 times more,

compared to the control that harbored the wild-type (WT) promoter (**Fig. 3.1b**). Consequently, all of them underwent a knock-out of the mutS mutation repair gene, which provided a mutation rate 1000 times higher than the MG1655 WT strain.

From each of the ten lines, including those newly generated and the the one with the original promoter, a LTEE was carried out. Each of the lines had 4 biological replicates, which were subjected

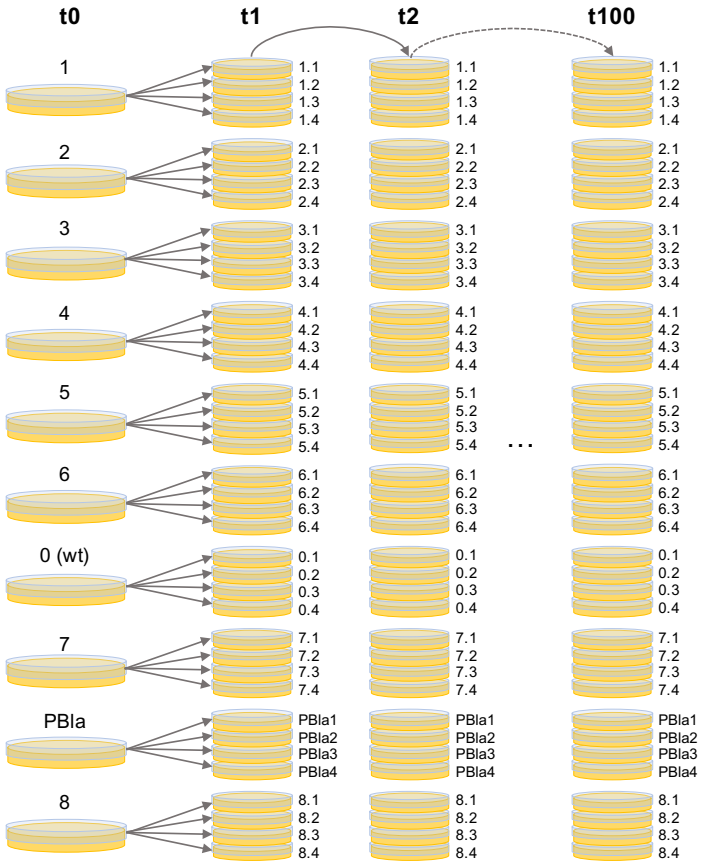

Figure 3.2 Representation of the clonal evolution. Four biological replicates are evolved from the initial 10 lines at t0 during 100 passages. A single cell was picked onto a new plate every day.

to daily single-cell passages for 100 passages (22.22 generations/day; 2222 generations) (**Fig. 3.2**). A "fossil record" was created by freezing cells from each line and replicate every 10 passages.

3.2.2. Determining changes in biological fitness: growth rates.

The relative biological fitness of each evolved line was determined at t100 (**Table 3.1**) in order to assess how it had changed in relation to the ancestral line and to help to choose a selected few to proceed with whole genome sequencing. Generally, we observed a higher variability between biological replicates in the lines that have a

Table 3.1 Relative fitness (w) values of each line, compared to their common ancestral line MG1655 ΔmutS at t0.

	t0	t100			
Line	**1**	**1.1**	**1.2**	**1.3**	**1.4**
w	0.877	0.989	1.122	1.117	1.139
Line	**2**	**2.1**	**2.2**	**2.3**	**2.4**
w	1.100	0.917	0.984	0.929	0.426
Line	**3**	**3.1**	**3.2**	**3.3**	**3.4**
w	1.025	0.556	0.795	0.820	0.901
Line	**4**	**4.1**	**4.2**	**4.3**	**4.4**
w	1.066	1.204	0.988	0.734	0.893
Line	**5**	**5.1**	**5.2**	**5.3**	**5.4**
w	1.037	0.925	1.055	1.055	0.915
Line	**6**	**6.1**	**6.2 t**	**6.3**	**6.4**
w	0.954	0.847	0.903	0.735	0.741
Line	**0**	**0.1**	**0.2**	**0.3**	**0.4**
w	0.996	0.289	0.575	0.645	0.606
Line	**PBla**	**PBla1**	**PBla2**	**PBla3**	**PBla4**
w	0.906	0.714	0.806	0.827	0.765
Line	**7**	**7.1**	**7.2**	**7.3**	**7.4**
w	0.914	0.686	0.722	0.827	0.554
Line	**8**	**8.1**	**8.2**	**8.3**	**8.4**
w	0.823	1.090	0.915	0.802	0.928

lower expression level than the WT promoter. Their biological fitness predominantly decreased throughout their evolution; while those that have a higher expression than the WT strain do not show such differences with respect to their corresponding ancestors or the line carrying the WT promoter (MG1655 ΔmutS t0). From these data, 7 lines were selected for whole genomic sequencing, in addition to their common ancestral one (MG1655 ΔmutS t0) (see growth plots of all lines in **Fig. 3.3**).

The chosen lines for sequencing were:

1. MG1655 ΔmutS at t0.
2. One replicate of the line originally expressing 0.238 times the expression of the WT. The name of this line is 2.4.
3. One replicate of the line originally expressing 0.273 times the expression of the WT. The name of this line is 3.1.
4. One replicate of the line originally expressing 0.649 times the expression of the WT. The name of this line is 6.1.
5. One replicate of the line originally expressing GroES/L at the WT level of expression, the name of this line is 0.3.
6. One replicate of the line originally expressing 1.822 times the expression of the WT. The name of this line is 7.3.
7. One replicate of the line originally expressing 6.222 times the expression of the WT. The name of this line is 8.1.
8. One replicate of the line originally expressing 6.222 times the expression of the WT. The name of this line is 8.3.

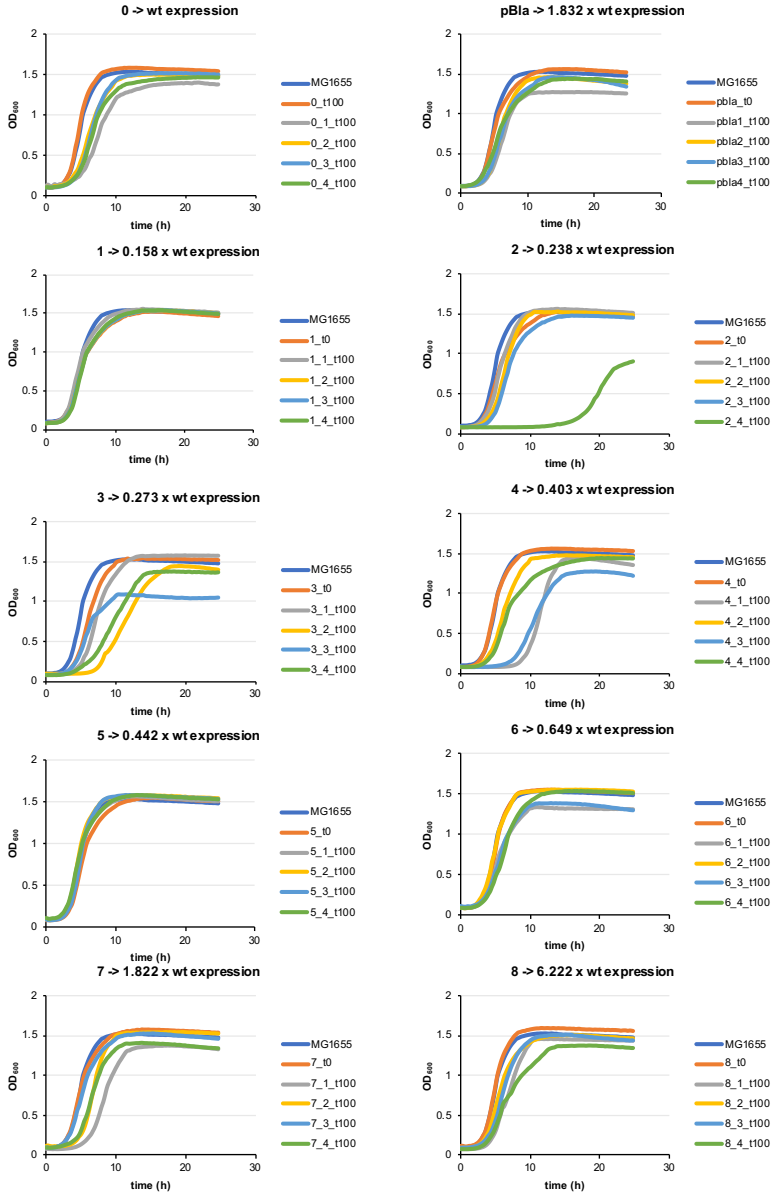

Figure 3.3 Growth curves for each of level of GroEL expression, compared to their respective ancestral line at t0 (orange) and an original MG1655 ΔmutS (dark blue). Each level of expression has 4 biological replicates (grey, yellow, light blue, and green). OD_{600} vs. time. Each line represents the median of three technical replicates for each curve.

3.2.3. Determining changes in GroEL expression: qPCR

At t100, the expression of GroES/L was verified by qPCR for each line chosen for sequencing. In those lines where the expression of the chaperone was partially silenced, as well as those that maintained the WT promoter, there was a slight increase in expression compared to its ancestral (for example, from 0.24x to

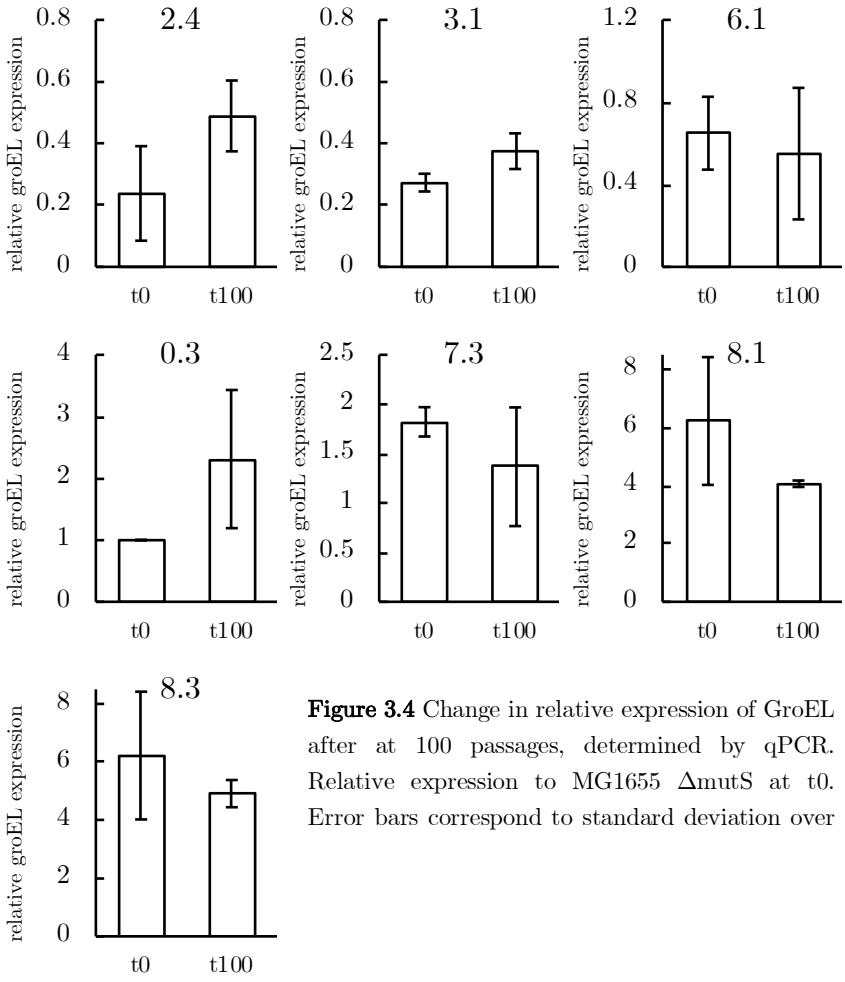

Figure 3.4 Change in relative expression of GroEL after at 100 passages, determined by qPCR. Relative expression to MG1655 ΔmutS at t0. Error bars correspond to standard deviation over

0.5x, in the 1x to 2.3x); interestingly, the line harboring the WT promoter (0.3) doubled GroEL expression. On the other hand, those that were originally expressed at 6.22x, at the end of the evolution express GroEL at around 5x the ancestral expression level. Such changes are probably due to the need to produce more chaperone to buffer lethal mutations or less to avoid the toxic effect of overexpression, respectively (**Fig. 3.4**).

3.2.4. Experimental evolution under soft bottleneck (1%), population evolution.

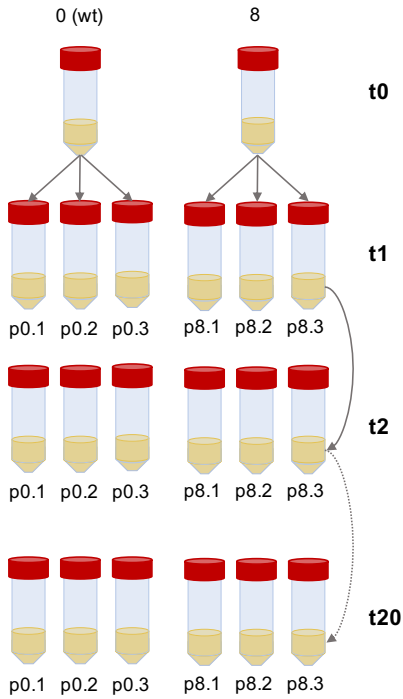

Figure 3.5 Representation of the populational evolution. Three biological replicates are evolved from the initial two lines at t0 during 20 passages. 1% of a grown culture was passaged onto fresh media every day.

A study was carried out with two lines (WT promoter and SPL 8 promoter), with three biological replicates each, evolving populationally through daily passages of 1% of the population, for 20 days (6.6 generations/day; 132 generations) (**Fig. 3.5**). At t20, we studied, using phenotypic microarrays, the possible emergence of changes in the metabolization of some carbon and nitrogen sources (EcoPlates). Phenotypic arrays allow us to observe changes in the cells' metabolism. These changes might be influenced by the

capability of enzymes to fold in other manners, changing their activity as a result. Having chaperones overexpressed, which improve correct protein folding, or the lack of enough of them, could modify the ability of the cells to consume different metabolic components, resulting in changes that we could observe using this type of array.

Figure 3.6 Picture (scanned) of the EcoPlates after incubation at 24h, purple dots indicate the reduction of the tetrazolium reporter dye. **a** Plate corresponding to the line carrying the WT promoter after 20 passes. **b** Plate corresponding to the line carrying promoter 8 after 20 passes.

For this purpose, we chose a population evolution, where natural selection is present, in contrast to a strong genetic drift where mutations are fixed stochastically. Consequently, measuring the reduction of the tetrazolium dye that occurs when the carbon or nitrogen source present in the well is metabolized, we could have measured if the level of expression of *groE* allowed a potential leap in the fitness landscape. However, no significant changes were observed between the lines with the parameters tested (see an example in **Fig. 3.6**).

3.2.5. Mutational dynamics of GroEL clients and non-clients

A total of 7 lines and the control (MG1655) at t0 were chosen for whole genome sequencing, as mentioned above. Comparing the sequences at end-point to their common ancestor at the initiation of the experiment (MG1655 ΔmutS), we found that the most mutations had been accumulated in the line 2.4, which only had approximately one fourth of the expression compared to de WT promoter and also showed the greatest lag time (**Fig. 3.3**). Contrarily, the line 3.1, which had 0.273 times the expression of the WT line, accumulated fewer mutations than any other. The remaining lines showed similar mutation accumulation.

Previous studies have shown that clients of GroEL have higher mutational robustness compared to non-clients. Following our previous GroEL client classification, we distinguish four different classes of GroEL clients, that is, proteins that need GroEL to correctly fold: I) all 252 GroEL/GroES interactors—that is, both facultative and obligate clients (35); II) 85 obligate clients only as defined by Kerner (35); III) 57 obligate clients as defined by Fujiwara (36); and IV) the 34 obligate clients classified (35) that do not depend on GroEL/GroES for solubility; as seen elsewhere (21)

We studied the ratio of mutations accumulated in clients vs. total mutations, to investigate whether even with fewer total mutations those lines that initially overexpressed GroE had accumulated more changes in known GroEL clients in proportion, but no clear trend was found either.

All in all, the sequencing results did not provide significantly different mutation accumulations in relation to groEL expression (Spearman's correlation, two-tailed P= 0.90837) (**Table 3.2**).

Table 3.2 Overview of the mutations accumulated in each of the lines which underwent whole genome sequencing. The different columns refer to each of the lines sequenced. The mutations are classified depending on the type of client of groE Class I through Class IV, the total of mutations in clients is highlighted in bold.

	SPL_Δmut S 2.4 t100	SPL_Δmut S 3.1 t100	SPL_Δmut S 6.1 t100	SPL_Δmut S 0.3 t100	SPL_Δmut S 7.3 t100	SPL_Δmut S 8.1 t100	SPL_Δmut S 8.3 t100
Class I	8	0	2	3	3	7	4
Class II	22	3	8	12	5	4	4
Class III	4	1	8	2	2	2	1
Class IV	8	1	2	1	3	5	1
TOTAL mutations in CLIENTS	**42**	**5**	**20**	**18**	**13**	**18**	**10**
TOTAL mutations	718	168	392	267	303	462	272
Mutations in non-clients	676	163	372	249	290	444	262
ratio client/mutations	6.213	3.067	5.376	7.229	4.483	4.054	3.817

3.3. Discussion

In this study we performed whole-genome sequencing to analyze changes in the mutation accumulation spectra under differential GroEL expression. We found that the expression of the chaperonin system groES-groEL, which is known to increase the resistance of bacteria to mutational insults under strong genetic drift (26), did not necessarily correlate to the mutation accumulation in this scenario, as our results did not provide a clear picture of any significant differences influenced by GroEL. Whole genome sequencing of the remaining clones evolved under the same study would increase our statistical power to infer more established conclusions.

It is known that overexpressing GroE has a high energetic cost for the cell (26). In periodically bottlenecked populations with a large

effect of genetic drift, high levels of GroE expression can be maintained (24). But our results suggest, in general terms, that if this overexpression is too high, it may decrease with time due to the metabolic burden that producing GroEL entails, even under strong bottlenecks (strong genetic drift); while if the expression is too low, it might increase with time as previously described in endosymbionts (28). Furthermore, previous studies had shown that the absence of the overexpression of GroEL led to extinction events in an LTEE (24), however, our study differs from the previous one in a key point: in the former study, overexpression was provided by a plasmid which needed the presence of chloramphenicol to be maintained and arabinose to induce the expression of GroEL, while the present study did not require the use of any antibiotic or alternative carbon source that could have other effects on the evolution (37). However, further work needs to be completed to elucidate this question in a confident manner. Regarding the populational evolution, it would be beneficial to characterize how metabolism is regulated in the two lines evolved. As indicated in previous work, metabolic erosion occurs in the same genetic background with or without overexpression of groE, thereby highlighting the mutational buffering aspect of groEL (38).

Additionally, GroEL's mechanism of buffering may be more complex than previously thought, and the set of client proteins of this chaperone may be larger than currently defined. The list of known GroEL clients is likely to miss many proteins that interact with GroEL in a transient manner (24). Further research is needed to confirm this.

3.4. Materials and methods

3.4.1. Bacterial Lines

Escherichia coli K-12 substr. MG1655 was obtained from I. Matic (INSERM U571, Paris, France) through J. Blazquez's group at Centro Nacional de Biotecnología (CSIC, Madrid, Spain). The mutS deletion strain was obtained in house. This deficiency in MutS (a component of the mismatch repair system of *E. coli* that recognizes and binds to mispaired nucleotides allowing for their removal by further action of MutL and MutH) and has a predicted mutation rate that is 1,000-fold higher than WT (39, 40). Cells were grown routinely in Luria–Bertani (LB) media (Pronadisa #1551; 1% bacto-tryp- tone, 1% NaCl, 0.5% yeast extract).

The acquisition of lines with differential expression levels of *groE*'s operon was based on "*Escherichia coli* strains with promoter libraries constructed by Red/ET recombination pave the way for transcriptional fine-tuning" (41). Using the method described in this paper in conjunction with λ-Red recombineering (42), we generated different promoters that substituted the wild-type *groE* promoter in *E. coli* K-12 MG1655, thereby obtaining several lines with different expression levels of the chaperone system. The primers employed to carry out this genetic engineering were: P1_yjeH (65bp) "TGAAATGTGAGGTGAATCAGGGTTTTCA CCCGATTTTGTGCTGAGTGTAGGCTGGAGCTGCTTC" and P2_SPL_groE (121bp) "TATCTGTTATGGGTGACGCCG GAGCTTACGTGGTTTCCCGGCTGGTGACCNNNNNTATG GGGANGNNNNNNNNNNNNNNNCTTTCAAGGNNTAAN NNNNCATATGAATATCCTCCTTAG" and P2_pBla_groE (119bp) "TATCTGTTATGGGTGACGCCGGAGCTTACGTGG TTTCCCGGCTGGTGACCAGAGAAATTGTCTCATGAGCG

GATACATATTTGAATGTATTTAGAAAAACATATGAATA TCCTCCTTAG". Amplification conditions were: 1x 94ºC 3 min; 30 cycles x 92ºC seconds, 55ºC 30 seconds, 72ºC 30 seconds, followed by a final cycle at 72ºC during 4 minutes; using as the template the pKD3 plasmid (42). Mutants were sequenced by Sanger with the oligo: "GCCGATTGTCTGTTGTGCCC" which targets the Kanamycin resistance gene.

Consequently, the Kan^R cassette was eliminated by recombinase flippase (FLP) recombination mediated by pCP20 (42).

For the knock-out of the gene mutS, a similar protocol was carried out. The primers used for this purpose were: P1_mutS_for (75bp) "TATGAGTGCAATAGAAAATTTCGACGCCCATACGCCCA TGATGCAGCAGTATCTCGTGTAGGCTGGAGCTGCTTCG" and P2_A_mutS (74bp) "GATAGCAAAAGACTATCGGGAA TTGTTATTACACCAGGCTCTTCAAGCGATAAACATATGA ATATCCTCCTTAG". Amplification conditions were: 1x 94ºC 2 minutes; 35 cycles x 92ºC 30 seconds, 60ºC 30 seconds, 68ºC 90 seconds; followed by a final cycle at 68ºC during 4 minutes, using as the template the pKD3 plasmid (42).

Both PCR were treated with DpnI (1h, 37 ºC) to avoid false positives, as this enzyme is sensible to methylation and only processes that DNA replicated by *E. coli*.

3.4.2. Evolution Experiment

Evolution was performed according to two different propagation procedures: 1) Through single-cell bottlenecks in solid media (agar plates; 1.5% Bacto-agar European grade; Pronadisa #1800), or 2) through serial transfer of large bacterial populations in liquid LB

media. A total of 32 bottlenecked and 6 large population lineages were established. We passaged each lineage serially every 24 h by streaking a single colony, in the case of bottlenecked populations, or by sampling 0.1 ml of the large population into 10 ml of fresh media. We passaged all lineages from clonal evolution during 100 days, corresponding to approximately 2.2K generations of the bacteria (~22 generations per day for bottlenecked bacterial lineages and 6.6 generations per day for large populations). Population evolution was carried out during 20 days. A glycerol stock of each lineage was prepared every ten passages and stored at -80 ºC to establish a "bacterial fossil record."

3.4.3. Growth Rate and Fitness Measurements

Growth parameters for bottlenecked lines were evaluated using the Bioscreen C plate-reader system (Oy Growth Curves Ab Ltd, Helsinki, Finland). Each population was precultured overnight at the corresponding temperature and media, and used to inoculate 200 µl of fresh medium (LB) to an OD595 of 0.06–0.07, distributed in 100-well Honeycomb plates. Populations from each time point were grown at least in triplicate in the same plate. Each experimental run was conducted including negative control (*i.e.*, blank fresh media) and positive control (*i.e.*, ancestral lines), and using at least three replicates per 100-well Honeycomb plate. Each run consisted of two Honeycomb plates that allowed testing up to 200 cultures. Positive controls were used to allow for interplate and interrun comparison. Plates were incubated at 37 °C with continuous shaking (medium force) in the Bioscreen C. Growth was monitored for a period of 24–48 h taking OD_{600} measurement every 15 min. Growth rate was determined after OD_{600} normalization following a previous formula (43). Growth rates were

averaged across replicated cultures with 3 replicates. From these data, relative fitness of the clonal lines was determined the ratio of the growth rate of the evolved line versus the growth rate of the ancestor lineage (44).

3.4.4. GroEL Measurement and Quantification by qPCR

RNA extractions were performed using the RNAeasy Mini Kit. Genome wipeout and retrotranscription were performed by using QuantiTect Rev. Transcription Kit (Qiagen). Reference genes were chosen by using the *Geneinvestigator* software (45). Those that have a steady and similar expression to groE's were selected, and the glutathione reductase (b3500) showed the best performance. qPCR was performed in a Applied Biosystem's AB7500 Fast. Data analysis using the provided software by Applied Biosystem. In this experiment, we chose de $2^{-\Delta\Delta Ct}$ method described by Livak (46). The primers used were as follows: for groEL "GGCTGACAAGAAAATCTCCAAC" and "CCGCTGCTGATCATCGCTG". This pair had a 102% of efficiency according to the Applied Biosystems Software. For the reference gene, the primers used were "TGAGCATCTGGATTACAGCAAC" and "CGCGGTGAAAGAGGATTTATAC", with 100% efficiency according to the Applied Biosystems Software.

3.4.5. Deep Sequencing and DNAseq Data Analysis

Whole-genome sequencing (WGS) was performed using paired-end Illumina sequencing at two different sequencing facilities. The

extraction of genomic DNA was performed with the QIAmp DNA mini kit (Qiagen, Venlo [Pays Bas], Germany) for a QiaCube automatic extractor using bacterial pellets obtained from approximately 10ml cultures (from frozen stock or sampled directly from the most recent bacterial culture) of the lineages from bottlenecked populations. Clonal DNAseq libraries were constructed using the TrueSeq DNA polymerase chain reaction-free HT sample preparation kit (Illumina) and labeled with individual indices to allow for running in a single lane. Quality and quantity of clonal DNAseq libraries were assessed using a 2100 Bioanalyzer (Agilent). All clonal DNAseq libraries were paired-end sequenced on an Illumina HiSeq2000 platform using a 2 100 cycles configuration (this service was provided by LifeSequencing SL., Valencia, Spain).

After removing a sample from the culture after passage of interest for freezing, the remaining culture was pelleted. Genomic DNA was extracted from the frozen bacterial cell pellets with the Illustra bacteria genomicPrep mini spin kit (GE Healthcare). DNAseq libraries were constructed with a Nextera XT DNA sample preparation kit (Illumina) and labeled with individual indices to allow for multiplexed running in a single lane. A sequencing run was performed in a MiSeq Benchtop Sequencer using 2 150 bp with 300 cycles pair-end reads configuration. MiSeq sequencing was performed at Valgenetics SL sequencing facility (Valencia, Spain).

The ancestral strain, *E. coli* K-12 substr. MG1655 ΔmutS, was already sequenced and assembled in previous work (24).

Sequencing reads were converted from Illumina quality scores into Sanger quality scores. We then used the breseq v 0.24rc4 (version 4) pipeline (47) for aligning the Illumina reads to our *E. coli*

parental genome and for identifying single nucleotide polymorphisms (SNPs) and indels (using bowtie2 (48)). Individual runs of breseq, with junction prediction disabled but otherwise default parameters, were performed for the ancestral sequence, as well as for each of the evolved lines. Finally, in house scripts were written and run to create tables containing all SNPs and indels for each lineage (including the ancestral), as well as assigning GroEL client information.

3.4.6. Phenotypic arrays

For phenotypic arrays, EcoPlates (EcoPlates: Catalog No. 1506) were employed. Cells were cultured over night at 37 ^{o}C and agitation at 220 rpm. Afterwards, two consecutive washes were completed by centrifuging the cells at 4500 rpm during 15 minutes; using 10 mL Phosphate-buffered saline (PBS) to resuspend the cells after centrifugation. Then cells were resuspended to an OD_{600} of 0.45 in PBS. From this suspension, 150 µL were loaded into the Ecoplates. These were further incubated at 37 ^{o}C, measuring OD_{590} at 0, 4, 12, 24 and 48 h, images of the plates at 24h were collected.

3.4.7. Statistical analysis

To compare the different levels of GroE expression obtained after engineering, a one-way analysis of variance (ANOVA) test was performed. Different groups were generated to perform the comparisons. *P* values were used to assess the statistical significance.

To measure the strength and direction of the relationship between two variables, Spearman's correlation was performed. *P* value was used to assess the statistical significance..

3.5. References

1. Rutherford SL, Lindquist S. Hsp90 as a capacitor for morphological evolution. *Nature*. 1998;396(6709):336-342.

2. Keys DN, Lewis DL, Selegue JE, et al. Recruitment of a hedgehog regulatory circuit in butterfly eyespot evolution. *Science*. 1999;283(5401):532-534.

3. Zákány J, Duboule D. Hox genes in digit development and evolution. *Cell Tissue Res*. 1999;296(1):19-25.

4. Queitsch C, Sangster T a, Lindquist S. Hsp90 as a capacitor of phenotypic variation. *Nature*. 2002;417(6889):618-624.

5. Hayden EJ, Wagner A. Environmental change exposes beneficial epistatic interactions in a catalytic RNA. *Proc Biol Sci*. 2012;279(1742):3418-3425.

6. Tokuriki N, Jackson CJ, Afriat-Jurnou L, Wyganowski KT, Tang R, Tawfik DS. Diminishing returns and tradeoffs constrain the laboratory optimization of an enzyme. *Nat Commun*. 2012;3(1):1257.

7. Visser JAGM, Hermisson J, Wagner GP, et al. Perspective: Evolution and detection of genetic robustness. *Evolution*. 2003;57(9):1959-1972.

8. Gaut BS. Evolution Is an Experiment: Assessing Parallelism in Crop Domestication and Experimental Evolution. *Mol Biol Evol*. 2015;32(7):1661-1671.

9. Todd AE, Orengo CA, Thornton JM. Plasticity of enzyme active sites. *Trends Biochem Sci*. 2002;27(8):419-426.

10. DePristo MA, Weinreich DM, Hartl DL. Missense meanderings in sequence space: a biophysical view of protein evolution. *Nat Rev Genet*. 2005;6(9):678-687.

11. Zeldovich KB, Chen P, Shakhnovich EI. Protein stability imposes limits on organism complexity and speed of molecular evolution. *Proc Natl Acad Sci U S A*. 2007;104(41):16152-16157.

12. Fersht AR. Nucleation mechanisms in protein folding. *Curr Opin Struct Biol*. 1997;7(1):3-9.

13. Winklhofer KF, Tatzelt J, Haass C. The two faces of protein misfolding: gain- and loss-of-function in neurodegenerative diseases. *EMBO J*. 2008;27(2):336-349.

14. Hartl FU. Protein Misfolding Diseases. *Annu Rev Biochem*. 2017;86:21-26.

15. Wang X, Minasov G, Shoichet BK. Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *J Mol Biol*. 2002;320(1):85-95.

16. Tokuriki N, Stricher F, Serrano L, Tawfik DS. How protein stability and new functions trade off. *PLoS Comput Biol*. 2008;4(2).

17. Fromer M, Shifman JM. Tradeoff between stability and multispecificity in the design of promiscuous proteins. *PLoS Comput Biol*. 2009;5(12).

18. Studer RA, Christin PA, Williams MA, Orengo CA. Stability-activity tradeoffs constrain the adaptive evolution of RubisCO. *Proc Natl Acad Sci U S A*. 2014;111(6):2223-2228.

19. Hartl FU, Bracher A, Hayer-Hartl M. Molecular chaperones in protein folding and proteostasis. *Nature*. 2011;475(7356):324-332.

20. Tokuriki N, Tawfik DS. Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature*. 2009;459(7247):668-673.

21. Williams TA, Fares MA. The Effect of Chaperonin Buffering on Protein Evolution. *Genome Biol Evol*. 2010;2(1):609-619.

22. Young JC, Barral JM, Hartl FU. More than folding: Localized functions of cytosolic chaperones. *Trends Biochem Sci*. 2003;28(10):541-547.

23. Rutherford SL. Between genotype and phenotype: protein chaperones and evolvability. *Nat Rev Genet*. 2003;4(4):263-274.

24. Sabater-Muñoz B, Prats-Escriche M, Montagud-Martínez R, et al. Fitness trade-offs determine the role of the molecular chaperonin GroEL in buffering mutations. *Mol Biol Evol*. 2015;32(10):2681-2693.

25. Bogumil D, Dagan T. Cumulative impact of chaperone-mediated folding on genome evolution. *Biochemistry*. 2012;51(50):9941-9953.

26. Fares MA, Ruiz-González MX, Moya A, Elena SF, Barrio E. GroEL buffers against deleterious mutations. *Nature*. 2002;417(6887):398-398.

27. Fares MA, Moya A, Barrio E. Adaptive evolution in GroEL from distantly related endosymbiotic bacteria of insects. *J Evol Biol*. 2005;18(3):651-660.

28. Fares MA, Moya A, Barrio E. GroEL and the maintenance of bacterial endosymbiosis. *Trends in Genetics*. 2004;20(9):413-416.

29. Dorrity MW, Cuperus JT, Carlisle JA, Fields S, Queitsch C. Preferences in a trait decision determined by transcription factor variants. *Proc Natl Acad Sci U S A*. 2018;115(34):E7997-E8006.

30. Iyengar BR, Wagner A. GroEL/S Overexpression Helps to Purge Deleterious Mutations and Reduce Genetic Diversity during Adaptive Protein Evolution. Lu J, ed. *Mol Biol Evol*. 2022;39(6):msac047.

31. Lenski RE. Convergence and Divergence in a Long-Term Experiment with Bacteria. *Am Nat*. 2017;190(S1):S57-S68.

32. Cooper VS, Lenski RE. The population genetics of ecological specialization in evolving Escherichia coli populations. *Nature*. 2000;407(6805):736-739.

33. Wiser MJ, Ribeck N, Lenski RE. Long-Term Dynamics of Adaptation in Asexual Populations. *Science.* 2013;342(6164):1364-1367.

34. Bruger EL, Marx CJ. A decade of genome sequencing has revolutionized studies of experimental evolution. *Curr Opin Microbiol.* 2018;45:149-155.

35. Kerner MJ, Naylor DJ, Ishihama Y, et al. Proteome-wide analysis of chaperonin-dependent protein folding in Escherichia coli. *Cell.* 2005;122(2):209-220.

36. Fujiwara K, Ishihama Y, Nakahigashi K, Soga T, Taguchi H. A systematic survey of in vivo obligate chaperonin-dependent substrates. *EMBO J.* 2010;29(9):1552-1564.

37. Schwarz S, Kehrenberg C, Doublet B, Cloeckaert A. Molecular basis of bacterial resistance to chloramphenicol and florfenicol. *FEMS Microbiol Rev.* 2004;28(5):519-542.

38. Aguilar-Rodríguez J, Fares MA, Wagner A. Chaperonin overproduction and metabolic erosion caused by mutation accumulation in Escherichia coli. *FEMS Microbiol Lett.* 2019;366(10):121.

39. Bjedov I, Dasgupta CN, Slade D, le Blastier S, Selva M, Matic I. Involvement of Escherichia coli DNA Polymerase IV in Tolerance of Cytotoxic Alkylating DNA Lesions in Vivo. *Genetics.* 2007;176(3):1431-1440.

40. Turrientes MC, Baquero F, Levin BR, et al. Normal Mutation Rate Variants Arise in a Mutator (Mut S) Escherichia coli Population. *PLoS One*. 2013;8(9):e72963.

41. Braatsch S, Helmark S, Kranz H, Koebmann B, Jensen PR. Escherichia coli strains with promoter libraries constructed by Red/ET recombination pave the way for transcriptional fine-tuning. *Biotechniques*. 2008;45(3):335-337.

42. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A*. 2000;97(12):6640-6645.

43. Warringer J, Blomberg A. Automated screening in environmental arrays allows analysis of quantitative phenotypic profiles in Saccharomyces cerevisiae. *Yeast*. 2003;20(1):53-67.

44. Lenski RE, Rose MR, Simpson SC, Tadler SC. Long-Term Experimental Evolution in Escherichia coli. I. Adaptation and Divergence During 2,000 Generations. *Am Nat*. 1991;138(6):1315-1341.

45. Hruz T, Laule O, Szabo G, et al. Genevestigator V3: A Reference Expression Database for the Meta-Analysis of Transcriptomes. *Adv Bioinformatics*. 2008;2008:1-5.

46. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2ˆ(-ΔΔCT) method. *Methods*. 2001;25(4):402-408.

47. Deatherage DE, Barrick JE. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol Biol*. 2014;1151:165-188.

48. Philipson C, Davenport K, Voegtly L, et al. Brief Protocol for EDGE Bioinformatics: Analyzing Microbial and Metagenomic NGS Data. *Bio Protoc*. 2017;7(23).

General discussion

Genetically engineered bacteria with novel regulatory circuits (1, 2), heterologous enzymes and pathways, or sensory layers (2) have the potential to revolutionize biotechnology by providing solutions to particular applications and generate new knowledge from a synthetic perspective. The potential applications of genetically engineered bacteria are vast and include *e.g.* the production of pharmaceuticals, the cleaning of environmental pollutants, and the generation of biofuels. However, one overlooked question is whether the resulting bacteria perform as intended in variety of environments or with time. Important issues relate to the operability regime of these genetic modifications, the response ability of bacteria to cope with multiple changes that can occur in a non-predictable way, and the evolutionary stability of the engineered bacteria (3–5). In this thesis, we sought to better understand these aspects in order to identify principles that can be used to predict the effects of different changes and develop strategies to mitigate them.

In this thesis, we have considered a recently engineered genetic system that allows programming the cell to work as a minimal computer (arithmetic logic unit) in order to analyze its operability regime (6). This system involves transcriptional and post-transcriptional regulations. In particular, we have studied the analog behavior of the system, the effect of physicochemical changes in the environment, the impact on cell growth rate of the heterologous expression, and the ability to maintain the arithmetic functioning over time. Conclusively, our results suggest i) that there are wide input concentration ranges that the system can correctly process, the resulting outputs being predictable with a

simple mathematical model, ii) that the engineered circuitry is quite sensitive to temperature effects, iii) that the expression of heterologous small RNAs is costly for the cell, not only of heterologous proteins, and iv) that a proper genetic reorganization of the system to reduce the amount of heterologous DNA in the cell can improve its evolutionary stability.

More in detail, we exploited a pre-characterized transcriptional layer for RNA circuit engineering and utilized RNA folding routines to accurately predict riboregulatory performance from energetic and structural calculations. This allowed us to use a simple mathematical model based on algebraic equations to predict the output responses of different induction conditions. Our results highlight the importance of temperature as an environmental factor that can significantly impact circuit performance due to the interaction between transcriptional and post-transcriptional regulation. To address this, we suggest using a circuit implementation with tightly-regulating transcription factors only in the sensory layer and a pure RNA-based actuation layer, which may help to limit the impact of temperature changes. In addition to temperature, we also found that the expression of heterologous sRNAs can be costly for cells and lead to growth defects, as well as genome-wide transcriptomic perturbations. This suggests that care should be taken when designing RNA circuits to minimize the demand for ribonucleases and avoid disrupting cellular functions. Finally, thanks to our evolutionary experiments, we discovered a trade-off for synthetic RNA circuits, which require high-copy number plasmids or T7 promoters to function properly but can also be costly for cells and prone to accumulating deleterious mutations.

Overall, our results provide a foundation for predicting the effects

of various changes on regulatory RNA circuits and developing strategies to mitigate them. These findings will be useful for researchers and engineers working on the development and application of RNA circuits for biotechnology and related fields (2, 7).

In this thesis, we have also studied how bacterial populations responded in time and space to environments with antimicrobial compounds. In particular, we focused on environments defined by the presence of macroscopic HKUST crystals (*i.e.*, mm-based nanomaterials based on coordinated metals and organic ligands) (8). These materials were shown to release significant amounts of copper in saline medium at a slow rate, which was exploited to study how the growth of a bacterial population was affected. In addition, we analyzed the bacterial response when these nanostructured crystals of observable size were loaded with chloramphenicol to exert a joint metal-antibiotic action, going beyond the traditional oligodynamic effect (a more challenging environment). By performing such studies, we expect to contribute to the development of sustainable, targeted antimicrobial approaches that might be used to effectively control bacterial populations in a range of contexts.

In particular, our study demonstrated the potential of a metal-organic framework (MOF) called HKUST-1 as a means of controlling bacterial populations. We found that MOFs can slowly decompose in biological media, releasing antimicrobials that can act against bacteria. We also showed that MOFs can be loaded with antibiotics, *e.g.*, chloramphenicol, to enhance its antimicrobial action. We analyzed the spatiotemporal bacterial response following a combined experimental and theoretical approach in a such a complex and challenging environment involving metal-based

materials in both liquid and solid media, and found that bacterial densities were greatly reduced. Simple mathematical models of cell growth rate and molecule diffusion explained the data.

Overall, our results suggest new means to control bacterial populations, which could be a promising strategy in further applications, either in the industry or in the clinic (9). Of note, the ability to load the MOFs with antibiotics and other bioactive molecules, as well as the possibility of modifying them post-synthetically, opens up a wide range of possibilities for the development of tailored antimicrobial approaches. However, it is important to note that there are some limitations to this technology. For example, the size of the ligand constrains the types of guest molecules that can be loaded into the material. This may limit the range of bacteria that can be effectively targeted with this approach. Additionally, while we demonstrated the effectiveness of the MOF in controlling bacterial populations in the laboratory, it will be important to test its effectiveness in more complex environments, such as *in vivo*, and also assess if we can move from bacteriostatic to bactericidal effects, that may have better prospects. Our results mainly show a bacteriostatic effect on lab strains, which is only a first step. Future work should be focused on testing the ability of this type of materials (MOFs and also other nanostructured materials) to repress pathogenic bacteria *in vitro* and *in vivo* (10).

Finally, we have decided to study in this thesis how bacteria are able to maintain a given function (or phenotype) at long times, *i.e.*, when the mutational load may be a challenge. To this end, we focused on studying the activity of a key factor in the cell, such as GroEL (11). Of note, GroEL is a protein chaperonin complex that plays a crucial role in maintaining protein folding and function in

cells (12). According to previous work, it has the ability to buffer or purge mutations that may occur in proteins, making it an important factor in cellular evolution and survival (13). As such, we hypothesized that by constructing a library of GroEL expressions and assessing the impact on evolvability we might provide valuable insights into the role that GroEL plays in protein folding and function, as well as its potential role in buffering or purging mutations.

To study the role of GroEL in evolution, we performed an evolutionary experiment. We used a regular strain and a standard culture medium. By analyzing the genetic makeup of the cells at the end of the experiment (here, 100 passages), we can gain insights into the types of mutations that are being selected for and how they may contribute to the evolution of the population. Importantly, we performed whole-genome sequencing to analyze the genetic profile of the different lines. Unfortunately, the results of the whole-genome sequencing were not conclusive and did not provide a clear picture of the types of mutations that were being selected during the experiment as a result of differential GroEL expression. This may be due to a variety of factors, such as that GroEL functions in complex or that there was little to none selective pressure (14). More lines could also be sequenced and analyzed to increase the statistical power.

Despite the lack of conclusive results, our evolutionary experiment and whole-genome sequencing may provide valuable insight into the role of GroEL. Further research is needed to fully understand the role of GroEL in evolution and to identify the specific mechanisms by which it may impact protein folding and function. This information could have important implications for the development of bacterial cells more robust to mutations, thereby

with greater evolutionary stability for applications such as in bioproduction, as well as for our understanding of evolution at the molecular level.

All in all, genetically engineered bacteria could have a significant impact on various industries, the clinic, and further areas of society (15–18). However, it is important to carefully consider the potential risks and benefits of using genetically engineered bacteria, as well as the ethical and social implications of their use. It is also important to ensure that appropriate safety measures are in place to protect human health and the environment (19). This thesis tries to advance our knowledge on how bacteria would respond in a variety of scenarios, once their genetic engineering has been accomplished. On top of that, ecological interactions with other bacteria, either natural or engineered, would also be important (20). Indeed, the prediction of the coexistence of microorganisms in a particular environment is instrumental for several biotechnological applications. A thorough understanding at different levels (molecular, cellular, and ecological) of how bacterial metabolic strategies are encoded and deployed will lead to exciting changes in our society.

References

1. Xiong LL, Garrett MA, Buss MT, Kornfield JA, Shapiro MG. Tunable Temperature-Sensitive Transcriptional Activation Based on Lambda Repressor. *ACS Synth Biol.* 2022;11(7):2518-2522.

2. Sudarsan N, Hammond MC, Block KF, et al. Tandem riboswitch architectures exhibit complex gene control functions. *Science.* 2006;314(5797):300-304.

3. Sleight SC, Bartley BA, Lieviant JA, Sauro HM. Designing and engineering evolutionary robust genetic circuits. *J Biol Eng.* 2010;4(1):1-20.

4. Nuismer SL, Layman NC, Redwood AJ, Chan B, Bull JJ. Methods for measuring the evolutionary stability of engineered genomes to improve their longevity. *Synth Biol.* 2021;6(1): ysab018.

5. Menuhin-Gruman I, Arbel M, Amitay N, et al. Evolutionary Stability Optimizer (ESO): A Novel Approach to Identify and Avoid Mutational Hotspots in DNA Sequences while Maintaining High Expression Levels. *ACS Synth Biol.* 2022;11(3):1142-1151.

6. Rosado A, Cordero T, Rodrigo G. Binary addition in a living cell based on riboregulation. *PLoS Genet.* 2018;14(7): e1007548.

7. Qi LS, Arkin AP. A versatile framework for microbial engineering using synthetic non-coding RNAs. *Nat Rev*

Microbiol. 2014;12(5):341-354.

8. Chui SSY, Lo SMF, Charmant JPH, Orpen AG, Williams ID. A chemically functionalizable nanoporous material [Cu3(TMA)2 (H2O)3](n). *Science.* 1999;283(5405):1148-1150.

9. Sharabati M al, Sabouni R, Husseini GA. Biomedical Applications of Metal−Organic Frameworks for Disease Diagnosis and Drug Delivery: A Review. *Nanomaterials.* 2022;12(2):277.

10. Patra JK, Das G, Fraceto LF, et al. Nano based drug delivery systems: recent developments and future prospects. *J Nanobiotechnology.* 2018;16(1).

11. Sigler PB, Xu Z, Rye HS, Burston SG, Fenton WA, Horwich AL. Structure and function in GroEL-mediated protein folding. *Annu Rev Biochem.* 1998;67(1):581-608.

12. Fares MA, Ruiz-González MX, Moya A, Elena SF, Barrio E. GroEL buffers against deleterious mutations. *Nature.* 2002;417(May):398.

13. Iyengar BR, Wagner A. GroEL/S Overexpression Helps to Purge Deleterious Mutations and Reduce Genetic Diversity during Adaptive Protein Evolution. Lu J, ed. *Mol Biol Evol.* 2022;39(6):msac047.

14. Sabater-Muñoz B, Prats-Escriche M, Montagud-Martínez R, et al. Fitness trade-offs determine the role of the molecular chaperonin GroEL in buffering mutations. *Mol Biol Evol.*

2015;32(10):2681-2693.

15. Singh JS, Abhilash PC, Singh HB, Singh RP, Singh DP. Genetically engineered bacteria: An emerging tool for environmental remediation and future research perspectives. *Gene*. 2011;480(1-2):1-9.

16. Riglar DT, Silver PA. Engineering bacteria for diagnostic and therapeutic applications. *Nat Rev Microbiol*. 2018;16(4):214-225.

17. Piñero-Lambea C, Ruano-Gallego D, Fernández LÁ. Engineered bacteria as therapeutic agents. *Curr Opin Biotechnol*. 2015;35:94-102.

18. Sanghvi G, Thanki A, Pandey S, Singh NK. Engineered bacteria for bioremediation. *Bioremediation of Pollutants*. Elsevier; 2020:359-374.

19. Snow AA, Andow DA, Gepts P, et al. Genetically engineered organisms and the environment: Current status and recommendations. *Ecol Appl*. 2005;15(2):377-404.

20. Bajic D, Sanchez A. The ecology and evolution of microbial metabolic strategies. *Curr Opin Biotechnol*. 2020;62:123-128.

Conclusions

This Ph.D. dissertation is intended to expand our knowledge on how bacteria deploy spatiotemporal responses. That is, bacteria change its gene expression programs with time and also in response to fluctuations in external conditions that take place as result of a different space. In this regard, this thesis has reached the following three main conclusions:

1) We studied how environmental and genetic changes affect the functionality of a synthetic genetic circuit, which is a proxy of a particular phenotype. Our results contribute to a better understanding of the operability regime of engineered circuitries based on regulatory RNAs. Moreover, our results allow disclosing design principles that can be used to predict the effects of perturbations and develop strategies to mitigate them.

2) We studied the bacterial behavior in complex environments. In particular, in media where nanostructured materials appear. This is important because these materials represent an alternative way to fight bacteria and might be widely used in the future. Our results show that bacterial population control can be achieved by the employment of metal-organic frameworks (MOFs) prepared to be observable (size $>$ 0.2 mm). The slow release of antimicrobials, in the form of metals, organic compounds, and antibiotics, allows tuning the control of bacterial growth in liquid and solid media.

3) We studied the contribution of the protein chaperone system GroEL in the channeling of mutations, knowing that previous work pointed out that buffering and purging are apparent opposite effects of this system. In particular, we assessed the impact of

different expression levels of GroEL on the mutational load of bacteria by performing a long experimental evolution. However, our results did not support a clear picture. In this regard, further work is required to disclose the evolutionary role of GroEL in *E. coli*.

Acknowledgments

Special thanks to my thesis supervisors, Guillermo Rodrigo, Beatriz Sabater-Muñoz and Christina Toft, whose help has been indispensable in the fulfillment of this thesis.

The most gratitude to Mario Fares, whose passion for science undoubtedly helped ignite mine.

I would also like to thank all the people I have had the pleasure to work with in Mario Fares' lab, and also Guillermo Rodrigo's lab. Especially, María Prats, Victor Berlanga, Flo Mattenberger, Lucas Goiriz, María Heras, Roswitha Dolcemascolo, Raúl Ruiz, Alejandro Requena, and Rosa Márquez; who have become more than colleagues, being proud of calling them friends.

And of course, many thanks to my family, especially Jaume, who have been incredibly supportive throughout the years of my adventures in science.