

NEWCASTLE UNIVERSITY



DOCTORAL THESIS

**A unified platform for experimental
and computational biology**

Charles WINTERHALTER

*A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy*

in the

School of Computing Science
Institute for Cell and Molecular Biosciences

January 2019

Declaration of Authorship

I, Charles WINTERHALTER, declare that this thesis titled, "A unified platform for experimental and computational biology" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: *Charles Winterhalter*



Date: *January 13, 2019*

"It would not be much of a universe if it wasn't home to the people you love."

Stephen Hawking

"A man must for years examine for himself great piles of superimposed strata, and watch the sea at work grinding down old rocks and making fresh sediment, before he can hope to comprehend anything of the lapse of time, the monuments of which we see around us."

Charles Darwin (1859)

Abstract

Charles WINTERHALTER

A unified platform for experimental and computational biology

In natural sciences, the correct engineering of a system's chemical, biological and physical properties may allow it to sustain life. Bioengineering cells is probably one of the most complex challenges of biological research; yet, the little we do know about the nature of life is sufficient to guide scientific research, and to explore the elements beyond the apparent simple proliferation of living cells. Although Mendel first characterised the concept of genetic heredity over 150 years ago, we only recently became able to perform tailored genetic modification of living organisms. The development of digital technologies, in particular, has positively influenced the quality and reproducibility of experimental results emerging from biological assays. However, the use of any equipment may require the need for a specific expertise in order to perform a given experimental procedure. Therefore, multidisciplinary research can bring benefits to all fields of science by helping the development of analytical methods that cross the boundaries of individual disciplines. This emerges as a systematic view of scientific problems, and relies on the adequation and integration of results from different research areas. Nevertheless, there is a complex interface between hard sciences that often creates a gap between experimental and theoretical models.

In this thesis, we explored synthetic biology approaches and created a unified platform to fill this gap. We propose the first barcoding platform (Bac2code) that allows the identification and the tracking of bacterial strains. In order to facilitate communication between researchers, we developed a barcode system in DNA that physically links bacteria to their genetic description. We designed DNA barcodes as bioorthogonal elements, elaborated a universal cloning strategy to integrate these sequences in Gram-negative and Gram-positive bacteria, and demonstrated their stability over time. Through a generic protocol, any barcoded strain can later be identified via a single sequencing read. With the engineering of a synthetic circuit library, we built a biorepository of genetic constructs for our barcoding platform. These biological devices were optimised based on the closest achievable interface between experimental biology and

computational results. Following their characterisation, and in the context of intercellular communication, we studied the behaviour of small cohorts of bioengineered cells at the microscale in microfluidics. We pushed the biological and physical boundaries of engineering techniques to the maximum, in order to observe physiological changes between bacteria separated by distances down to $20\mu m$. However, we also showed that we reached a technological barrier, where even the use of nanoscale features was found insufficient to maintain cells isolated under high cellular density. Yet, microfluidics remains a remarkable technology, and we propose the expansion of barcoding methods to automated systems, which would allow serial barcode integration and documentation retrieval at any one time.

Here, we developed and tested a barcoding method to ensure the cohesion of experimental and computational biology resources. We demonstrated its use by the *in vitro* assembly and the *in vivo* or *in silico* characterisation of a series of genetic circuits via different techniques. The research output of this thesis is realised as a step forward in interdisciplinary studies, and is now being adapted to reach a larger community of users as a startup company.

Acknowledgements

I would like to thank my supervisor Natalio Krasnogor for showing me the joys of a PhD. For her technical help and friendship throughout this experience, I would also like to acknowledge my second supervisor Birgit Koch. For their support and insightful discussions, I shall as well acknowledge Heath Murray, Philip Aldridge, Henrik Strahl, Yulia Yuzenkova and Pamela Gamba. I wish to thank Newcastle University for providing me the means to carry out cutting edge research in a currently unstable political context, and the Engineering and Physical Sciences Research Council (EPSRC, grants EP/L001489/1 and EP/N031962/1) and Newcastle University's School of Computing Science for the funding of my research.

None of this work would have been possible without the constant support of my family. Although they are one thousand miles away, Anne, Bernard, Chloé, Célia, Arnaud and Pierre showed me that love has no boundaries; they were the sun that lightened my rainy days. In particular, I would like to thank my parents Anne and Bernard, who have always been incredibly supportive and led me to where I stand now.

For making Newcastle such a lovely city to live in, I would especially like to thank my English brother Calum and the Daves (Forrest and Barnacle). I would also like to thank Jurek, who was always there at the right time, and Tom, who supported me when I needed. Without any doubt, I would not have reached this point if it was not for the motivation I received from William, Boris and Guillaume.

Finally, last but not least, I would like to acknowledge my better half Amy. It would not have been the same without her, and no one could have shown me as much support as she did. I got used to my "one puppy a day" encouragement pictures whilst writing up, and I don't know what I will do now this thesis is written. Life has shown me that the best things happen unexpectedly. Getting a PhD is one thing, but I did not expect that it would also show me the one that always makes me look forward to tomorrow.

Contents

Declaration of Authorship	iii
Abstract	vii
Acknowledgements	ix
Contents	xi
1 Introduction	1
1.1 Design in synthetic biology	4
1.1.1 Cellular stochasticity	4
1.1.2 Host/circuit compatibility	6
1.2 Thesis problem statement and objectives	7
1.3 Plan	10
1.4 Contributions to this thesis	11
2 Background and genetic library engineering	13
2.1 Engineering the genetic code	13
2.1.1 DNA	14
2.1.2 RNA	15
2.1.3 Proteins	16
2.2 Circuit replication	19
2.2.1 Replicative plasmids	19
2.2.2 Selective pressure	19
2.2.3 Chromosomal recombination	21
2.3 Transcription machinery	21
2.3.1 Promoters	21
2.3.1.1 Constitutive promoters	22
2.3.1.2 Arabinose-dependent promoter	22
2.3.1.3 Quorum-sensing promoter	23
2.3.2 Terminators	23

2.4	Translation machinery	24
2.4.1	Ribosomes	24
2.4.2	Ribosome binding sites	27
2.4.3	Transcription factors	27
2.4.4	Protein degradation tags	27
2.5	Fluorescence integration	28
2.5.1	Fluorescent reporter proteins	28
2.5.2	Photostability	30
2.6	Genetic circuits built in this study	31
2.6.1	Microbiology techniques	31
2.6.2	Plasmids	31
2.6.3	Strains	37
3	DNA barcoding in bacteria	43
3.1	Filling the gap between <i>in silico</i> and <i>in vitro/in vivo</i> biology	43
3.1.1	Sequencing technologies	44
3.1.2	A biological version control system	44
3.1.3	Online documentation platform: Bac2code	45
3.1.4	Barcode minimal identifiers	45
3.2	Bacterial referencing via DNA barcodes	46
3.2.1	DNA encoding	48
3.2.1.1	Barcode sequences	48
3.2.1.2	Bio-orthogonality	49
3.2.1.3	Additional barcode features	51
	Universal primer	51
	Synchronisation sequence	52
	Checksum	52
3.2.2	Cloning of DNA barcodes	52
3.2.2.1	Considerations related to the design and use of DNA barcodes	52
3.2.2.2	Overview of the cloning method	53
3.2.2.3	Finding new cloning sites	57
3.2.2.4	<i>In vitro</i> assembly of recombinant DNA	59
3.2.2.5	<i>In vivo</i> integration protocols	62
3.2.3	Retrieval of DNA barcodes	62
3.3	Large-scale screening setup	62
3.3.1	Continuous flow systems	63

3.3.2	High-throughput screening	68
3.4	DNA barcodes <i>in vivo</i> characterisation	70
3.4.1	Cloning method	70
3.4.2	Growth characteristics	72
3.4.2.1	Spot-plating assay	72
3.4.2.2	CFUs counts	72
3.4.3	Barcodes stability over long periods of continuous flow culture	72
3.4.4	Barcodes stability in high-throughput subcultures	74
3.4.5	Barcodes readability	76
3.5	Considerations in the study of DNA barcodes	80
3.5.1	Barcode integration	80
3.5.1.1	Setup of a universal cloning method	80
3.5.1.2	Extending the barcoding technique	81
3.5.2	Large scale characterisation studies	82
3.5.2.1	Continuous flow system	82
3.5.2.2	High-throughput assays	83
3.5.3	Barcode design and sequencing data	84
3.6	Summary	84
4	Genetic circuits engineering	87
4.1	Bacterial communication via quorum sensing	87
4.1.1	Quorum signaling molecules: the <i>lux</i> model	88
4.1.1.1	LuxR regulator	88
4.1.1.2	LuxI autoinducer	89
4.1.2	Inducible quorum devices	89
4.2	Construction of genetic circuits	90
4.2.1	Engineering the <i>lux</i> operon	90
4.2.2	Genetic circuits library	91
4.2.2.1	Setup of bacterial terminators	91
4.2.2.2	Choosing adequate ribosome binding sites	91
4.2.2.3	Inducing quorum molecule production	93
4.2.2.4	Detecting quorum molecules	95
4.2.2.5	Cotranscribed fluorophores	95
4.3	Single cell and population scale <i>in vivo</i> characterisation	97
4.3.1	Production of quorum compound	97
4.3.1.1	P_{TET} and P_{BAD} promoters	97
4.3.1.2	<i>E. coli</i> expression strains	100

4.3.1.3	Continuous time-lapse of quorum circuit activation . . .	102
4.3.1.4	Testing for AHL production	103
4.3.2	Quorum sensing response	106
4.3.2.1	LuxR regulation	107
4.3.2.2	Induction via sender genetic circuits	111
4.3.2.3	Detector/amplifier circuits	112
4.4	Mathematical modelling	113
4.4.1	Non-linearity in biology	113
4.4.2	Modelling via the Hill equation	114
4.4.3	Data normalisation	115
4.4.4	Quorum sender devices modelling	116
4.4.5	Sensor strains modelling	116
4.4.6	Transfer function modelling	118
4.4.7	Optimised bioengineered circuits	120
4.5	Summary	123
5	The making of microfluidics devices for bacterial screening	127
5.1	Introduction to microfluidics studies	127
5.1.1	The use of microfluidics chips with bacteria	128
5.1.2	Small world effect	128
5.1.3	PDMS substrates	129
5.1.4	Scaling up bacterial screening	129
5.2	Design of biochips	130
5.2.1	Photolithography considerations	131
5.2.2	Microfluidics chip compartments	133
5.2.2.1	Physical considerations associated with PDMS structures	134
5.2.2.2	Observation chambers	135
5.2.2.3	Channels	135
5.2.3	Fluid dynamics	137
5.2.3.1	Node analysis	137
5.2.3.2	Finite element modelling	138
5.2.4	The making of biochips	138
5.3	Channel-type chips	139
5.3.1	Design	139
5.3.1.1	Cellular flow	139
5.3.1.2	Unique identifiers	139
5.3.2	Fabrication of PDMS devices	141

5.3.2.1	SEM screening	141
5.3.2.2	Adequate flow verification	144
5.3.3	<i>In vivo</i> studies	144
5.3.3.1	Cell size control	144
5.3.3.2	Limiting design factors	147
5.4	Surface-type chips	148
5.4.1	Design features	149
5.4.1.1	Improving biological fitness	149
5.4.1.2	Creating thinner diffusion features	149
5.4.1.3	Correcting design by fluid dynamics	150
5.4.2	Characterisation of PDMS devices	153
5.4.2.1	SEM screening	153
5.4.2.2	Liquid flow check	157
5.4.3	<i>In vivo</i> cellular fitness	161
5.4.3.1	Cellular loading	161
5.4.3.2	Bacterial growth stability over time	163
5.5	Intercellular communication in microfluidics	165
5.5.1	Microscale reaction-diffusion systems	165
5.5.1.1	PDMS device fabrication	165
5.5.1.2	<i>E. coli</i> growth in observation chambers	166
5.5.1.3	Diffusion system limitations	168
5.5.2	Wiring and controlling bacterial communication	168
5.5.2.1	Quorum signal repression	169
5.5.2.2	Pulse generator devices	169
5.6	Summary	171
6	Integrating bioengineering approaches in synthetic biology	173
6.1	Bac2code online barcoding platform	173
6.1.1	Possible actions on the Bac2code website	174
6.1.2	Integration of <i>in silico</i> resources	177
6.1.3	Integration of barcoded strain specifications	177
6.2	Barcoding in microfluidics	179
7	Concluding remarks and further prospects	181
A	Biological protocols	183

B	Recombining barcode DNA in <i>B. subtilis</i> and <i>E. coli</i>	189
B.1	Transformation in <i>B. subtilis</i>	189
B.2	Transformation in <i>E. coli</i>	190
C	DNA barcoding kit	193
C.1	Assembling recombinant barcode DNA	194
C.1.1	Primer-free PCR assembly	194
C.1.2	Recombinant DNA amplification	195
C.1.3	Purification	195
C.2	Transforming <i>B. subtilis</i> with recombinant barcodes	195
C.2.1	Barcoding cells	196
C.2.2	Removing the selection marker	196
C.3	Transforming <i>E. coli</i> with recombinant barcodes	196
C.3.1	Propagate the helper plasmid	197
C.3.2	Barcoding cells	198
C.3.3	Removing the selection marker	198
C.4	Retrieving DNA barcodes	199
C.4.1	Harvest bacterial genomic DNA	200
C.4.2	Barcode DNA amplification	200
C.4.3	Purification and barcode retrieval	200
C.5	Troubleshooting	200
D	Making PDMS microfluidics chips	203
D.1	Intermediate-PDMS template	203
D.2	Cleaning microfluidics wafers	205
D.3	Single-use soft-PDMS	206
	Bibliography	207

List of Figures

1.1	Simplified model of bioengineering studies in synthetic biology	3
1.2	Analogy between electrical and biological circuits	5
1.3	Barcoding platform engineered in this study	9
2.1	The central dogma of molecular biology	14
2.2	Biochemical structure of DNA	15
2.3	The double-helical structure of DNA	16
2.4	Structure of the alanine tRNA	17
2.5	Proteins maturation	18
2.6	Expression states of the arabinose promoter	23
2.7	<i>V. fischeri lux</i> operon	24
2.8	<i>E. coli rrnB</i> T1 and T7 bacteriophage terminators fusion	25
2.9	Ribosome structure and mechanism of action	26
2.10	Automated plate readers	29
2.11	Different common fluorescent protein UV spectra	30
3.1	Barcoding cells	44
3.2	Bac2code online platform	46
3.3	Considerations for DNA barcodes design	47
3.4	Different types of barcodes	48
3.5	DNA barcodes genetic structure	49
3.6	DNA barcodes redundancy	51
3.7	Method to barcoding bacteria	54
3.8	Single and double crossovers events	55
3.9	Schematics of the SOE-PCR	56
3.10	Integration sites for barcodes cloning	58
3.11	Standard <i>E. coli</i> growth curve	64
3.12	Setup of a chemostat	66
3.13	Chemostat assay overview	67
3.14	<i>B. subtilis</i> growth on high throughput equipment	69

3.15	Barcode sequence retrieval	71
3.16	Plate-spotting assay comparing wild-type vs. barcoded bacterial strains.	73
3.17	Minimal colony forming units per barcoded strain	74
3.18	OD measurements in continuous culture	75
3.19	High-throughput screening of barcoded <i>B. subtilis</i>	77
3.20	High-throughput screening of barcoded <i>E. coli</i>	78
3.21	Barcodes alignment	79
3.22	Peristaltic pump schematics	83
4.1	Genetic circuits assembly	90
4.2	Deconvoluting the <i>lux</i> operon	92
4.3	Plasmid templates	93
4.4	Ribosome binding sites characterisation	94
4.5	P_{TET} and P_{BAD} inducible promoters	94
4.6	Robust green/red fluorescence quorum emission/detection	96
4.7	Uninduced sender genetic circuits	98
4.8	Induced sender genetic circuits	99
4.9	P_{TET} and P_{BAD} sender devices in flow cytometry	100
4.10	Flow cytometry fluorescence levels in <i>E. coli</i> MG1655 and BW25113 . . .	101
4.11	Averaged individual sender fluorescence in <i>E. coli</i> BW25113 and MG1655	102
4.12	Reporting AHL production via red fluorescence	104
4.13	Reporting AHL production via green fluorescence	105
4.14	Bioluminescence control plasmids	105
4.15	Testing for production of AHL	106
4.16	Reporting AHL detection via red fluorescence	108
4.17	Flow cytometry fluorescence levels in receiver genetic circuits	109
4.18	Reporting AHL detection via differential LuxR regulatory	110
4.19	Bioluminescence induced by sender circuits	111
4.20	Cotransformation of sender and receiver circuits	112
4.21	Hill abstraction in biological models	113
4.22	Sender genetic circuits modelling	117
4.23	Sensor genetic circuits modelling	119
4.24	Sender genetic circuits transfer functions	121
4.25	Receiver genetic circuits transfer functions	122
4.26	Fold inductions of genetic circuits	123
4.27	Most sensitive AHL sender/detector strains	124

5.1	Intercellular communication in small microbial cohorts	130
5.2	Reynolds number and impact on fluids flow	131
5.3	Silicon wafer mask alignment marks	132
5.4	Microfluidics chips design approach	133
5.5	Producing microfluidics chips	134
5.6	<i>E. coli</i> growth in a 2D-chamber	136
5.7	Standard microfluidics fluids analysis methods	137
5.8	Schematics of channel-type microfluidics chips	140
5.9	Scanning Electron Microscopy (SEM)	142
5.10	Channel-type SEM results overview	143
5.11	Channel-type microfluidics chips flow control	145
5.12	Cells in channel-type biochips	146
5.13	Overgrowth in channel-type biochips	148
5.14	Surface microfluidics chips design	151
5.15	Surface chips deep flow channels simulation	152
5.16	Overview of surface microfluidics chips flow simulation	154
5.17	Optimisation of flow dynamics near bacterial traps	155
5.18	Surface microfluidics chips in the SEM observation chamber	156
5.19	Surface microfluidics chips PDMS defects	157
5.20	Chip CWqs21 SEM overview	158
5.21	Surface microfluidics chips SEM overview	159
5.22	Surface microfluidics chip fluid dynamics colorimetric test	160
5.23	Surface microfluidics chips in culture medium	161
5.24	Surface microfluidics chips cellular loading	162
5.25	Surface microfluidics chips in time course assays	164
5.26	Surface biochips <i>in vivo</i> limitations	167
5.27	Pulse generator simulations	170
6.1	Bac2code typical users and actions	175
6.2	Barcoding with Bac2code	176
6.3	Retrieval of barcoded strain profile on Bac2code	176
6.4	Batch orders of DNA barcodes	177

List of Tables

2.1	Plasmids used in this thesis	32
2.2	Strains used in this thesis	38
3.1	Mapping between hexadecimal characters and nucleotide triplets for the encoding of DNA barcodes	50
3.2	Barcodes cloning sites in <i>E. coli</i> and <i>B. subtilis</i>	57
3.3	Primers used for barcoding	60
3.4	Barcoded strains growth rate in <i>E. coli</i> and <i>B. subtilis</i> compared to wild- type strains.	73
4.1	Fluorescent proteins used as genetic circuits reporter.	95
4.2	Genetic circuits fitting in non linear regression	118

List of Abbreviations

α	alpha
β	beta
λ	lambda
μ	micro
μF	microfarad
μg	microgram
μl	microlitre
μm	micrometre
Ω	ohms
σ	sigma
%	percent
1D	one Dimension
2D	two Dimensions
3D	three Dimensions
A	Adenine
SAM	S -adenosyl/methionine
acyl-HSL	acyl-Homoserine Lactones
AHL	<i>N</i> -Acyl-Homoserine Lactone
aTc	anhydro Tetracycline
bp	base pair
C	Cytosine
C_{α}	carbon alpha
CAD	Computer-Aided Design
CFU	Colony Forming Unit
COOH	carbon hydroxyl group
<i>cf.</i>	<i>confer</i>
$^{\circ}C$	degree Celsius
DNA	Deoxyribo-Nucleic Acid
dNTP	deoxynucleotide
dsDNA	double-stranded DNA
EAN	European Article Number
EDTA	Ethylene-Diamine Tetraacetic Acid
EM-CCD	Electron Multiplying Charge Coupled Device
<i>e.g.</i>	<i>exempli gratia</i>
EtBr	Ethidium Bromide
fmol	femtomole
g	grams

G	Guanine
gDNA	genomic DNA
GFP	Green Fluorescence Protein
h	hours
h-PDMS	hard PDMS
HTSF	High Throughput Screening Facility
kb	kilobase
K_D	dissociation constant
kV	kilovolt
L	litres
LB	Luria Bertani
Mb	Megabase
M	Molarity
Mg₂SO₄·7H₂O	magnesium sulphate
ml	millilitre
mm	millimetre
mM	millimolar
MM	Minimal Medium
min	minutes
mRNA	messenger-RNA
ms	millisecond
mTor	millitorr
NA	Nutrient Agar
NASA	National Aeronautics and Space Administration
NEB	New England Biolabs
ng	nanogram
NGS	Next Generation Sequencing
nm	nanometre
nM	nanomolar
NaCl	sodium chloride
NaOH	sodium hydroxide
NH₂	amine group
OD	Optical Density
OD_{600}	OD at 600nm
PCR	Polymerase Chain Reaction
PDMS	Polydimethylsiloxane
PPE	Personal Protective Equipment
psi	pounds per square inch
RBS	Ribosome Binding Site
RFU	Relative Fluorescence Units
RNA	Ribo-Nucleic Acid
RNAP	RNA polymerase
rpm	rotations per minute
rRNA	ribosomal RNA
SBOL	Synthetic Biology Open Language

SEM	Scanning Electron Microscopy
sec	seconds
SM	Starvation Medium
SMM	Spizizen Minimal Medium
SOB	Super Optimal Broth
SOC	Super Optimal broth with Catabolite repression
SOE-PCR	Splicing Overlap with Extension PCR
s-PDMS	soft PDMS
sRNA	small RNA
T	Thymine
TAE	Tris Acetate EDTA
TBE	Tris Borate EDTA
TE	Tris EDTA
Tris-HCl	Tris hydrochloride
tRNA	transfer-RNA
U	Uracyl
UV	Ultra Violet
UPC	Universal Product Code
UUID	Universally Unique Identifier
vs.	versus
w/v	wet volume
x	times

Pour mes parents Anne et Bernard
Winterhalter, à qui je dois tout.

Chapter 1

Introduction

This thesis is focussed on the unseen majority that evolves in our environment: bacteria. The huge number of these unicellular organisms surrounding us, invisible to the naked eye, is essential for the proliferation of life on Earth. Bacteria may affect other species and either be beneficial, detrimental, or both, depending on a specific biological context[229]. Regardless of their effect, all bacteria follow the same basic biological rules; they display a certain behaviour based on the molecular processes that are directed by the central dogma of biology (conversion of deoxyribonucleic acid (DNA) to ribonucleic acid (RNA) and proteins, described in Chapter 2)[39]. For instance, in common bacterial cells such as *Escherichia coli*, there are over 4,000 proteins that form a complex interaction network, based on information encoded from a genome of ~ 4.6 million nucleotides (individual bases of DNA). For a single cell, that is a lot to handle. Yet, a long time before humans even populated the Earth, bacterial strains have evolved and developed techniques to adapt their growth to an enormous range of environments[184]. Humans, in turn, later started developing techniques to engineer bacteria for specific applications. Although we have been using bacteria for centuries, we did not appreciate until recently the complexity that lies behind their various, seemingly simple processes. For instance, the maturation of milk is coordinated with the growth of a set of bacterial species, and these bacteria are now of major importance for dairy industries[64]. Over time, scientists have developed a thorough understanding of the cellular mechanisms that happen at the microscale, and it is nowadays possible to program cells with a specific cellular behaviour, based on the engineering of the genetic code and environmental factors.

In 1953, Watson and Crick published the structure of DNA, but there was still at the time a tremendous amount of biological processes to be understood[236]. However, scientists had already started thinking about cells as small devices that carry out specific functions. For instance, the physicist Feynman postulated in 1960 that cells are analogous to machines and that they are, at a small scale, capable of an impressive

amount of operations[66, 159, 221]. This vision suggested that cells may act as molecular machines, which undertake a certain set of actions in response to different stimuli in order to produce a variable output. It was noneless 30 years later, with the development of sequencing technologies, that synthetic biology emerged as a discipline of real interest in the field of biological research. Synthetic biology aims to create robust biological circuits, with a well-known function and predictable behaviour[27, 28]. It includes the development of computational tools that help the understanding of experimental processes, and the optimisation of biological device functions. In fact, most synthetic biology studies are driven by the creation of tools to solve variable problems[33, 168]. Many applications are aimed towards the production of therapeutics to cure various conditions or diseases affecting humankind, but some other studies, for instance the terraformation project led by the National Aeronautics and Space Administration (NASA), use bacteria to transform the outer space into potential Earth-like habitats. There are thus a plethora of experimental techniques available for genetic modification, which may be used to engineer bacteria for completely different kinds of objectives. Altogether, the engineering of novel synthetic devices is generally directed towards getting a better understanding of life, or to have a specific use in a production process such as pharmaceuticals, or biofuel companies. However, there are a number of restrictions that emerge from the nature of living cells[24, 218]. Therefore, as shown in Figure 1.1, the engineering of a biological device is often a cyclic process, where several cycle iterations need repeating before obtaining well-behaved circuits.

In the context of synthetic biology, the development of description and analysis tools should aid the rapid engineering of novel synthetic devices. However, this type of multidisciplinary research involves a profound understanding of the elementary genetic elements of bacteria, and often requires circuit optimisation to obtain robust constructs *in vivo*. On one side, computational studies should help the modelling and prediction of physiological changes; these are hidden behind stochastic noise represented by living microorganisms and their associated subcellular processes. On the other side, biologically engineered constructs should be portable and also provide replicable behaviour in various genomic context. Throughout this thesis, we adopted these principles and developed a series of genetic devices and screening platforms, in order to facilitate the connection between experimental and computational biology. In the following section, we provide more details about the context of this study: the design of artificial circuits in synthetic biology.

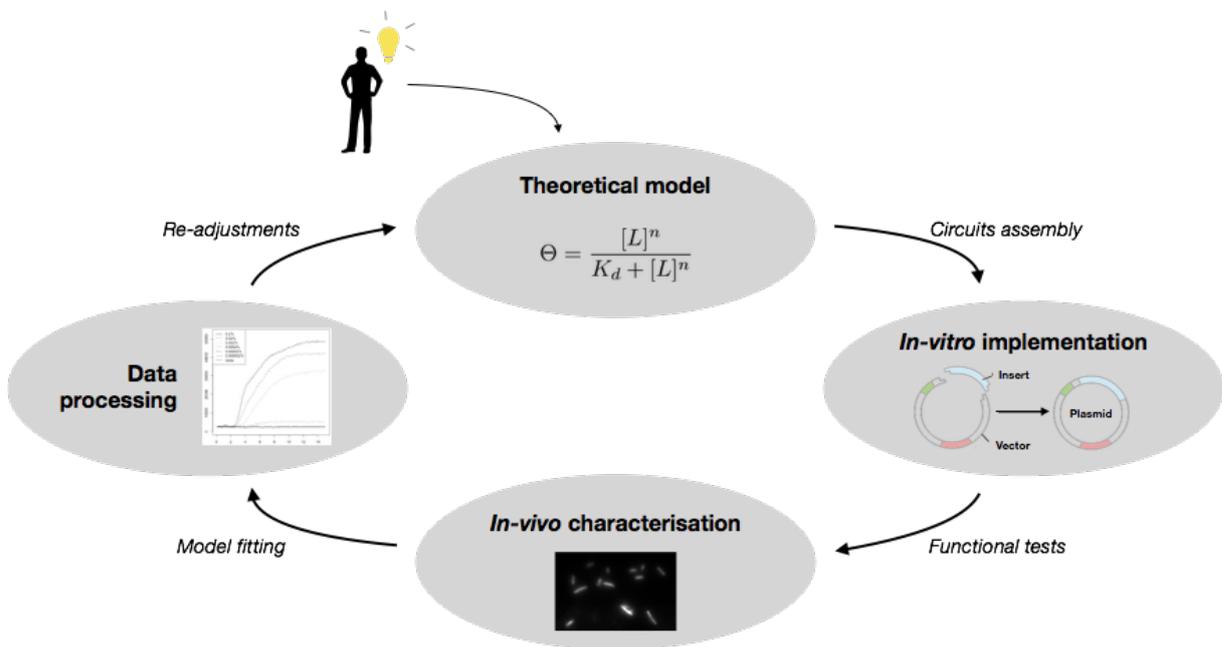


FIGURE 1.1: The model starts with a theory over the circuit design and mathematical rules dictating its behaviour. Theoretical models can be tested in the wet lab: first, parts necessary for the circuit are functionally assembled, then the circuit behaviour is followed in living cells. Fitting between experimental data and theoretical model helps re-evaluating the circuit and improving it in other iterations of this engineering process.

1.1 Design in synthetic biology

Synthetic biology and computer science are two intertwined disciplines. This is because bioengineered cells can display the same characteristics as micro-/nanocomputers. Specific algorithms are associated with different cellular mechanisms, and genetic programming becomes feasible by changing modular subunits within molecular devices[218, 185, 123, 6]. To exemplify this approach, we use the analogy between a simple electrical circuit and a light production device in bacteria (Figure 1.2). On the electric board, if the switch is turned on, the light bulb should receive current generated by the battery and start emitting light. If no light is observed, then it is easy to either change the light bulb or the battery, which are the most likely causes of fault. In synthetic biology, we build modular devices that can be adapted to reproduce similar systems. There should thus be known components and methodological controls that allow the troubleshooting of any uprising issue (no light emission protein observed for instance). Since individual cells are the tiny batteries powering genetic devices, it is important to grow them in optimal conditions in order to provide these circuits reliable environmental conditions. Genetic modification of specific circuit transcription or translation units can then be compared to replacing a light bulb, but at the molecular level. Therefore, there are key parameters that allow for the construction of stable biological devices, and optimising one element is analogous to one or multiple cycles through the development of synthetic circuits presented in Figure 1.1. By the end of this optimisation process, biological devices should be thoroughly tested, well-characterised and show a good response to changes in the environment.

1.1.1 Cellular stochasticity

When studying genetic circuits in a cellular context, there are a lot of processes operating at the same time as the specific device functions. Therefore, there are constant perturbations within cells that restrict our ability to obtain precise parameters over synthetic circuit behaviour[167]. Without the use of external methods of analysis, biological research would be like looking for a needle in a haystack. Fortunately, the development of characterisation methods for single cell or population scale measurements facilitated our understanding of biological processes. As displayed in Figure 1.2B, one can use intrinsic biological system properties to activate light emission in bacteria. Actually, this is a common approach in synthetic biology, where a reporter

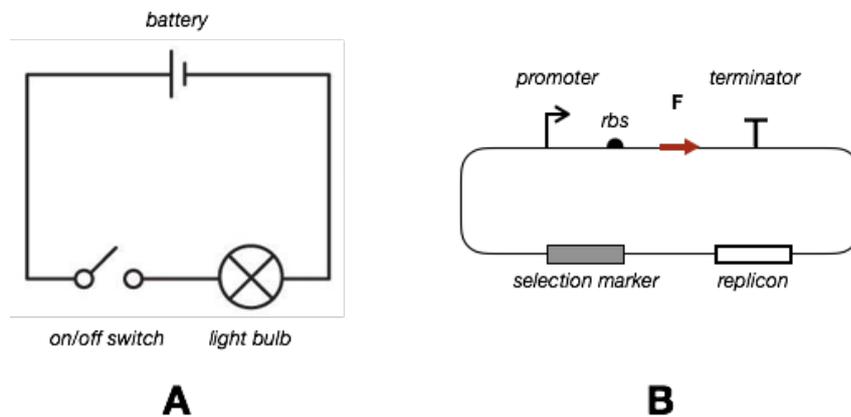


FIGURE 1.2: (A) A simple classic electrical system lighting up a bulb upon activation of a switch if wired through a power supply (battery). (B) A generic biological circuit, or plasmid, replicating itself with a replicon and selected for with a specific marker. This plasmid drives the expression of a protein of interest (fluorescent reporter F) through transcription (from the promoter to the terminator) and translation (via ribosomes recognising a ribosome binding site, or rbs).

provides variable illumination given different cellular states. Since bacteria are individual microscopic organisms, cellular signals often need amplification to be perceived, and the use of fluorescent reporters allows an easy detection of physiological changes via a range of qualitative and quantitative methods[112]. External tools thus enable us to analyse and to interpret changes invisible to the naked eye, but resulting from significant biological activity performed at the microscale. These physiological changes later need to be compared to the proposed theoretical model. Nevertheless, the background noise behind experimental data may sometimes impair with the use of automated analysis methods, and it makes the matching of experimental data to computational models a trickier process.

Bacteria represent a stochastic environment: cellular responses may vary from experiment to experiment, and sometimes lead to false-positive and false-negative results. This highlights the importance of performing biological replicates in order to obtain meaningful data. However, the complementation of biological studies with *in silico* resources usually facilitates the identification of true-positives. For instance, fluorescence measurements derived from *in vivo* experiments can be analysed, modelled by mathematical models and compared to theoretical results in order to improve genetic devices[176, 128]. Generally, this is achieved with a low modelling level of abstraction, better to model biological processes, that is based on simulating the main features of the central dogma of biology for a specific function (cf. Chapter 2). In practice, this

allows us to overcome the noise imposed by millions of surrounding molecules, and to truly characterise a device function. Yet, the modelling process of synthetic constructs is sometimes context specific, and may not provide an exact prediction of the observed physiological changes in different conditions. Thus, the documentation of *in silico* resources linked to biological devices should account for specific features, which allow the easier characterisation of novel circuits.

1.1.2 Host/circuit compatibility

In the microscopic world, any individual that does not fit the environment is quickly eliminated by natural selection. In all biological assays, cells can sustain a specific metabolic load, which is directly connected to their growth profile. Therefore, in synthetic biology, any defect in bacterial growth is usually synonym of an increased metabolic load, resulting from the functionality of a genetic circuit creating crosstalk between host and specifically encoded elements[79, 185]. Back to the analogy between genetic and electrical circuits presented in Figure 1.2, leaving the light on at all times would likely reduce the life time of a light bulb, besides leaving a salty bill. To avoid this situation at the molecular level, we thus try to limit the impact that genetic devices may display on cellular metabolism, since it is the same metabolism that is also responsible for the functionality of a given construct. Hence, in the development of any biological circuit, it is of utmost importance to be aware of its potential metabolic load, and to control the impact it may create on other specific cellular processes.

So far, we have provided an overview of biological circuit design in synthetic biology, and explained how cells may be programmed to perform different kinds of actions. Differential behaviour is obtained by the setup of an ordered nucleotide sequence that contains regulatory regions, required for a specific biochemical phenomenon to take place *in vitro/in vivo*. However, many genome and DNA sequences are of unknown function, and some may be deleted without altering the bacterial fitness. In contrast, attempting to remove other unknown regions may also be detrimental to the overall metabolism. Although, *per se*, the succession of individual nucleotides may not provide direct information about its function, it still follows generic patterns that are recognised and used by the cellular machinery. Therefore, a possibility to minimise the potential for interactions between synthetic circuits and host metabolism is to design bioorthogonal circuits[129]. This consists of engineering a nucleic acid sequence that displays minimal homology with the elements of nature, making it biologically inert. In this area of synthetic biology called DNA programming, nanoengineering methods must account for context specificity (where a circuit evolves), and circuit functionality

(what a circuit does). It is usually only after a rigorous optimisation of these features that bioengineered devices may provide robust behaviour in different cellular contexts.

Synthetic biology is a unique discipline in biological research that has numerous advantages resulting from its systematic engineering principles. Novel synthetic devices are constructed by assembling and inserting specific modules into larger systems, which are in turn connected into a complex biochemical network within cells. This incredible complexity implies that nearly all applications in synthetic biology require the optimisation of biological devices. By doing so, it increases the chances of getting a theoretical model and equivalent experimental data providing the same results. Ultimately, genetic devices rely on a specific sequence of nucleotides. This DNA sequence may encode functional or biologically inert features, but should, in all cases, provide stable information. In a bacterial context, this involves considering a stochastic system where multiple molecules and underlying processes may interact with a circuit functionality. Therefore, in order to achieve specific goals in synthetic biology, it is essential to control and to protect the integrity of the bioengineered devices, and to fully document their variable behaviour.

1.2 Thesis problem statement and objectives

Synthetic biology represents the combination of engineering, computer science and biology. Multidisciplinary studies bring together the expertise of various disciplines, and this implies the use of standards for a better coordination. However, the interface between individual research areas needs to be crystal clear for a smooth orchestration of multiple project strands. For most fields of science, the communication between experts in different domains is not trivial because it requires the abstraction of precise mechanisms into formal entities and concepts. For instance, DNA modelling tools such as the Synthetic Biology Open Language (SBOL) greatly help the *in silico* characterisation of genetic circuits, and their virtual description[202, 194, 37]. These resources are tailored to simplify the communication between molecular biologists and computer scientists, and allow the investigation of further computational studies associated with a synthetic device[23, 128]. Due to the noise and the stochastic environment imposed by living cells, manipulating microorganisms themselves makes biology arguably the messiest of all hard sciences. Therefore, there is a multitude of platforms designed to help the understanding of specific biological processes. Nevertheless, such a number of resources produces large amounts of data, which may become difficult to track in

order to derive meaningful conclusions. Although these tools are connected to biological devices, there is no physical link between them, and only this could guarantee the perfect adequation of computational and experimental data. Hence, there is a gap between the *in silico* data and experimental studies that often delays the overall progress of research projects.

In this thesis, we focussed our interest on unifying wet laboratory and computational resources. The first research statement we tried to answer was the following:

(i) Is it possible to engineer a robust physical link between bacteria and external resources?

In order to answer this question, we designed a barcoding platform (Bac2code) that uses artificial and unique DNA sequences for the identification/documentation of bacterial strains, which is summarised in Figure 1.3. Throughout this work, we first endeavored the development of bioorthogonal DNA barcodes and their associated cloning protocols, and then verified artificial sequences stability *in vivo*. For the development of a barcode biorepository, we used a case study with a library of genetic devices; our second research question was thus stated as:

(ii) How to optimise and to integrate the documentation of bioengineered devices on a unified platform?

For the construction of the biorepository shown in Figure 1.3, we developed a series of biological circuits displaying intercellular communication features. For specific cellular process optimisation, we characterised bioengineered cells in different settings and iterated through the engineering cycle of synthetic biology multiple times via different techniques. In particular, we restricted our field of study to the microscopic scale and investigated microfluidics solutions for bacterial screening, which led us to our last scientific question:

(iii) How to design microfluidics platforms for microscale bacterial studies?

For this matter, we used the synthetic devices created for the biorepository, and presented a few microfluidic design examples that can be used for microscale communication of bacterial species. Finally, we also discuss the potential use of high-throughput platforms for the automation of barcoding processes. In a nutshell, we reviewed and adapted in this thesis the engineering of synthetic biology methods for the creation of a unified bacterial barcoding platform, in an attempt to bridge the gap between *in silico* and *in vitro/in vivo* biology.

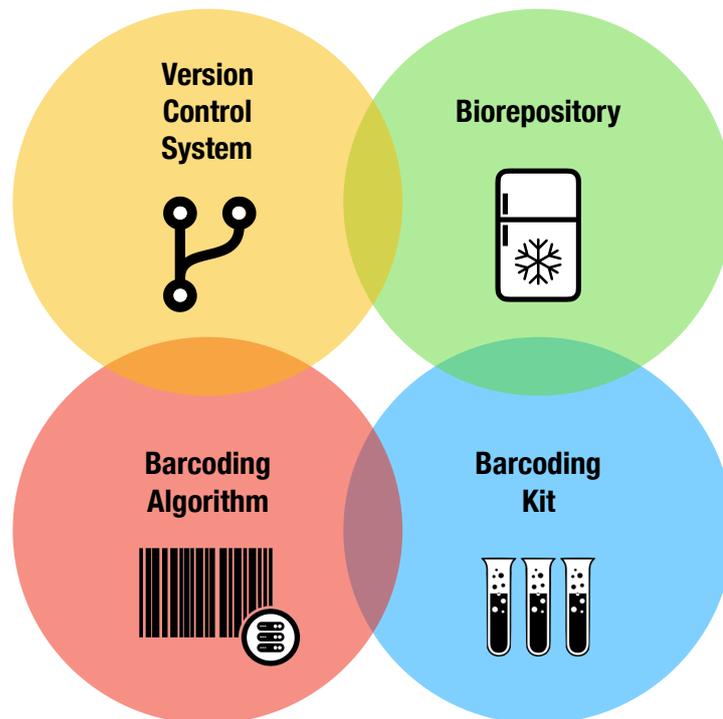


FIGURE 1.3: Main features of the Bac2code platform. Each coloured circle represents an area of optimisation in the development of the barcoding platform. The overlap between circles represents the interface that needed consideration for the integration of individual concepts.

1.3 Plan

We structured this thesis as follows:

- In Chapter 2 we review the fundamental processes of bacterial bioengineering. We used molecular biology methods detailed in Appendix A to produce a series of genetic variants, and provide a list of these constructs later characterised in Chapter 4.
- In Chapter 3 we describe the technical methods that were used to create the barcoding platform. In particular, we explain the concepts of the DNA barcode algorithm that translated biological rules into a computer program, which generates unique bioorthogonal DNA sequences. Secondly, we provide a thorough description of biological protocols that were elaborated to accommodate the propagation and retrieval of these artificial sequences. We present the properties of DNA barcodes and their study in two large scale assays. In order to demonstrate barcode stability, we grew barcoded strains over an extended number of replication cycles in multiple conditions, and provide their characterisation via chemostat and high-throughput systems. Here, we show the efficiency of the barcode integration method - available as a barcoding kit (cf. Appendix C) - and discuss the accuracy of DNA barcode retrieval.
- In Chapter 4 we review the optimisation of genetic circuits for signalling studies in bacteria. This chapter exemplifies the type of data hosted by our barcoding platform, representing our first biorepository parts (cf. Figure 1.3). We guide the characterisation of synthetic circuits via a case-study using a quorum sensing system allowing intercellular communication.
- In Chapter 5 we describe the technical methods that were used to fabricate microfluidics devices. We review microfluidics design and properties for bacterial studies, and provide optimised fabrication protocols (cf. Appendix D) for the setup of micron or submicron size circuits. We describe the results obtained for the fabrication and testing of high-throughput microfluidics devices: first, we provide an overview of the microfluidics chips manufacturing process; then, using the genetic circuits presented in Chapter 4, we discuss the observation of cellular communication at the microscale, taking place between distal microcolonies of bacteria.
- In Chapter 6 we expand on current and further developments of our barcoding platform. In particular, we discuss the online barcoding platform, and potential

applications to adapt barcoding techniques to large scale screening applications via microfluidics.

- In Chapter 7, we close this dissertation with a few concluding remarks and further prospects of this research.

1.4 Contributions to this thesis

The work on DNA barcodes was carried out in collaboration with Dr. Jurek Kozyra. He wrote the algorithm to encode bio-orthogonal sequences and to recognise barcode sequences from sequencing reads, and is also developing the online barcoding platform.

In the microfluidics chapter, the channel-type biochips were designed by Dr. Sunny Park and their silicon wafer was kindly provided by Dr. Lucy Eland. The silicon wafer for diffusion-type biochips was manufactured by INEX, a company specialised in photolithography based in Newcastle-upon-tyne.

The rest of this thesis, including experimental and computational approaches, was research I conducted by myself.

Chapter 2

Background and genetic library engineering

"A grain in the balance will determine which individual shall live and which shall die - which variety or species shall increase in number, and which shall decrease, or finally become extinct" – Charles Darwin (1859)

These words were written by the father of the theory of evolution, and explain the basis of natural selection. Although Darwin postulated and presented his theory over 150 years ago[42], it was not until the 1950s that the molecular structure of living cells was elucidated. Based on preliminary work from Franklin and Wilkins, Watson and Crick published in 1953 the crystal structure of DNA and later earned the Nobel price for Physiology and Medicine in 1962 for their discoveries[236, 39, 181]. After decades of labour to decipher the nature of life, scientists are yet to comprehend all the intricate mechanisms that are involved in the regulation of biological systems. In this thesis, we explored the prokaryotic kingdom[245] and bestowed our efforts to streamline its use for bioengineering studies. In this chapter, we provide the biological background required to understand the engineering of biological circuits in bacteria, and we explain the different notions that were used to build a library of genetic constructs, listed in the last section.

2.1 Engineering the genetic code

Genetic circuits evolve in living organisms and obey general biological rules that are driven by the flow of information from DNA to proteins. Often documented as a three-component system, the central dogma of biology relies on the transcription of DNA to RNA molecules, and the translation of certain types of RNA (messenger RNA or mRNA) to proteins (Figure 2.1)[38]. This highly connected network of DNA, RNA and proteins forms the essence of biological function. In the following subsections, we

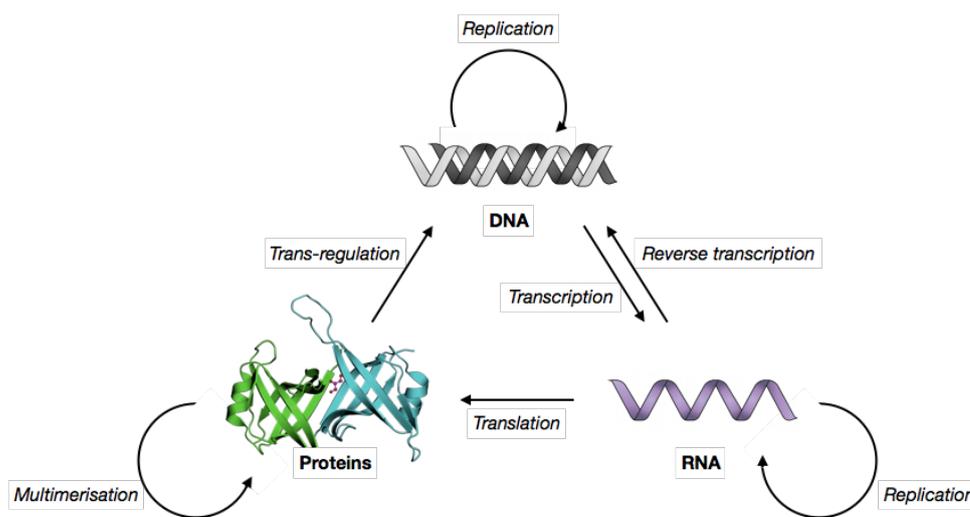


FIGURE 2.1: The central dogma of biology and its components. DNA is transcribed to RNA, that can be in turn reverse transcribed to DNA or translated to proteins. Both DNA and RNA can replicate through specific polymerases, and proteins can assemble into complexes to perform trans-regulation of DNA.

introduce the basic notions that need understanding in order to undertake biological studies.

2.1.1 DNA

The key to understanding life relies on genetic information encoded in DNA[236, 237]. Through this chemical machinery, cellular organisms can replicate and adapt their response to change in the environment. Four nucleotides form the core of DNA: adenine, guanine, thymine and cytosine, commonly abbreviated to A, G, T and C respectively. These nucleotides, or bases, are polymerised on sugar phosphate backbones, generally referred to as DNA strands (Figure 2.2). Strand polymerisation relies on the presence of a 5'-phosphate on individual nucleotides to be incorporated into a growing chain of nucleic acid. By convention, every strand has a 5'-3' orientation depending on the end containing the 5'-phosphate. The 3'-end displays a deoxyribose hydroxyl group, and is thus available to bond a phosphate group, thereby extending the DNA strand. This model depicts our chemical understanding of DNA primary structure.

Individual DNA strands form a duplex structure depending on sequence complementarity between A and T, and G and C, and naturally adopt a helical conformation as shown in Figure 2.3. In nature, the most common form of DNA is found as the B-form

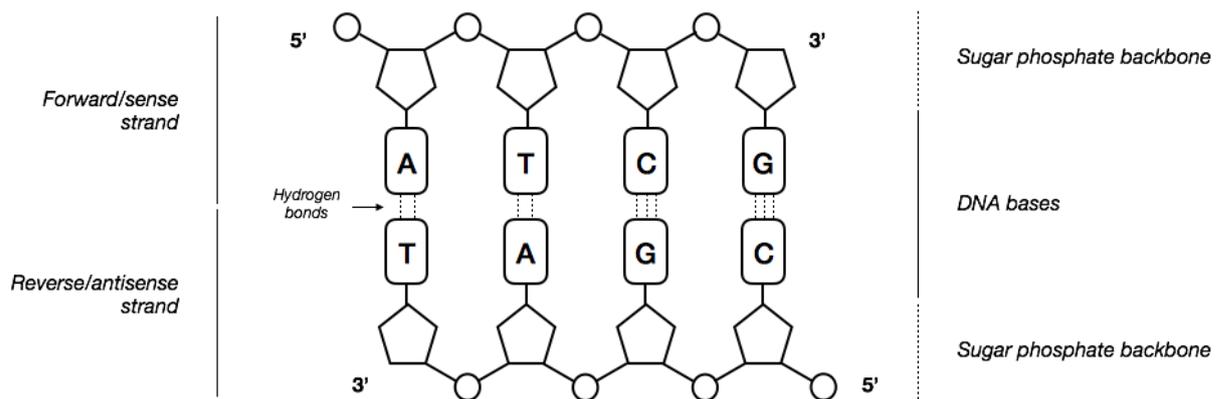


FIGURE 2.2: Two complementary strands of DNA consisting of DNA bases sitting on sugar phosphate backbones. Two or three hydrogen bonds are formed between individual bases depending on the AT or GC base pairing.

double-helix (Figure 2.3). Figure 2.2 outlines that AT and GC base pairing differ in the number of hydrogen bonds formed between individual nucleotides. While AT pairs account for two hydrogen bonds and allow for flexibility in DNA helices, GC pairs form three hydrogen bonds that are harder to break, and add stiffness to helices. Therefore, even if DNA structure is prone to form a double-helix, it can in fact adopt multiple conformations; these are dictated by specific nucleotide sequences. For instance, a single strand of DNA may display hairpin-like secondary structures, where complementary regions form short double helices (or stems), separated by single-stranded DNA loops of non-complementary bases. In the following chapters, we used these features as cofactors to coordinate biological behaviour.

2.1.2 RNA

RNA molecules follow the same principles as DNA, but represent its altered copy where uracyl (U) bases replace DNA-specific thymine nucleotides. The three other bases A, G and C remain unchanged. In cells, RNAs play major roles in the conversion of DNA to proteins, and are a means of regulation for the overall metabolism. Our current understanding of biology allows the division of the RNA world into different categories, dependent on molecule structure and associated functions. RNA is comparable to DNA in many aspects but usually evolves in single-stranded conditions. Since it shares the same basic properties as DNA, stable secondary structures can also emerge to stabilise RNA intermediates. The simplest form of them all is the hairpin model that was aforementioned for DNA. Using a series of these stem loops, a special

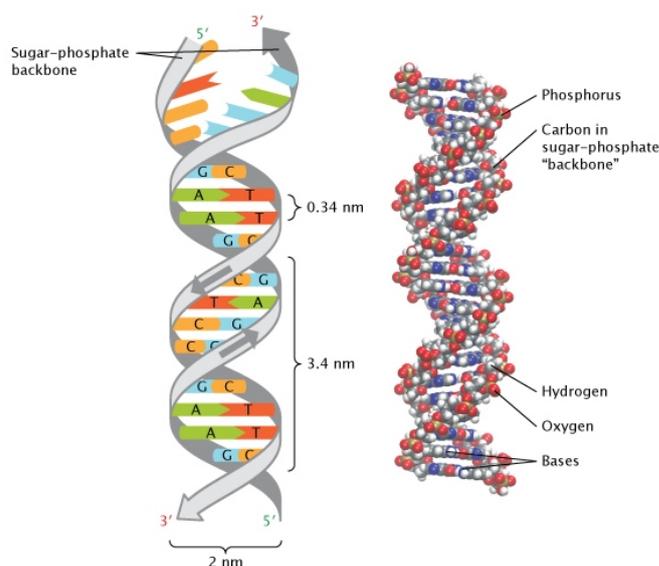


FIGURE 2.3: The three-dimensional double helix structure of DNA, correctly elucidated by James Watson and Francis Crick. Complementary bases are held together as a pair by hydrogen bonds and stacking interactions. Image from Pray (2008)[181]

type of RNA is the transfer-RNA (tRNA), known to incorporate individual amino acids into nascent proteins (or polypeptides) during translation (Figure 2.4). Its structure is complex but very conserved among species as it allows polymerisation of polypeptide chains from mRNA molecules, one of the many essential features for cellular proliferation[35]. Messenger-RNA is the only RNA used for protein translation, and yet, it only represents 1% to 5% of the total RNA in cells. Therefore, only a small percentage of total RNA is recognised by ribosomes and available as a valid template for protein synthesis.

2.1.3 Proteins

Proteins are the third type of polymer present in the central dogma of biology. As aforementioned, only mRNA molecules are used in the translation of RNA to proteins. Ribosomes sliding on mRNA sequences allow the polymerisation of polypeptides encoding protein sequence information. By triplets of nucleotides (or codons), ribosomes read mRNA sequences and recruit individual tRNAs to incorporate specific amino acids to nascent polypeptides[39]. All amino acids are based on a chemical backbone formed of three groups: the N-terminal amine (NH_2), the central carbon and

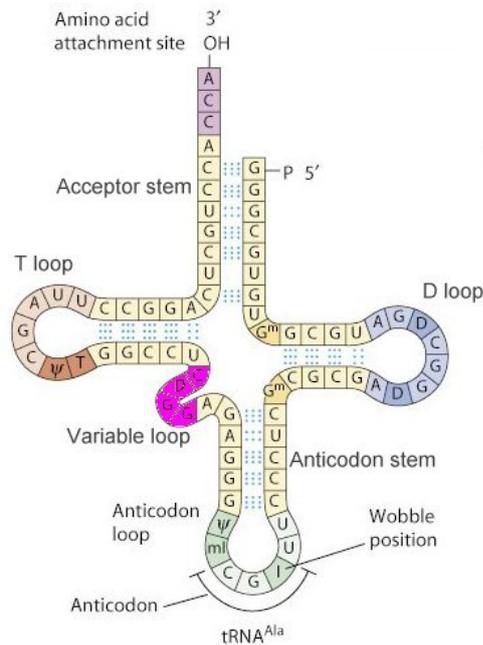


FIGURE 2.4: Alanine tRNA sequence conformation. Dotted lines between complementary bases denote hydrogen bonds between DNA strands. Different hairpin loops are used to stabilise the tRNA on the ribosome machinery. A specific anticodon matching an alanine codon (bottom loop) is used for the specific recognition of adequate tRNAs during translation.

the C-terminal carboxyl ($COOH$) groups. One of 20 available side chains can be integrated on the central carbon (or C_{α}), and each side chain is typically associated with specific physico-chemical properties[26]. During translation, tRNAs are specialised to recruit amino acids and to present them to sliding ribosome complexes for elongation of polypeptide chains. When an amino acid is added to a nascent polypeptide, its N-terminal region is fused to the peptide at its C-terminal end. In this process, a molecule of water is lost, and we thereby refer to individual amino acids in proteins as residues.

In translation, when protein sequences are first produced, they only form a linear sequence of residues, and are catalytically inactive. In order to gain functionality, proteins need to undergo a number of maturation processes that are associated with the formation of secondary, tertiary and quaternary structures. This sequential post-translational process starts by the arrangement of linear peptide chains into α -helices and pleated sheets (or β -sheets). These secondary structures can then be combined to display higher level tertiary structures, also known as folds. Several folds can be assembled into functional protein domains, with multiple domains interacting with one another (quaternary structure), eventually leading to protein multimerisation[213]. In

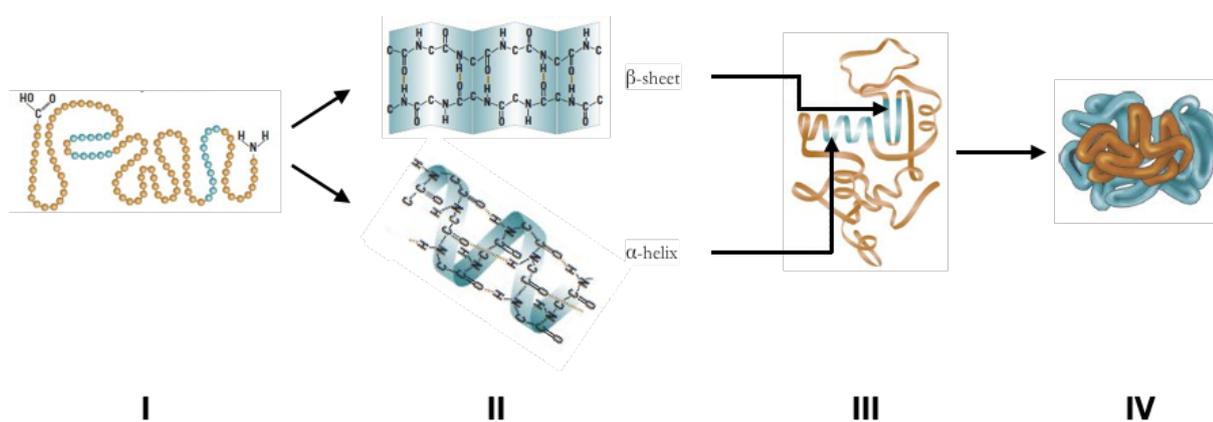


FIGURE 2.5: Protein maturation process. (I) Nascent polypeptides are synthesised and found as a chain of amino acid residues in a primary structure. (II) These are then assembled into α -helices and β -pleated sheets known as the secondary structure. Finally, several secondary structure motifs can be assembled into more complex tertiary (III) and quaternary structure complexes (IV). Adapted from [205].

Figure 2.5, we streamline the maturation process of proteins that leads to the activation of their catalytic activity. Under the right environmental conditions, a protein may display one or several active sites processing different regulatory functions. These functions are usually reached upon binding of ligands on proteins, which triggers specific enzymatic activity. These ligands may be DNA, RNA, proteins or other small metabolites. This regulatory process has further implications for protein function, and may in particular allow the activation or repression of various cellular responses.

The central dogma of biology drives our understanding of biological processes. Transcription and translation are essential for life to replicate, and bacteria are particularly good chassis organisms to study these processes. For instance, cloning circuits in *E. coli* only takes a couple of days, which eases the testing of genetic devices. In most applications in synthetic biology, modification of transcription (promoters/terminators) and translation (ribosome binding site (RBS) and small RNA (sRNA)) regulatory units allows the programming and optimisation of defined cellular functions behaviour [95, 153, 135, 156, 41, 188, 186]. In the next section, we review the key regulatory elements that were accounted for in the construction of a series of genetic devices.

2.2 Circuit replication

In order to modify living organisms, a given genetic circuit needs to account for its replication within the host it evolves in. The world of bacteria is competitive and non-replicative circuits cannot survive in such conditions. Any synthetic circuit must then account for a method to multiply along its host[113]. One method is to integrate genetic devices directly onto the host chromosome, which ensures a stable replication. By definition, cells need to maintain the chromosome (their genetic material) to survive[165]. However, chromosomes usually consist of a single copy per cell, and even if the use of rich media may form concatemers (duplicated fused copies of the chromosome[132]), the overall amount of DNA to be expressed remains very little[34]. To reach higher genetic circuit expression levels, we used vectors that carried biological sequences along a self-replicating high copy backbone.

2.2.1 Replicative plasmids

The backbone normally contains an origin of replication and an antibiotic marker to be selected for in bacterial replication[44]. Plasmid vectors are automatically recognised by their host cellular metabolism. On their own, plasmids are only inert pieces of DNA but some DNA sequences may become active in a biological context where cellular machinery recognises replication elements. Plasmids' origin of replication are usually characterised by an average number of DNA copies per cell. Low copy number plasmids may contain from one to 20 copies of the genetic material per cell whilst high copy number plasmids may reach up to 1000 copies per cell[30, 25, 253, 215]. In practice, one can design genetic devices and ensure they contain replicative elements that allow for *in vivo* propagation. However, this was not the only requirement for a plasmid to survive in bacterial populations: it also needed to be specifically selected for by the use of antibiotics.

2.2.2 Selective pressure

When a plasmid is inserted (or transformed) in bacteria, its replicative elements are detected by the cellular machinery. As all there is is genetic information, transformed DNA is yet another piece of material to be propagated within the host. Therefore, even if a genetic circuit can replicate on its own, it will not be kept in the long term if it does not provide a selective advantage. We usually use antibiotics to solve this issue. There are two types of antibiotics, that either target the cellular membrane or perturb

the central metabolism[127, 59]. Similar to the toxin/antitoxin concept, antibiotics are toxic to bacteria unless they harbor a specific "antitoxin" (the resistance gene product) that inactivates the antibiotic. Two of the most commonly used antibiotics are ampicillin and chloramphenicol[55, 2]. If two genetic circuits were to be combined, these may be used in conjunction with each other in order to select for the replication of each device.

Although there are possibilities to combine several vectors in a single host, a major issue may arise if they use the same replication mechanism[172]. A number of origins of replication have been discovered over the years and all share the same problem: if two circuits using an identical replication unit are combined, the cellular metabolism would not be able to differentiate between the circuits and copy-numbers would constantly fluctuate. This scenario could result in losing one of the circuits. There are a number of compatible *E. coli* origins, however, relatively few are used in laboratories[100, 192]. In this study, we focussed on the use of the plasmid-derived pUC[253] and p15A[207] origins of replication when needed to select for two vectors simultaneously.

Plasmids ease of use and limitations Plasmids are very easy to propagate in bacteria and can be harvested from lysed cells to be used for further genetic manipulation. This is possible due to their independent origins of replication from the chromosome, which allows maintenance of a stable DNA copy number in cells. From an engineering point of view, it is important to be able to change different parameters, such as how much genetic material we want as an input to a system. High copy number plasmids are good to reach substantial levels of expression and, therefore, a better signal of the circuit in the experiment. However, it is common to find some circuits to be toxic in live cells, and that either do not allow for bacterial growth, or that evolved through mutations within the circuit itself and/or the host. In order to avoid this kind of situation, an alternative to lighten cytotoxicity effects is to reduce the plasmid copy number[138]. Another issue may arise from plasmid size[141]. For instance, pUC19 is a plasmid commonly used as backbone, is found in high copy number in *E. coli* and consists of less than 2.5kb of DNA. If the cloning reaction only adds a few kilobases of synthetic circuit to the plasmid, it is unlikely to cost much more to the cells to replicate than the original version (based on the DNA size). However, a 30kb plasmid in high copy would probably not be very easy for bacteria to replicate. First, it would need to be transformed[93, 49, 8] and somehow permeated through the bacterial membrane and then, use the cellular metabolism to replicate its entire sequence. In high copy number settings, this would be > 20Mb of additional DNA to replicate per cell, while the *E. coli* chromosome varies from 4.5 to 5.5Mb; cells would have to replicate

five times as much DNA as they normally do. To avoid this kind of scenario, it is always important to think how we want to propagate a circuit in cells, whether that is the main effector of the system, if it needs to be complemented with other units or if it is cytotoxic. With careful planning, plasmids are a real advantage to test combinatorial libraries of genetic parts. However, in the case of DNA barcodes that were not meant to be expressed in cells, we placed artificial circuits directly on the chromosome via recombination.

2.2.3 Chromosomal recombination

In biology, strands of DNA can be exchanged by the action of specific enzymes called recombinases. These enzymes carry out recombination where a region of homology is detected in DNA, which recruits one or several proteins that cleave and modify pieces of genetic material[234]. A recombination mechanism can be either site-specific or not. This means that certain recombinases are only active when they identify specific recognition sequences (site-specific recombination, *e.g.* lambda-red[180, 102], Cre[252] or CrispR[134] systems). However, other enzymes are solely based on DNA strand homology recognition and exchange, and are often found as a native means of recombination in bacterial species (for instance, the recA system[122]). We discuss in Chapter 3 further aspects of recombination for the insertion and update of barcode sequences.

2.3 Transcription machinery

In DNA, gene coding sequences display functional features on one strand only, and are subject to a specific orientation on the genome. In bacteria, the codirectionality of gene coding sequences allows the transcription of large operons, which generally encode biologically related components.

2.3.1 Promoters

The first key element in transcription is the recruitment of the RNA polymerase (RNAP) at a promoter sequence. In *E. coli*, these sequences are relatively simple and are responsible for the melting of double-stranded DNA (dsDNA) to allow RNA polymerisation on the non-coding strand. The driving force of transcription, the RNA polymerase, is a highly processive protein complex that slides on the DNA coding strand and polymerises an RNA molecule. The sigma (σ) factor is a functional domain of the RNAP

with the objective of finding a relatively stable site on DNA where RNA polymerisation can start[87, 257, 256]. In bacteria, these sites are often characterised with their associated -35/-10 sequences (or boxes), which refers to the relative positioning of sequences with respect to the transcription start site, referred to as +1. With regards to the +1, the further -35 site is the main effector to reach open-complex conformation of the RNAP. Closer -10 sites usually define the rate of dsDNA opening (melting). As a rule of thumb, there are usually 17bp between the -35 and -10 boxes, and 7bp from the end of the -10 to the +1[105]. However, promoter engineering is a tedious task and combining supposedly ideal sequences does not always produce a functioning device. Besides, promoters can be subject to operator sites that attract activators or repressors of DNA transcription by modifying the contact between DNA and RNAP.

2.3.1.1 Constitutive promoters

Promoters have different affinity with the RNAP due to the sequence combinations they may display. In *E. coli*, the -35 'TTGACA' and -10 'TATAAT' sequences form the consensus for the maximum level of open-complex formation and DNA melting respectively[212]. Constitutive promoters are usually very conserved within a same species and only vary from one or two nucleotides per -35/-10 box. These polymorphisms, however, may have a direct impact on a promoters capacity to recruit RNAP molecules, which creates a range of promoter activities (or strengths)[45]. In this study, we used Anderson's library of *E. coli* promoters[10]. Although constitutive promoters hardly differ from their characteristic basal transcription start rate, some other promoters can have more complex structures and involve cofactors to change their relative activity.

2.3.1.2 Arabinose-dependent promoter

The arabinose promoter (PBAD) is an example of a highly responsive inducible promoters where a regulatory protein (AraC) activates its expression state when bound to specific DNA boxes. AraC is constitutively expressed upstream of PBAD in the antisense direction and recognises several sites, one partly overlapping with the -35 box[90, 204]. As shown in figure 2.6, AraC can bind to proximal (I1, I2 half sites) and distal (O1, O2) sequences. Any combination other than binding I1 and I2 abolishes the transcription rate by not leaving enough room for the RNAP to sit on -35/-10 boxes. Only an allosteric change in AraC (provided by arabinose) allows binding to I2 and activation of transcription.

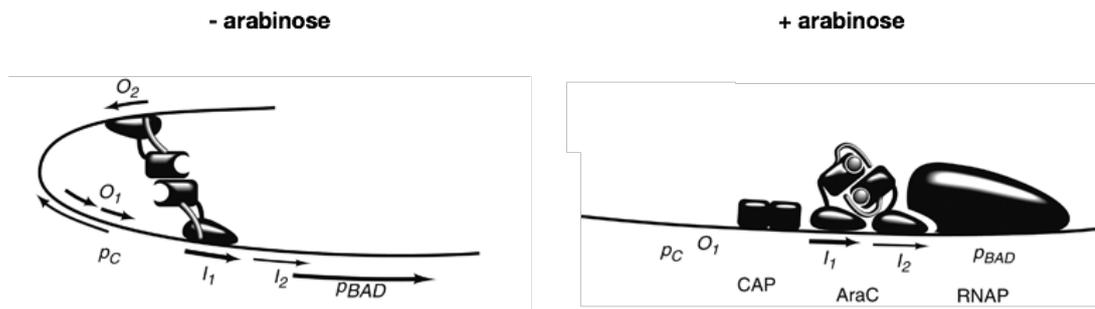


FIGURE 2.6: In the absence of arabinose (left panel), AraC binds to I1 and a distal operator site O2 and forms an inactive dimer. In the presence of arabinose (right panel), arabinose binds to AraC proteins and changes AraC conformational state to I1 and I2 sites, which facilitates RNAP loading onto DNA. Adapted from[90]

2.3.1.3 Quorum-sensing promoter

Bacterial populations growth can be regulated by certain metabolites called quorum sensing signals[51, 240]. Quorum molecules synchronise a number of biological pathways and regulate bacteria at the population level. Acyl-homoserine lactones (acyl-HSL) production is controlled by a bipartite promoter resembling PBAD in a number of ways. As for the role of AraC in PBAD, LuxR acts as a cofactor in order to modulate the promoter activity depending on the presence of acyl-HSL[209]. The P_{L-lux} promoter controls LuxR levels and has a weak activity. The leaky P_{R-lux} promoter encodes LuxI, an autoinducer synthase catalysing formation of acyl-HSL which, in turn, forms an active complex with LuxR and increases its binding activity to P_{R-lux} , and thereby production of inducer. This type of circuit is often referred to as a positive feedback loop, where induction of the system tends to increasingly amplify the inducer signal production. On a population scale, it is only over a certain environmental concentration of acyl-HSL that bacteria become synchronous. For our studies, we focussed on the 3-oxo-C6-HSL: the main quorum molecule in *V. fischeri*.

2.3.2 Terminators

We have previously described how DNA can form secondary structures such as stem loops and vary from a standard double-helix conformation. These structures are sequence dependent and generally harbor long GC stretches that allow stable structures to emerge. Hairpin loops form as physically stable secondary structures and are a system to stop sliding RNAP complexes in their elongation phase[139]. When processing RNAP complexes encounter such secondary structures, it looks like a building block

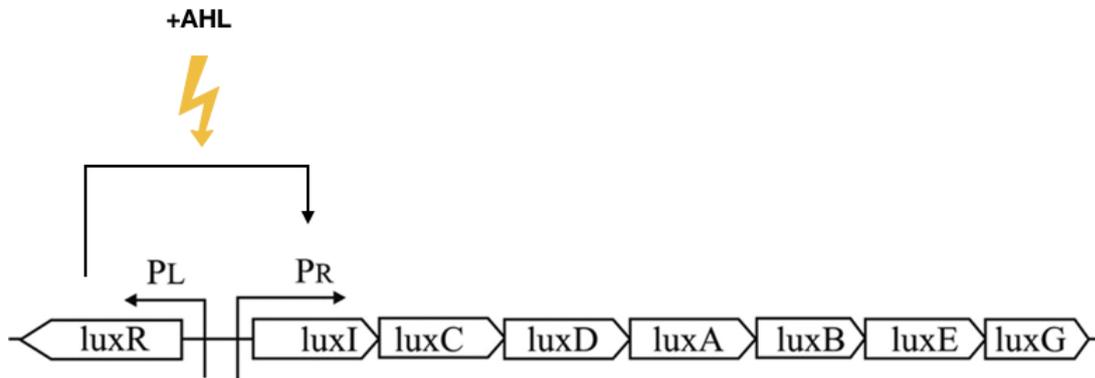


FIGURE 2.7: Structure of the *lux* operon in *V. fischeri*. A bipartite promoter drives the expression of LuxR (via P_L , or $P_L - lux$) and activates the expression of autoinducer and bioluminescence genes via P_R , or $P_R - lux$, in the presence of the quorum signal AHL.

too stable to open and they simply fall off of DNA and terminate mRNA transcription. These so called terminator structures are a check-point ensuring that only certain operons are transcribed and overall genetic regulation is observed. In genetic circuits with consecutive promoters, each controlling some protein expression, terminators are the only way to ensure minimal crosstalk between genetic elements. Without terminators, the RNAP complexes would carry on transcription and be likely to interfere with further regulatory units. In this study, we used a fusion of the *E. coli* *rrnB* T1[174] and bacteriophage T7[246] terminators. Figure 2.8 shows the predicted fold of this terminator fusion, which appears to be very stable *in vivo* and achieves consistently good transcription termination.

2.4 Translation machinery

There are three types of molecules involved in the central dogma of biology: DNA, RNA and proteins. We have introduced so far which components are used to engineer transcription, and how nascent RNA molecules can be produced from DNA.

2.4.1 Ribosomes

The cellular machinery relies on a big family of ribosomal proteins involved in a process called translation. Ribosomes are large RNA/protein complexes that recognise some specific RNA sequences and initiate biosynthesis of proteins from mRNA molecules[117].

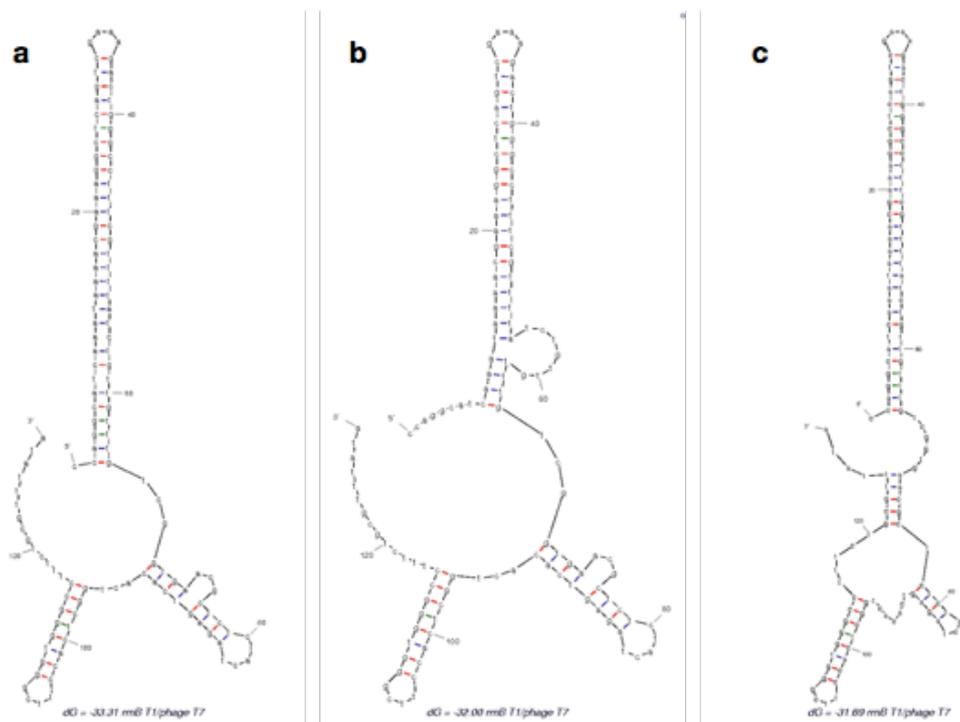


FIGURE 2.8: Predicted DNA fold with three predominant conformations at 37°C (a) $\Delta G = -33.31$, (b) $\Delta G = -32.00$ and (c) $\Delta G = -31.69$ kcal/mol. Plots generated by mFold[263].

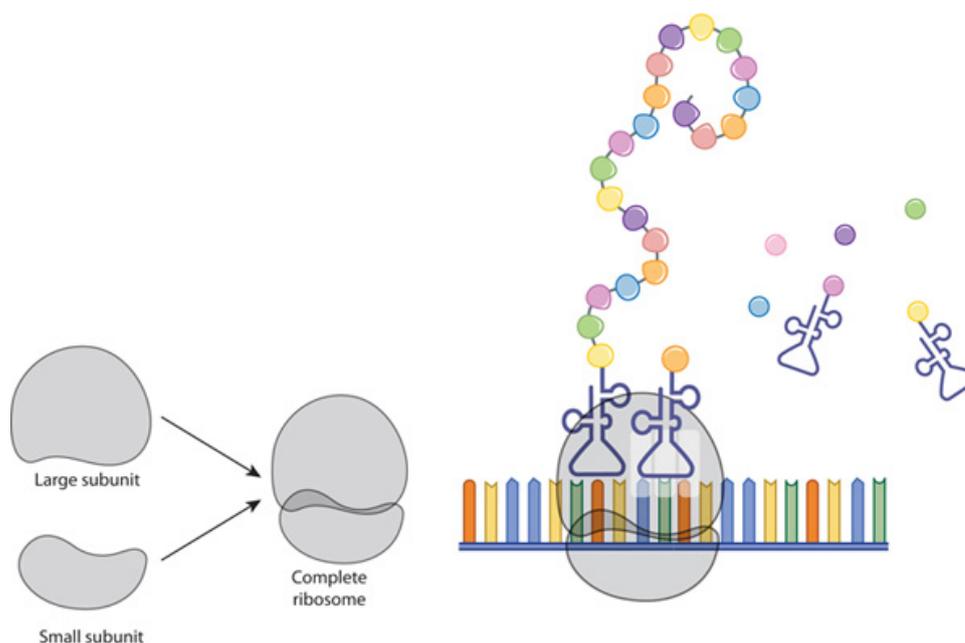


FIGURE 2.9: (A) Simplified small and big subunits and their assembly. (B) Subunits assemble on a mRNA during translation to incorporate amino acids to a native polypeptide chain chain via specialised tRNA. Adapted from [61]

When they find a binding site on a mRNA sequence, two subunits are assembled into a sliding complex that reads ribonucleotides three at a time (codons) from a start consensus 'AUG' sequence. This start codon initiates peptide synthesis and tRNA molecules bring individual amino acids to the ribosome, depending on a specific anticodon to incorporate the right residue in the native polypeptide chain (Figure 2.9)[61]. The role of ribosomes is to recognise specific sequences upstream of protein coding sequences, to translate the coding regions and to achieve translation. Ribosomes sliding on mRNAs stop translation when they encounter a stop codon ('UAG', 'UAA' or 'UGA'). However, there is not necessarily a system as in DNA, where secondary structures would make ribosomes fall off from RNA molecules. This said, bacteria have made very good use of this feature. In bacterial genomes, most genes are encoded in operons where a single promoter drives the expression of several downstream genes. In operons, intergenic regions that separate the different proteins are usually small and coding sequences sometimes overlap. Having little to no spacer on mRNA intergenic regions ensures that ribosomes currently translating one gene of an operon are likely to recognise the next start sequence and translate the entire operon. Nevertheless, ribosomes are sensitive to the presence of small RNAs (sRNA) that cover some of the mRNA regions, and may create roadblocks for the assembly and/or sliding of ribosomes.

2.4.2 Ribosome binding sites

In synthetic biology, one can modulate protein expression on the transcriptional and translational levels. While promoter engineering is not trivial (transcription regulation), changing ribosome recognition sequences (RBSs) is sometimes an easier way to regulate protein levels[200]. As RNAP complexes recognise DNA sequences to start transcription, ribosomes assemble sequentially on mRNA molecules at a RBS. Ribosomes are ribonucleoproteins, which means they are made of a mixture of RNA (ribosomal RNA, or rRNA) and proteins. In *E. coli*, the small RNA subunit 16S recognises a consensus sequence 'AGGAGG/A' located about 8bp upstream of a protein start codon. The formation of a complex between the ribosome small subunit and RNA in turn recruits the bigger ribosomal subunit to initiate the search for a start codon[117]. We generally represent a RBS by a core 6bp combination, but combinatorial libraries of RBSs have showed that RBS neighbouring sequences are also important in the regulation of translation. In this study, we mainly tested four RBS sequences of different strengths (or affinity to start translation) and adapted the circuits with weaker RBSs in order to accommodate the translation of multiple proteins simultaneously.

2.4.3 Transcription factors

In bacterial genomes, a lot of operons express proteins that may in turn regulate their own promoter activity. These are transcription factors that affect the rate at which the RNAP may recognise and process intracellular mRNA molecules. Transcription inhibition is often achieved by hybridisation of regulatory proteins to DNA near or within promoters and by loop formation due to transcription factor multimerisation[90, 209, 164]. In general, transcription factors dimerise in order to activate or repress the activity of certain promoters. In the example given in Figure 2.6, AraC strongly binds to different operator sites located around the promoter and only facilitates the access to the RNAP when bound to L-arabinose[204]. In the context of synthetic biology where genetic circuits should be remotely controlled, the activity of such transcription factors together with promoter engineering allows different regulatory units to turn on and off, given certain environmental cofactors. Circuits built in this study involved the use of AraC, TetR, LuxR, and cI transcription factors.

2.4.4 Protein degradation tags

In bacterial cells, there are no intracellular compartments that split DNA from RNA and proteins, as opposed to eukaryotes. Hence, all cellular products are mixed and

held together by a membrane in what is called the cytosol[228]. It is easy to see how the space within the cytosol can become crowded: if a lot of proteins are produced from an exogenous genetic circuit, then translation spends a lot of resources on expressing target proteins but the cell has less energy to replicate its own metabolism. Therefore, in genetic circuits, it is necessary to minimise the long term impact protein overexpression may have over the cellular background[56]. In order to reduce the metabolic load of protein overproduction, some tags can be attached to proteins in order to attract intracellular protein degradation complexes[81, 9, 86]. The *E. coli* system is widely used in microbiology and based on the ClpXP and ClpAP proteases. In this study, we tagged proteins with the peptide 'AANDENYALVA' C-terminal regions for rapid degradation of protein products. Proteins tagged with this peptide achieve a half-life of $T_{\frac{1}{2}} = 40 - 45min$, which allows for degradation of proteins that could be stable for days otherwise, and is essential for reliable and precise time-lapse measurements.

2.5 Fluorescence integration

In synthetic biology, a lot of methods make use of reporter systems to visualise biological responses to changes. Within cells, fluorescence emission is a way to report the activity of some regulatory units potentially localised at certain loci[21]. Modern technologies provide multiple types of fluorescence detection methods that allow for taking measurements from single cell to population level over a discrete or continuous time scale. From an engineering point of view, the coupling of fluorescence emission with metabolic processes helps to track progression of the behaviour of genetic circuits[200]. It makes invisible processes quantifiable, which is essential for the modelling of biological devices.

2.5.1 Fluorescent reporter proteins

To study bacteria, we often coexpress fluorescent reporters with regulatory units that aim to express specific signals. If fluorescence can be observed, then it means that target signals should be expressed too. In a functional genetic design, fluorescence measurements are a good indicator of the state of biological processes. For instance, automated equipment such as plate readers now usually integrate several filters to process absorbance, fluorescence and other types of light emission in an incubator over long periods of time. In this system, 96 or 384 samples are usually loaded onto a microplate (Figure 2.10) and growth characteristics can be assessed simultaneously[222].



FIGURE 2.10: Automated plate readers for *in vitro/in vivo* characterisation. On the left, a BMG plate reader that integrates absorbance and fluorescence measurements to follow bacterial cultures over time. On the right, a Biotek plate reader that is specialised in the making of precise growth curves.

This method is particularly efficient for screening of libraries and allows a large number of replicas in characterisation studies. However, it does not capture single-cell resolution details that can be achievable with microscopes[80]. In this study, we used different fluorescent reporters to follow genetic circuit activation upon external signal reception, and characterised constructs from the single cell to the population scale (data presented in Chapter 4).

The choice of fluorescent protein needs to account for the fact that several signals associated with different cellular processes may need to be detected at the same time. If so, reporters should be distinguishable from one another to observe interpretable metabolic signals. This is possible because the wavelength at which a fluorescent protein is excited is often a narrow window within ultra violet light (UV). However, as shown in Figure 2.11, excitation and emission spectra of different fluorophores may sometimes overlap. For instance, green and yellow fluorescent proteins have overlapping spectra, and would be a poor combination in order to differentiate fluorescence from disparate cellular processes. Nevertheless, coupling one of these with red fluorescence would ensure that detected signals would come from a single fluorescent protein. In this study, we mainly used green and red fluorescent proteins and optimised genetic constructs for a better signal to noise ratio based on fluorescence output.

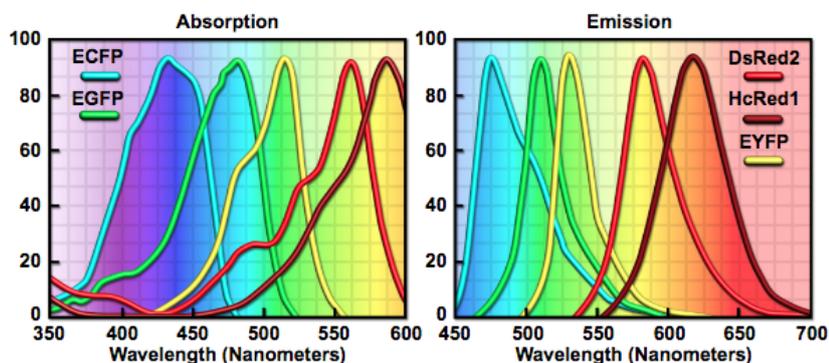


FIGURE 2.11: The excitation spectrum (left panel) represent what light should be shined at a protein to observe a quantum response. This response is characterised by the emission of a slightly different type of light (right panel). A few common proteins are represented by areas of the plot with specific peaks of fluorescence excitation/emission[53].

2.5.2 Photostability

In synthetic biology, genetic circuits can be represented as interconnected regulatory networks and fluorescence is a way to report differential biological states, linked to the activity of a circuit. While target proteins can be fused with fluorescent proteins at their N- or C-terminal fragment, protein fusions sometimes affect the activity of the target protein. In the context of this thesis, we kept fluorescent proteins as coexpressed reporters in order to keep other enzymes in their native form. Fluorescent proteins are photostable molecules that can be excited when exposed to UV light and emit fluorescence as a response[118]. The first isolated fluorescent protein originated from the jellyfish *A. victoria*, and it encodes the green fluorescence protein (GFP). Now, a number of fluorescent proteins are available for use and proteins mainly differ in terms of maturation time, photostability and brightness[210]. For practical reasons, brighter proteins are better for long-term fluorescence detection since the signal they emit does not need a long UV exposure time to produce a detectable response[108, 19]. At the same time, this limits physical damage that UV light can impose on living cells. Fluorescent proteins are an easy reporter system to use to study living cells and the choice of specific proteins should be made carefully.

2.6 Genetic circuits built in this study

2.6.1 Microbiology techniques

For the construction of genetic circuits and nanodevices, we used standard microbiology techniques as described in Appendix A. DNA isolation was usually performed using miniprep kits supplied by Qiagen and enzymatic reactions based on the New England Biolabs (NEB) catalogue. When possible, basic microbiological protocols were derived from the Green and Sambrook manual[83].

2.6.2 Plasmids

For *E. coli* and *B. subtilis* studies, we used or built the plasmids listed in Table 2.1. Plasmids were always propagated in *E. coli DH5 α* at 37°C unless stated otherwise in the plasmid notes (for instance, temperature sensitive replicons). The R6k origin of replication plasmids were propagated in an *E. coli DH5 α λ pir* strain to allow them to replicate[133]. Plasmids were usually constructed in *DH5 α* and characterised in *E. coli* MG1655 or BW25113 strains[12].

TABLE 2.1: Plasmids used in this thesis

Source	Label	Template	Modification	Marker	Note
[195]		pACYC184		Cam, Tet	Cloning template
[253]		pUC19		Amp	Cloning template
[238]		pBR322		Amp	Cloning template
[187]		pSB1A2		Amp	Cloning template
[187]		pSB1C3		Cloning template	
[187]		pSB1AK3		Amp, Kan	mCherry LuxR with LVA degradation tail
[90]		pBad33		Cam	Cloning template for low-copy num- ber circuits with inducible Pbad sys- tem
[43]		pKD46		Amp	thermosensitive plasmid with 30deg growth and propagation
Demuris		pIJ790		Cam	thermosensitive plasmid with 30deg growth and propagation
[32]		pCP20		Amp	Frt-site recombinase, temperature sensitive
This study	pCWqs01	pACYC184	B0015	Cam	Cloning template for low-copy num- ber circuits
This study	pCWqs02	pSB1AK3-B0015	B0015 x3	Amp	Cloning template for high-copy num- ber circuits

This study	pCWqs03	pACYC184-B0015		Cam	sender-CI without TetR constitutive expression
This study	pCWqs04	pSB1AK3-B0015x2	Prlux-B0034-mCherry	Amp	mCherry under control of Prlux
This study	pCWqs05	pSB1AK3-B0015x2	J23119-B0034-LuxR	Amp	LuxR under the control of strong constitutive promoter
This study	pCWqs06	pSB1AK3-B0015x2	J23119-B0034-LuxR-Prlux-B0034-mCherry	Amp, Kan	mCherry LuxR with LVA degradation tail
This study	pCWqs07	pSB1C3	PltetO1-B0034-LuxI	Cam	LuxI with LVA degradation tail
This study	pCWqs08	pSB1C3	B0034-GFP _{lva} -B0015	Cam	GFP with lva degradation tail
This study	pCWqs09	pACYCB15	J23119-B34-TetR	Cam	TetR with LVA degradation tail
This study	pCWqs10	pSB1C3	Plteto1-B34-LuxI-B34-GFP-B15	Cam	inducible signal of C6-HSL couple with GFP, LuxI and GFP both with LVA degradation tail
This study	pCWqs11	pACYCB15	Plteto1-B34-LuxI-B34-GFP-B15-J23119-TetR	Cam	LuxI GFP and TetR with LVA degradation tail
This study	pCWqs12	PACYCB15 and pSB1AK3		Cm, Amp, Kan	cotransformation of s-3oc6hsl and r-3oc6hsl
This study	pCWqs13	pSB1A2	B34-luxI.B34.GFP.B15	Amp	LuxI and GFP with LVA degradation tags

This study	pCWqs14	pACYCB15	Ptet.B34- luxI.B34.GFP.B15	Cam	LuxI GFP with LVA
This study	pCWqs15	pACYCB15	Ptet.B34- luxI.B34.GFP.B15- J23119.TetR	Cam	LuxI GFP and TetR with LVA
This study	pCWqs16	pSB1C3	Ptet.B34- luxI.B34.GFP.B15	Cam	LuxI and GFP with LVA
This study	pCWqs17	pBad33	B34- luxI.B34.GFP.B15	Cam	LuxI and GFP with LVA
This study	pCWqs18	J23100	B34-GFP _{lva} -B15- pSB1C3	Cam	GFP with LVA degradation tag;
This study	pCWqs19	J23104	B34-GFP _{lva} -B15- pSB1C3	Cam	GFP with LVA degradation tag
This study	pCWqs20	J23117	B34-GFP _{lva} -B15- pSB1C3	Cam	GFP with LVA degradation tag;
This study	pCWqs21	J23107	B34-GFP _{lva} -B15- pSB1C3	Cam	GFP with LVA degradation tag;
This study	pCWqs22	pBad33	B30-sfGFP	Cam	sfGFP; col2
This study	pCWqs23	pBad-sfGFP	b30-luxI	Cam	LuxI with LVA tag
This study	pCWqs24	B30-mCherry- pUC19	luxR.b30	Amp	coding parts of 3oc6hsl receiver
This study	pCWqs25	B30-mCherry- pUC19	J23104	Amp	mCherry reporter
This study	pCWqs26	LuxR.b30.b30- mCherry-pUC19	pLuxRI	Amp	assembled receiver for 3oc6hsl

This study	pCWqs27	luxR-pluxRI-mCherry	b30-sfGFP _{lva}	Amp	sfGFP with LVA degradation tag
This study	pCWqs28	pBad-luxI-sfGFP	b30-mCherry _{lva}	Cam	luxI and mCherry with LVA tags; colony 2
This study	pCWqs29	pBad-luxI-sfGFP	b30-sfGFP _{lva}	Cam	luxI and sfGFP with LVA degradation tags; isolated col3
This study	pCWqs30	pUC18r6k-Tn7T-Kan	b30-luxI	Amp, Kan	insertion of luxI sender device under Pbad control into ApaI sites
This study	pCWqs31	luxR-PluxRI-sfGFP	pUC18r6k-Tn7T-Kan	Amp, Kan	insertion of luxR receiver device under Plux-R/L control into ApaI sites
This study	pCWqs32	AraC-pBad-luxI-mCherry	pUC19	Amp	luxI and mCherry with LVA degradation tag; col5
This study	pCWqs33	AraC-pBad-luxI-sfGFP	pUC19	Amp	luxI and sfGFP with LVA degradation tag; col2
This study	pCWqs34	AraC-Pbad-luxI-sfGFP	pUC18r6k-Tn7T-Kan	Amp, Kan	insertion of luxI sender device under Pbad control into ApaI sites; col23
This study	pCWqs35	luxR-PluxRI-sfGFP	pACYC184	Cam	sfGFP with LVA degradation tag
This study	pCWqs36	luxR-pluxRI-sfGFP	pBR322	Amp	sfGFP with LVA degradation tag
This study	pCWqs37	pUC19-luxR-pLpR-sfGFP	luxI	Amp	luxI and sfGFP with LVA degradation tag
This study	pCWqs38	pUC19	pRlux-sfGFP	Amp	sfGFP with LVA degradation tag

This study	pCWqs39	luxR-J23104- pRlux-sfGFP	pUC19	Amp	luxR and sfGFP with LVA degradation tag; synthetic receiver with luxR strong constitutive expression; col 5
This study	pCWqs40	luxI	pUC19-luxR- J23104-pRlux- sfGFP	Amp	luxR, luxI and sfGFP with LVA degradation tag; synthetic receiver with strong luxR expression and luxI amplification; col3
This study	pCWqs41	luxR-pLpR-sfGFP	pUC19	Amp	luxR and sfGFP with LVA degradation tag; original synthetic receiver with LVA-tagged cofactor; col 3
This study	pCWqs42	luxI	pUC19-luxR- pLpR-sfGFP	Amp	luxR, luxI and sfGFP with LVA degradation tag; original synthetic receiver with LVA-tagged cofactor and luxI amplifier; col 4

2.6.3 Strains

For *in vivo* studies, we used *Escherichia coli* and *Bacillus subtilis* species. As part of the Portabolomics project (portabolomics.ico2s.org), our basic strains of *Escherichia coli* MG1655 and *Bacillus subtilis* 168CA were fully sequenced by MiSeq shotgun genome sequencing. Table 2.2 lists the strains used and created in this thesis. As detailed for plasmids, strains were grown at 37°C unless stated otherwise. This table does not include all *E. coli* plasmid propagation and expression strains. For the study of genetic devices in microfluidics, we expressed plasmids in *E. coli* BW25113 (Chapter 4) and fabricated 33 models of biochips (Chapter 5).

TABLE 2.2: Strains used in this thesis

Source	Label	Species	Name/description	Phenotype	Marker	Note
Portabolomics		<i>E. coli</i> MG1655	MG1655			Wild-type <i>E. coli</i> ; strain sequenced for Portabolomics project
Portabolomics		<i>E. coli</i> BW25113	BW25113			Wild-type <i>E. coli</i> ; strain sequenced for Portabolomics project
Portabolomics		<i>B. subtilis</i> 168 CA	168CA	<i>trpC2</i>		Wild-type <i>B. subtilis</i> ; strain sequenced for Portabolomics project
[260]		<i>B. subtilis</i> ZPM6	ZPM6	<i>trpC2 amyE::(zeo-Pxyl-mazF)</i>	Zeo20	derived from <i>B. subtilis</i> 1A751; contains miniMazF cassette with Pxyl promoter
This study	sCPW01	<i>E. coli</i> MG1655	pKD46		Amp	thermosensitive plasmid with 30deg growth and propagation; col4

This study	sCPW02	<i>E. coli</i> MG1655	pIJ790		Cam	thermosensitive plasmid with 30deg growth and propagation; col8
This study	sCPW03	<i>E. coli</i> MG1655	Barcode selection <i>E. coli</i>	::(<i>barcode-cat</i>)	Cam	better propagation with Cam17 (portabolomics strain 4);
This study	sCPW04	<i>B. subtilis</i> 168 CA	Barcode selection <i>B. subtilis</i>	<i>trpC2</i> ::(<i>barcode-zeo-Pxyl-mazF</i>)	Zeo	portabolomics strain 5
This study	sCPW05	<i>B. subtilis</i> 168 CA	Barcode loopout <i>B. subtilis</i>	<i>trpC2</i> ::(<i>barcode</i>)		sCPW04 barcoded strain without selection cassette (portabolomics strain 6);
This study	sCPW06	<i>E. coli</i> MG1655	Barcode loopout <i>E. coli</i>	::(<i>barcode</i>)		sCPW03 barcoded strain without selection cassette (portabolomics strain 7)
This study	sCPW07	<i>E. coli</i> BW25113	chromosome sender GFP	<i>tn7-AraC-pBad-luxI-sfGFP-tn7</i>	Kan	insertion in BW25113 chromosome; col 8
This study	sCPW08	<i>E. coli</i> BW25113	chromosome sender mCherry	<i>tn7-AraC-pBad-luxI-mCherry-tn7</i>	Kan	insertion in BW25113 chromosome; col 3

This study	sCPW09	<i>E. coli</i> BW25113	chromosome receiver GFP	<i>tn7-luxR-pluxRI-sfGFP-tn7</i>	Kan	insertion in BW25113 chromosome; col 11
------------	--------	------------------------	-------------------------	----------------------------------	-----	---

Here, we presented the basis of biological processes and a simplified version of the factors that affect the cellular metabolism. There are many ways to optimise genetic circuits, but the best way is to build them rationally, and to keep them simple. Biological functionality depends on a number of factors, and while different circuits may produce the same output, the simplest one will often be easier to propagate in cells. In the context of bacterial signalling by fluorescence, we characterised a set of genetic devices and describe their optimisation and use in Chapters 4 and 5. In the next chapter, we provide a walkthrough the development of DNA barcodes.

Chapter 3

DNA barcoding in bacteria

The work presented in this chapter was performed in collaboration with Dr. Jurek Kozyra, who developed the *in-silico* encoding method of DNA barcodes.

3.1 Filling the gap between *in silico* and *in vitro/in vivo* biology

There is a vast amount of techniques that are used to characterise biological devices. These include computational studies, or *in silico* methods, that produce biological models that aim to describe experimental data. However, there is sometimes a narrow link between models and precise *in vitro/in vivo* data. Consequently, this makes the matching of theoretical and experimental measurements more challenging, and may result in results irreproducibility. In particular for the field of biomedical research, data reproducibility is a major issue for scientific progress[250, 13]. Currently, the authentication of biological strains is based on a set of sparse documentation (protocols, simulations, sequences...) and a relatively well-labelled test tube[69]. This usually results in a gap between wet lab and computational biology.

In this chapter, we detail the creation of DNA barcodes as bacterial identifiers, their associated cloning method for barcode insertion/update, and their stability *in-vivo*. We focussed on the prokaryotic kingdom and studied chromosomally-encoded barcodes in the context of two distinct types of bacteria: *E. coli* and *B. subtilis*. In order to assess the integrity of DNA barcodes, we used two types of assays - hereinafter referred to as chemostat and high-throughput studies - followed by the targeted sequencing of barcode sequences.

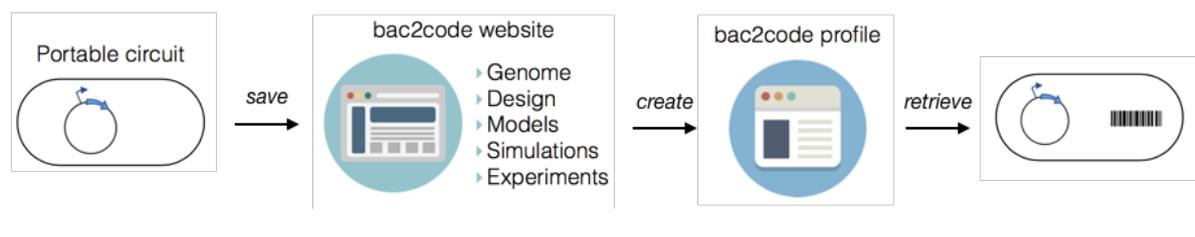


FIGURE 3.1: Barcoding cells: a specific genetic circuit can be tagged and saved along computational/experimental resources on the bac2code server. These tags, known as DNA barcodes, allow retrieval of full documentation for any barcoded strain via a single sequencing reaction.

3.1.1 Sequencing technologies

Over the last 40 years, the development of DNA sequencing methods - from Sanger to Next Generation Sequencing (NGS) approaches - has greatly helped the understanding of biological sciences[201, 104]. For instance, bacterial species identification can be streamlined by rRNA sequencing, which can provide information about the class of microorganism a sample may belong to. These rRNA sequences are highly conserved among species but display different polymorphisms, often characteristic of diverging species[115, 189]. However, precise strain identification remains difficult, and apart from wide-scale whole genome sequencing, there is not a single, short and inert DNA fragment that could be used as a standard identifier for individual species. As an advance towards the grouping of *in silico* resources to relevant strains, we created a barcoding platform that, via a single Sanger sequencing read, can provide thorough documentation about the studied microorganism.

3.1.2 A biological version control system

The Bac2code system aims at standardising the storage of information about biological circuits. A barcoded strain on its own is associated with a set of key characterisation features, and making the link between strains allows to recreate the structure of a version control repository. In the field of computer science, version control software eases the organisation of files and allows a better coordination between multiple project strands. This is because to produce reliable results, pieces of code need to be embedded in specific files and organised in a defined manner to be compiled. It works the same way in the development of biological devices for synthetic biology. Design modularity helps to create several genetic constructs, and to combine these in an organised fashion for higher level assembly. On the software side, version control systems

streamline the error-tracking mechanism and allow for "roll-back" solutions, where unfunctional code can be reverted to a previous working state. However, in biology, the tracking of a parental strains information in complex genetic backgrounds can become a tedious process. By the use of a standard DNA barcode identifier, we recreated the architecture of a version control software where barcoded strains are automatically linked to one another; these are used to reconstruct a strains lineage and facilitate the retrieval of parental sequences. Figure 3.1 displays an overview of our unified barcoding platform (bac2code). A dedicated server stores resources about bacterial strains (*in silico* and *in vitro/in vivo* protocols) and links them to a specific DNA tag (barcode) anchored in the bacterial chromosome. These tags are synthetically engineered DNA identifiers and provide the missing link between biological samples and all types of external documentation.

3.1.3 Online documentation platform: Bac2code

We are currently developing the web-platform called Bac2code that hosts all barcoded strain documentation. The construction of the barcode database and online platform are led by Dr Jerzy Kozyra. When users retrieve new strain information on the Bac2code server, a targeted sequencing read, or a barcode sequence only, allow access to a specific barcoded strain profile. On the online server (<http://bac2code.com>), a set of key strain features are highlighted to provide an easy access to computational and wet lab data. A basic strain is characterised by its genetic information (the DNA sequences of the chromosome and eventual genetic circuits) and different types of assays that provide supporting information. We aim to design a modular system that could be used as a laboratory notebook and also integrate external tools such as SBOL and other *in silico* resources, as outlined in Figure 3.2. Practically, a biologist in possession of a barcoded strain should be able to access all of its *in silico*, *in vitro* and *in vivo* information and browse the strains lineage in a single Bac2code profile. Further aspects of the online Bac2code platform are discussed in Chapter 6.

3.1.4 Barcode minimal identifiers

In this study, we focussed on the biological interface of synthetic DNA barcodes engineering. As shown in Figure 3.3, the design of DNA barcodes had to comply with many considerations from both computational and biological points of view. In the following sections, we describe methods that were used to design DNA barcodes and their integration in the biological context of two model bacterial species: *E. coli* and *B.*

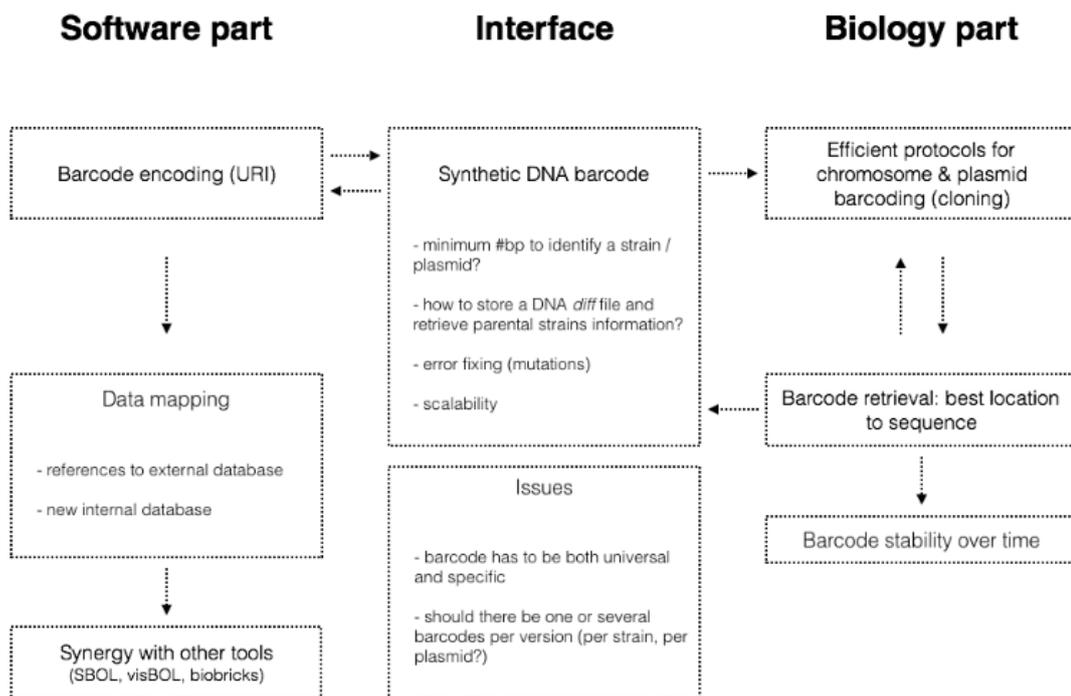


FIGURE 3.3: Design of a barcoding platform. The design of DNA barcodes was split in three parts: the software stage, the biological stage and the interface between the two. The software stage (left) assures a stable encoding of DNA barcodes and the easy storage of links to computational data. The biological stage (right) is represented by the setup of an unobstrusive cloning method. The interface (middle panel) is a set of parameters required to define the structure of DNA barcodes.

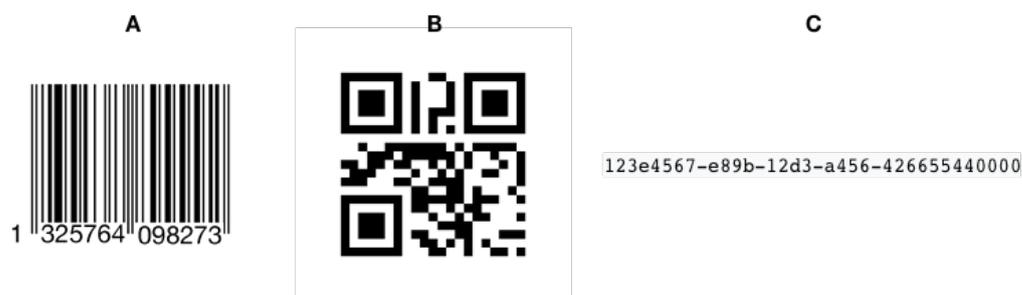


FIGURE 3.4: Different types of barcodes. (A) EAN barcode, (B) QR-code and (C) universally unique identifier (UUID).

3.2.1 DNA encoding

3.2.1.1 Barcode sequences

In order to build DNA barcodes, it was necessary to choose an appropriate DNA encoding method. Barcodes are found everywhere in everyday life: from train tickets to products packaging. For instance, european article number (EAN) and universal product code (UPC) barcodes are used worldwide - in particular in Europe - and aim to reference all sorts of consumer goods (EAN and UPC barcodes). All barcodes are, *per se*, a visual and/or alphanumerical sequence that references some item (Figure 3.4). The encoding method behind our DNA barcodes relies on universally unique identifiers (UUIDs)[1]. Initially used by Microsoft to create encoded object identifiers, we used the UUID version 4 algorithm to generate random 32-hexadecimal character sequences. Because these are randomly generated, there are about $5.3 * 10^{36}$ different possibilities of UUIDs of this type[244]. Version 1 and 2 are connected to the date/time and MAC addresses of the machine generating UUIDs, which would have been unnecessary for our purpose. Version 3 and 5 are related to namespace identifier and name, and are encoded by hashing algorithms MD5 and SHA1. These would only make sense if all strains had a unique name, which cannot be guaranteed at the scale of multiple laboratories. Therefore, we chose the random UUID version 4 to help us set a potentially large-scale system for the identification and encoding of bacterial species profile.

Based on the redundancy of the genetic code, we encoded each UUID hexadecimal character as a chain of three nucleotides, which resulted in constructing 96bp DNA barcodes (Figure 3.5). We were inspired by the way cells naturally perform translation from 64 codons to encode 20 amino-acids, but we set an equal distribution of DNA triplets per hexadecimal character to obtain four codons per character (Table 3.1). To

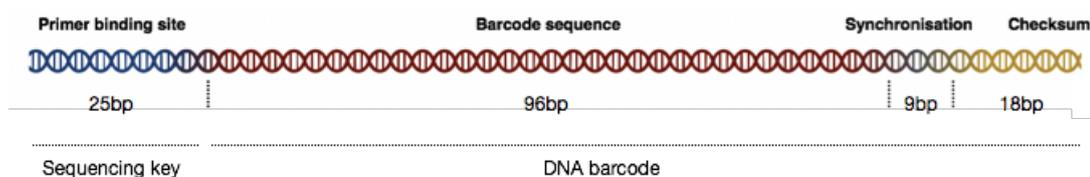


FIGURE 3.5: DNA barcodes genetic structure. DNA barcodes can be read through targeted sequencing at a universal primer binding site. Their structure is divided in subsequences that correspond to a specific UUID (96bp), a recognition synchronisation sequence (9bp) and a checksum for UUID read inaccuracies (18bp).

apprehend sequencing result variability, we also integrated an error-correction mechanism to account for truncated or partially incorrect reads. This was done via a 9bp recognition sequence that our decoding algorithm searches for, in order to align sequencing reads of variable lengths. Once this sequence is found, its upstream region maps a UUID reference encoded as 96 nucleotides, while downstream 18 bases correspond to the checksum for errors handling (details in Additional barcode features paragraph). Barcodes generated with our platform are kept in a secure database and mapped to sequencing reads in order to retrieve strain profile content. Should there be any misalignment between a sequencing read and a UUID, the checksum is used as a secondary encoding method to provide better chance of retrieving strain documentation in poor sequencing conditions.

3.2.1.2 Bio-orthogonality

We introduced in chapter 1 how biological elements may interfere with DNA, and how abiological sequences may be designed to minimise the impact of the introduction of synthetic DNA in cells. In the design process of DNA barcodes, we reduced the possibility of biological interaction between synthetic DNA and cellular molecular machinery as much as possible. While our encoding method created pseudorandom DNA barcodes, most sequences would still display biological features to some extent. As a standard tool for cloning, restriction enzymes are molecular subunits that generally recognise sequences of 4 to 8 nucleotides and cleave these sequences. In the prospect of DNA barcodes, these were exactly the type of features that needed removal. During barcode generation, we removed over 180 recognition sequences from barcodes, which covered a vast majority of known restriction sites (Figure 3.6B). However, we omitted to remove one non-specific enzyme site, as it would appear in most sequences. If it

TABLE 3.1: Mapping between hexadecimal characters and nucleotide triplets for the encoding of DNA barcodes

Character	Triplet 1	Triplet 2	Triplet 3	Triplet 4
0	AGC	GTA	CCG	CGT
1	ACC	TAT	GAC	GGC
2	ATT	GGT	TAG	TCT
3	AGA	GAT	GTG	TAC
4	GTT	TAA	TCA	CTG
5	GGG	ATC	CTC	TCC
6	AAC	CGG	GCC	GCG
7	AAG	AGT	CCT	GCT
8	GAG	GTC	TGC	TTC
9	CCC	AGG	CAA	CGA
a	ATG	GAA	GGA	TGA
b	CCA	GCA	AAT	TCG
c	ACT	CTA	TTA	TTG
d	CAG	CAT	CTT	TGT
e	ATA	CGC	TGG	ACA
f	AAA	ACG	TTT	CAC

were to be removed, it would have detrimental effects on the DNA encoding algorithm. These would cause a greedy convergence towards solutions where a single nucleotide would be repeated over the entire barcode sequence. Since this enzyme restriction site was promiscuous along *E. coli* and *B. subtilis* chromosomes, we thus considered that it would not be an issue if it were present on barcode sequences too. Figure 3.6 shows a conflicting case of DNA sequence containing a restriction enzyme site, and how it could be removed by using redundancy of the DNA encoding method. For each UUID hexadecimal character, there were 4 possibilities of nucleotide triplets (Figure 3.6A, note that the final encoding map shown in Table 3.1 was updated). Therefore, our algorithm allowed use of an alternative triplet in conflicting cases, where a sequence had to be changed to alleviate any potential biological interaction. In a similar fashion, and for sequencing needs this time, we limited the appearance of single- or di-nucleotide repetitions in barcode sequences. These tend to decrease the quality of sequencing reads, especially in Sanger sequencing[98]. Besides practical consideration for sequence retrieval, removing repetitive features also minimised the chances of mutations appearing due to DNA polymerases sliding over long DNA stretches of the same nucleotide. Finally, a transcriptional/translational units screen was performed by removing any sequences that would locate a protein start codon in the proximity of potential ribosome binding sites. After all the biological information removal

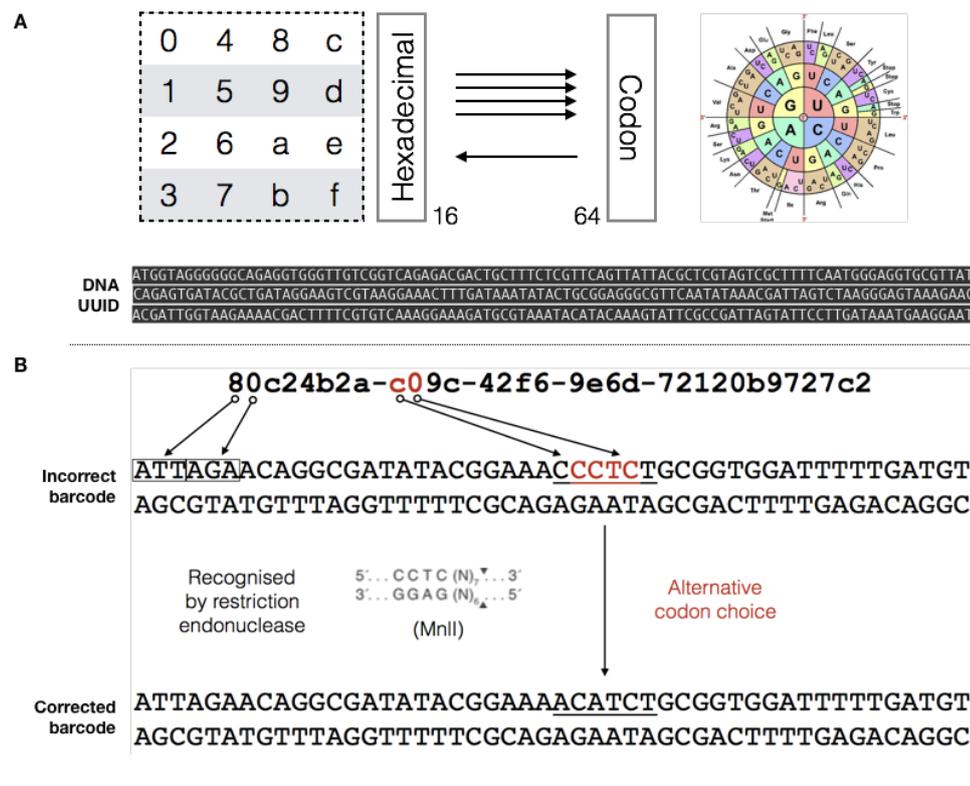


FIGURE 3.6: DNA barcode redundancy. (A) Each UUID hexadecimal character corresponds to four possible DNA nucleotide triplets (or codons), while each codon can only match one hexadecimal character. When UUID hexadecimal characters are converted to nucleotide triplets, undesirable sequences (for example, the *MnII* enzyme restriction site in (B)) are directly removed by choosing one of three alternative triplet possibilities that would comply with all the barcode design concepts.

layer of our algorithm, DNA barcode sequences often resulted in an average of < 50% GC-content. As it is known that AT base pairing contributes to destabilising DNA helices[251], a lower than average GC content favours DNA instability. Unstable DNA sequences would not naturally be expected to encode biological information, and following biological orthogonality concepts, the development of abiological DNA barcodes allowed us to build inert sequences that should not interact with the metabolism of bacteria.

3.2.1.3 Additional barcode features

Universal primer All barcoded strains present a DNA barcode preceded by a primer binding site. This universal primer allows the targeted sequencing of any barcoded strain after polymerase chain reaction amplification (PCR) of the neighbouring species-specific genes. Our DNA encoding algorithm was mainly used for the production of

DNA barcodes from individual UUIDs and to minimise potential interactions with cellular components. However, we submitted the primer binding sequence to the same process too, in order to produce a primer that would be unlikely to be linked to the molecular machinery. Best primer sequences were analysed by GC-content (50% – 55%), dissimilarity to bacterial genomes and GC-clamp. A consensus of adequate GC content featuring a GC-clamp and the highest dissimilarity to genomic content was chosen as the universal primer. Finally, as a complementary verification method of barcodes integrity, we designed a second species-specific primer. This sequence is located downstream of barcodes and targeted by a reverse primer to sequence DNA antisense strand. Together, this set of primers allows for full recovery of barcode sequences in two sequencing reactions.

Synchronisation sequence Synchronisation sequences, located directly downstream of barcode UUIDs, are invariant sequences that are used to align sequencing reads and to retrieve barcode and checksum information. These sequences were subjected to the same rules as the primer design (without a GC-clamp constraint), but consisted of much smaller synthetic DNA fragments (9bp).

Checksum There are three checksum elements appended to barcode sequences. These checksums allow to verify that barcode sequences match a record in our database, and to retrieve their associated information. An individual checksum is a number encoded by the MD5 hashing algorithm and represented by the combination of six nucleotides. Three checksums are concatenated one to another to validate the identity of DNA barcodes. Two of them respectively represent the sum over the elements of first and second part of DNA barcodes, while the last checksum element is computed for the entire barcode sequence. Any conflicting result between a given sequence and the expected checksum would detect inaccurate or incomplete sequences. In which case, checksums are compared to barcode entries and blasted against the barcode database to find an associated profile.

3.2.2 Cloning of DNA barcodes

3.2.2.1 Considerations related to the design and use of DNA barcodes

In synthetic biology, the two most commonly used microorganisms for bacterial studies are *E. coli* and *B. subtilis*. To use barcodes as universal identifiers, we demonstrated barcode stability in these two model organisms. Although *E. coli* (Gram-negative) and

B. subtilis (Gram-positive) are two common strains of bacteria in laboratories, performing the same experiment in each microorganism often requires substantially different protocols. One of the main differences between these species is explained by the structure of their cell walls[54]; they have different thicknesses, and therefore processes that allow transport of DNA through the cell wall are different too. In order to create a generic cloning method, we gathered information about common cellular mechanisms for DNA uptake in Gram-negative/-positive species and formulated a consensus protocol to clone barcodes in these species.

Due to their different cellular structures, DNA uptake mechanisms differ between bacterial species, but still converge in the use of analogous genetic concepts. Although the manipulation of *E. coli* and *B. subtilis* may be distinct, both species rely on similar kinds of events for the remodelling of their genetic structure. In order to ensure stable replication of DNA barcodes, we cloned synthetic products on the bacterial chromosome. Although barcoding methods may be adapted to barcoding plasmids and other kind of DNA vectors, barcoding a strain directly onto its chromosome would ensure the natural propagation of artificial sequences. For these means, we devised a generic method to first, assemble barcodes *in vitro*, then to clone them in different species through an analogous system of homologous recombination. This approach provided several advantages: (i) there is no need for selective pressure to propagate the chromosome, (ii) it avoids barcode loss if it was placed on an unstable plasmid and (iii) this places DNA barcodes away from eventual genetic circuits propagated in bacterial strains, reducing the risk of unexpected proximity interactions. In *E. coli*, and to a further extent in *B. subtilis*, different cloning sites have been characterised and are known as points of insertion of genetic circuits. Therefore, it made sense to find new cloning sites located away from any such points of insertion or important genome regions. Altogether, we developed a strategy suitable for cloning in *E. coli* and *B. subtilis*, but this generic method could in theory be applied to a wider a range of Gram-negative/-positive organisms and yeast species. Last but not least, cloning via homologous recombination allows seamless insertion of barcode sequences and their update through the exact same procedure.

3.2.2.2 Overview of the cloning method

Cloning of DNA barcodes is a two step process: recombinant barcode *in vitro* assembly and *in vivo* bacterial chromosome propagation. Figure 3.7 streamlines the process of barcoding bacteria. We chose homologous recombination as the consensus mechanism for the best flexibility and ease of use to clone DNA barcodes[173]. Other approaches,

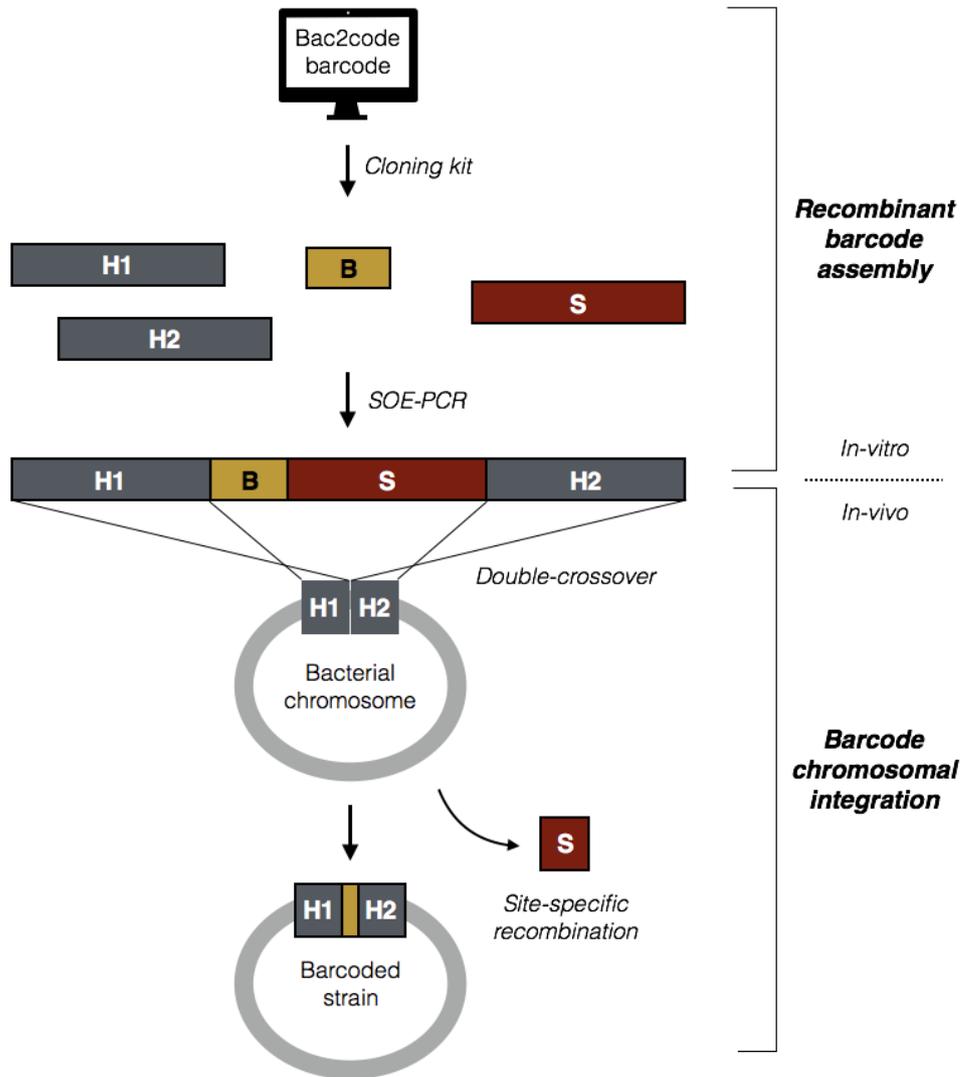


FIGURE 3.7: Method to barcoding bacteria. The barcode server (bac2code) supplies DNA barcodes provided with the necessary samples to assemble a recombinant barcode fragment by SOE-PCR (top panel). Recombinant DNA is integrated by homologous recombination at specific *E. coli* and *B. subtilis* loci and is curated from selective markers by site-specific recombination (bottom panel).

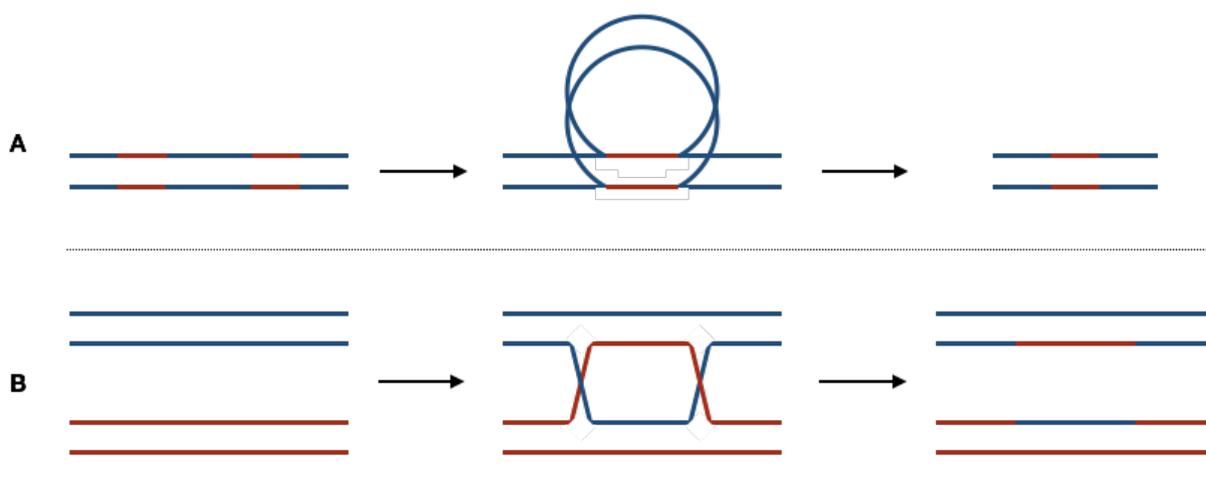


FIGURE 3.8: Single and double crossover events. (A) Single crossover: within a singular double strand of DNA, repeated or homologous regions (red) can be found separating different regions of DNA (blue). These homologous regions can hybridise and loop out DNA fragments, thereby losing an intermediary piece of genetic information. (B) Double crossover: when two double-stands of DNA share some homology (represented by the crossing points), pieces of DNA strands can be exchanged.

including CrispR amongst others, were also considered but using the most basic mechanism for DNA strand exchange appeared to be more scalable on a large scale[22, 146, 11, 58, 239, 252, 259, 102, 134, 160]. To promote this type of recombination, the easiest method was to place flanking regions of the chromosome upstream and downstream of specific circuits to propagate[191, 178]. When DNA is added to the cells, the chromosome replication machinery recognises homologous DNA arms and crosses them with parental DNA (Figure 3.8B). In this process, regions of the genome that separate homology arms are lost and replaced by recombinant DNA. The chromosome then harbors a synthetic DNA circuit, but there is a need for selective pressure to keep this modification as a dominant feature in an entire bacterial population. Therefore, within recombinant DNA fragments, we placed a selective marker downstream of barcode sequences and aimed at: (i) integrating DNA barcodes in the presence of a selection marker and (ii) removing the selection marker after isolation of barcoded strains. The latter alleviated the risk of incompatibility between potential selection markers, and reduced the amount of biological features to a minimal bioorthogonal barcode sequence.

Functional recombinant DNA fragments were obtained by a first *in vitro* step: this was achieved by splicing overlap with extension PCR (SOE-PCR, cf. Figure 3.9). In PCR, DNA mixed in different proportions with reagents is amplified and provides a good starting material for cloning experiments[241, 163]. Cloning via homologous regions required amplification of a homologous region with the host genetic material,

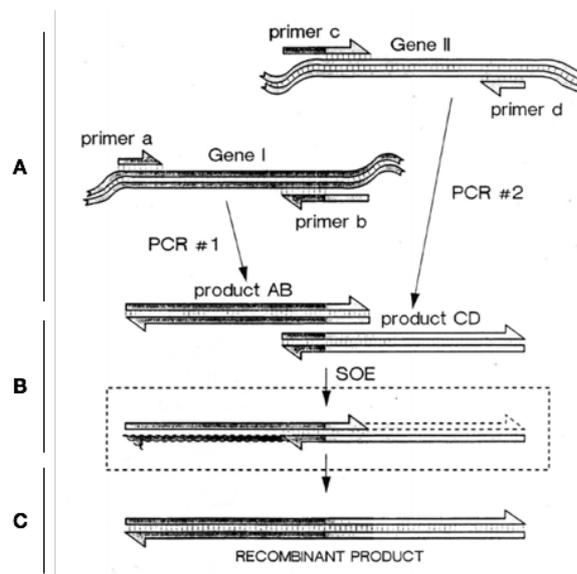


FIGURE 3.9: (A) Two DNA fragments are amplified individually. The genetic primer design allows two fragments to overlap at one end. (B) Assembly by splicing overlap extension of individual fragments. Fragments anneal at the overlap region whilst the polymerase fills in missing 3'-ends of fused products. (C) Purification of the fused product via a standard PCR with primers corresponding to 5'- and 3'-ends of the final product.

Adapted from [106]

and to introduce a disruption in native DNA sequences. In SOE-PCR[106], barcodes were assembled and selected by the fusion of homology arms to synthetic barcodes linked to a selection marker. After purification of recombinant DNA barcode fragments, bacterial cells were transformed and selected for the integration of barcode sequences via the expression of a selection cassette. By default, we used chloramphenicol in *E. coli* and zeocin in *B. subtilis*. However, alternative selection cassettes may be chosen instead, if these were to be incompatible with strain specific features. A counter-selection step follows the isolation of positive transformants (barcoded bacteria) and allows removal of selection markers by site specific recombination. This second step permits "looping-out" of selection cassettes. This mechanism is natural in *B. subtilis*; bacteria recognise genome fragments of homology and can undergo chromosomal reconfiguration via single-crossover events (Figure 3.8A)[258, 161]. Although *E. coli* is usually not capable of such reconfiguration, genetic tools can replace this process by expressing specific recombinase proteins. For both species, this resulted in the removal of selection markers and left a final piece of recombinant DNA as marker-free synthetic DNA barcodes.

TABLE 3.2: Barcodes cloning sites in *E. coli* and *B. subtilis*

Species	Upstream gene	Downstream gene	Genome location
<i>Escherichia coli</i> MG1655	pspG	qorA	331°
<i>Bacillus subtilis</i> 168CA	dusA/ybxH	ybxI	19°

3.2.2.3 Finding new cloning sites

We adopted a well-defined method to integrate biological concepts, and to find new suitable cloning sites in *E. coli* and *B. subtilis* species. First, we gathered data about known chromosomal insertion sites in each species[183]. We then postulated that DNA barcodes would have fewer chances to interact with cellular machinery if they were to be located in "dead-end" regions. Here, we mean an integration at converging ends of operons. As detailed in chapter 1, bacterial genomes are organised in operons that drive transcription of several genes on a single mRNA. Translation then takes place on these mRNA molecules, almost immediately. Therefore, placing barcode sequences at the 3'-end of oppositely orientated operons - together with filtering out biological sequences - would ensure that barcodes would not be actively transcribed in cells. This reasoning tackled the issue of a relative location for a cloning site, but did not account for specific biological context.

On the chromosome, interacting regulatory units are often found in neighbouring locations. Not all genes in the genome are essential, and studies have looked at reducing genomes to minimal working living organisms[77, 111, 121, 224, 180]. Therefore, for each species, we gathered and curated data about essential genes from the literature and ruled out locations in the genome that were neighbouring any essential units[125, 116, 43]. In a similar logic, we avoided locations near elements that were known to belong to important families of metabolism/membrane/transcription and translation proteins. Finally, in order to get a higher average of barcode copy number per cell, we searched for sequences that comply with the aforementioned considerations from the origin of replication (*oriC*) and picked sites within a 1Mbp window around *oriC*. This way, barcodes could be placed in safely replicable zones in a favourable biological context, where they would be unlikely to interact with proximal elements or to interfere with strain specific genetic circuits. Table 3.2 provides more precise information about integration sites location in *E. coli* and *B. subtilis*, while Figure 3.10 provides a visual of the different genomic features.

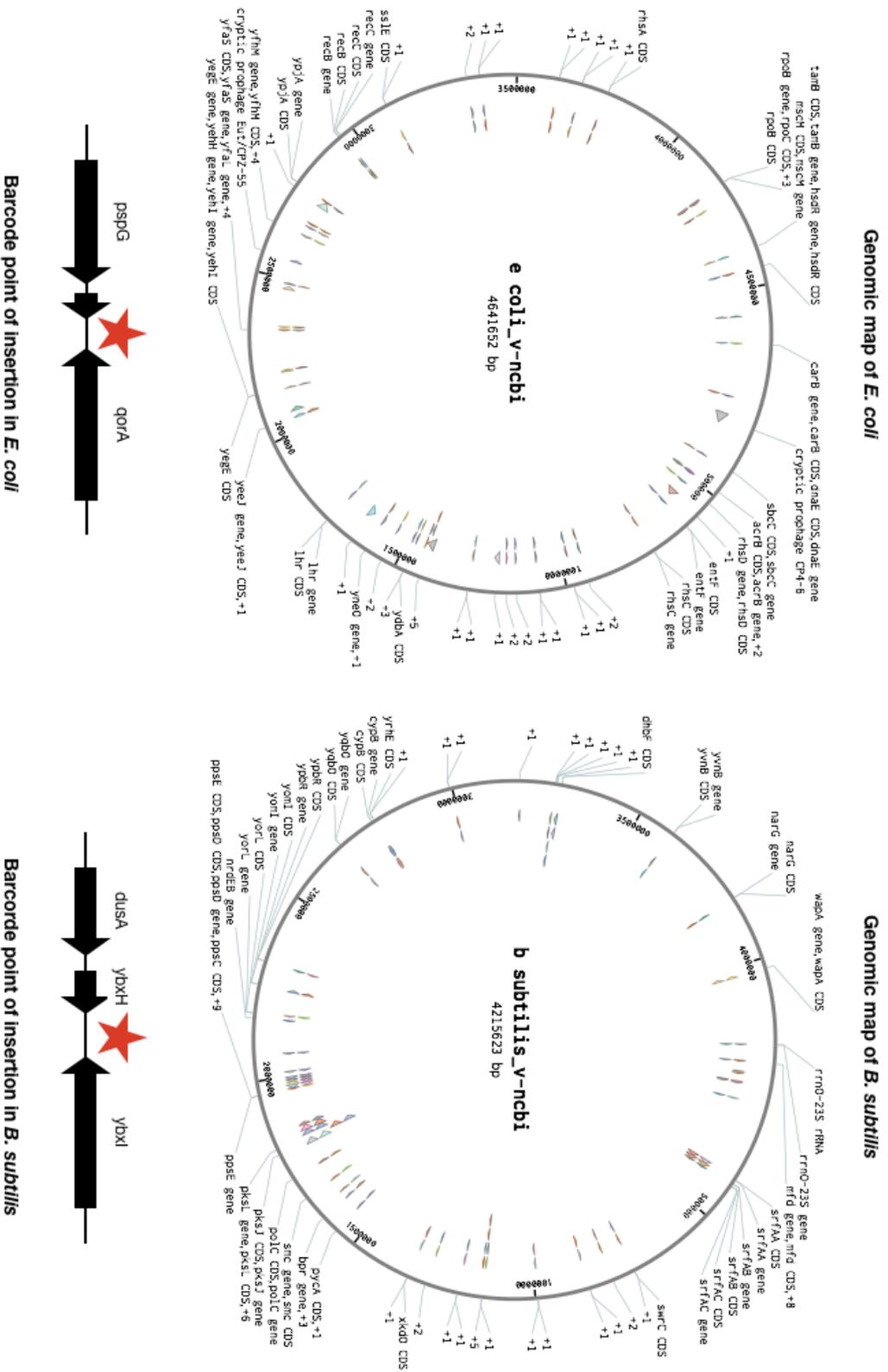


FIGURE 3.10: Integration sites for barcodes cloning. Both panels show the genome arrangement around barcode cloning sites indicated by a red star. (A) describes the *E. coli* neighbouring region while (B) shows the equivalent in *B. subtilis*. Arrow units represent nearby genes and show their left-to-right or right-to-left orientation.

3.2.2.4 *In vitro* assembly of recombinant DNA

Barcodes need to be logically assembled into cloning products in order to lead to viable replication in bacteria. All steps that were involved in PCR amplification of DNA fragments were performed using primers listed in Table 3.3. To minimise the risk of nucleotide misincorporation in PCR, one should use high-fidelity DNA polymerases at all times. Two homology arms were chosen and amplified from each recipient strain of *E. coli* MG1655 and *B. subtilis* 168CA. For homologous recombination, barcodes were flanked by DNA fragments - about 500bp in length - that corresponded to regions upstream and downstream of the species-specific insertion sites (Figure 3.10). A low melting temperature overlap between homology fragments and DNA barcode/selection cassette sequences allow a first PCR reaction (splicing-overlap) to anneal specific overlaps in an ordered fashion. This PCR was set with equimolar amounts of left-/right-homology arms, individual DNA barcodes and selection marker sequences (75fmol each), but without primers. At the melting temperature, individual fragments would anneal to one another. In the following extension cycle, the DNA polymerase would fill gaps and regenerate complementary strands. Individual fragments were thus sequentially assembled into longer ones, up to the correct recombinant DNA arrangement. A second PCR was then used in order to "clean" the assembly, and specifically amplify recombinant products (Figure 3.9). This was done with a pair of primers annealing at a higher temperature corresponding to the targeted amplification from the ends of the recombinant DNA. Because this method allowed intermediary products to form, for instance homology fragments with selection cassette but without barcode, a gel purification was recommended to isolate the exact DNA fragment that represented recombinant products. Recombinant DNA was obtained with this purification step and 500ng at least should be extracted for successful bacterial transformation.

TABLE 3.3: Individual PCR fragments generated to clone barcodes. Fragments that were amplified represented either left-/right-homology (respectively Left-H and Right-H), selection cassette (Selection) or Barcode sequences. Gram represents the species primers are designed for: *E. coli* (Gram-negative, -) and *B. subtilis* (Gram-positive, +).

Fragment	Gram	Template	Forward primer	Reverse primer	T_A	Size
Left-H	-	gDNA	GTAGTCAAATTCACCACGC	CGCAATAGTGACAGATTTGA- TTATC	62	443
Right-H	-	gDNA	G TTCCTATTCTCTAGAAAGT- ATAGGAACTTCGTGCTGTAC- CCTACATACAGC	GATCACCGGCGGTAAGAAAG	66	474
Selection	-	pACYC184	TTACGCCCCGCCCTG	G TTCCTATACTTTCTAGAGA- ATAGGAACTTCGAGACGTTG- ATCGGCACGTAAGAG	67	801
Barcode	-	Gblocks	TCAAATCTGTCACTATTGCG- TGGACATACATAGTATACTC- TGGTG	GGCGGGGCGTAATAAG	63	287
Left-H	+	gDNA	AGCAGAAGGAAAGGTGTTT	GGCTCTAAATCTCATGCTCA- AAAC	63	485
Right-H	+	gDNA	TATACCCAGGGAGACCCGGC- GACTAACCATTTAGGATG- TAATCAGGCCATAC	GTCAAGCGTGAATTCGAATC	62	527
Selection	+	gDNA	TTATCAGTCCTGCTCCTCGG	CGGGTCTCCCTGGGTATACG- AGTAGGTATTACTACCCAAT- CAGTACGTTAATTTTGGC	67	941

Barcode	+	Gblocks	GAGCATGAGATTTAGAGCCT- GGACATACATAGTATACTCT- GGTG	CCGAGGAGCAGGACTGATAAG	64	296
---------	---	---------	--	-----------------------	----	-----

3.2.2.5 *In vivo* integration protocols

The linear recombinant DNA obtained from the *in vitro* assembly was used to promote barcode double recombination on the bacterial chromosome. We developed two protocols that allowed the integration of recombinant barcode DNA in *E. coli* and *B. subtilis* strains, and provide these experimental procedures in Appendix B. These methods were used as a consensus and are applicable for any barcode that needs to be propagated.

3.2.3 Retrieval of DNA barcodes

As synthetic identifiers, DNA barcodes needed to offer the possibility of being addressable and retrieved. In the context of this study, this was achieved by Sanger sequencing. A number of sequencing techniques exist and work in different ways, but Sanger sequencing is probably the most used technique for sequencing small cloning products in laboratories[98]. To sequence a DNA fragment, Sanger sequencing adds fluorescently labelled dideoxynucleotides from a targeted locus defined via a single primer. This reaction is achieved with the DNA polymerase and means that nucleotides are incorporated in the 5'-3' direction. The directionality of the sequencing primer is therefore important in the targeting of sequencing products. In our approach, a uniquely addressable sequence is located upstream of all DNA barcodes and targeting this region allowed us to retrieve any DNA barcode on demand.

In this section, we have explained the methods that were set up in order to integrate DNA barcodes in bacteria. We designed DNA barcodes to fulfil a set of requirements including bio-orthogonality, a minimal size, *in vivo* stability and versatility of cloning methods. DNA assembly and transformation protocols were tested by four individual scientists who demonstrated positive integration of barcodes in different *E. coli* and *B. subtilis* species. The following sections provide a better understanding of barcode retrieval and their *in vivo* stability studies.

3.3 Large-scale screening setup

Barcode identifiers should be stable and their sequence should be able to be retrieved and intact after many cycles of DNA replication (*e.g.* cellular growth). Automated growth equipment allows serial growth of bacterial cultures at a high-throughput rate, and to take measurements of the biological fitness at regular intervals. A basic method

to observe whether cells behave normally is to grow them in favourable conditions and to observe their resultant growth curve[165, 68]. Bacteria are small and divide rapidly: about every 20min for classic laboratory *E. coli* and *B. subtilis* strains. Therefore, any abnormal replication due to the insertion of a genetic device should be recognisable by affecting specific growth rates. We opted for high-throughput techniques to screen for a large number of genetic isolates and followed them over unusually prolonged periods of growth to identify any barcode sequence instability.

The idea of integrating DNA barcodes in bacteria is only viable if barcode sequences can be retrieved, and still provide the encoded information they are meant to carry. In order to demonstrate barcodes stability, we employed two different methods: (i) imitating a "lab-like" situation, where strains would be grown and subcultured many times, and (ii) a more accurate approach to evaluate an exact number of replication events (*e.g.* barcodes replication), and subsequently drawing a link between barcodes replicability and eventual mutations appearing over time.

3.3.1 Continuous flow systems

As we detailed above, bacteria grow quickly and the need for fresh nutrient is a limiting factor in most biological applications. If the right molecules do not surround cells or if these molecules are in insufficient quantities, then bacteria slow down their growth in what is called a stationary phase (cf. Figure 3.11)[233]. Although it is always possible to redilute stationary cultures into fresh medium, bacterial growth is physiologically affected by periods of starvation (maximal cell density in the medium, no more growth)[158, 264]. One solution to this problem is to cultivate cells in continuous growth conditions via bioreactors. Bioreactors can be assembled in order to provide a constant supply of fresh nutrients to growing cultures and, therefore, allow a better tracking of growth experiments.

Bioreactors require a deeper understanding of cell growth as represented in Figure 3.11[262, 107]. In Figure 3.11, phase 3 shows the steepest increment in cell density over time. Derived by plotting optical density measures at $\lambda = 600nm$ over time, the exponential growth phase can be used to approximate cellular division rate[158, 147]. Knowledge of the frequency of cellular division (coupled with replication of one copy of the chromosome) is essential for the work with bioreactors. Chemostats, a specific type of bioreactor, are designed to have a continuous flow system which constantly provides fresh medium to an ongoing bacterial culture. To run this system, a manifold ensures that excess medium is extracted from growing cultures, and this volume must be equivalent to the fresh nutrient input rate[208, 226]. As a whole, chemostats allow

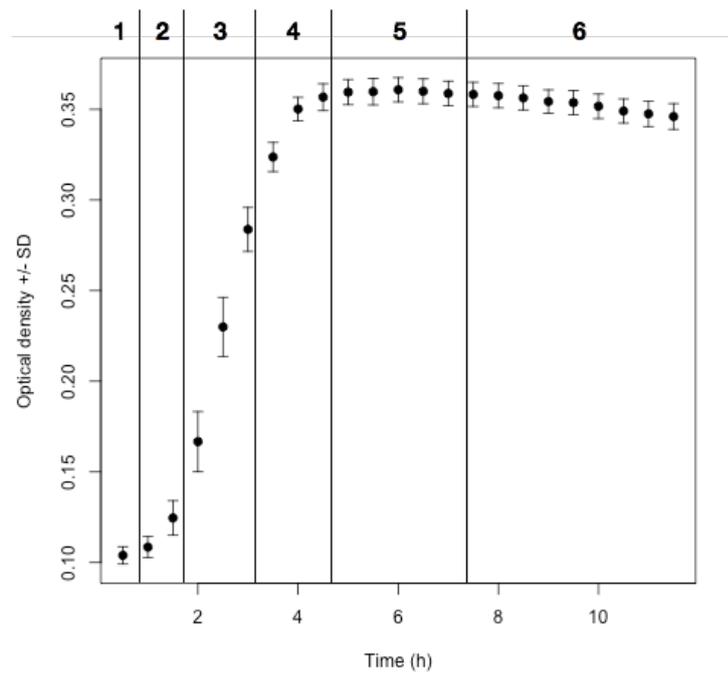


FIGURE 3.11: Bacteria undergo 6 consecutive phases during growth: (1) the lag phase (null growth), (2) growth acceleration, (3) exponential growth, (4) late-exponential growth (or retardation phase), (5) stationary phase (or no more growth) and (6) biological decay. Error bars show standard deviation based on 12 individual samples.

cultures to be grown at a constant rate that can be controlled via a pump settings, where the best rate is equivalent to bacteria growth rate measured in exponential phase. With the approximation of a culture growth rate, it is relatively straightforward to test and derive an appropriate chemostat dilution rate to maintain constant growth with a sufficient nutrient input, and to sustain near-optimal growth conditions for long periods of time.

Although mechanistic aspects of chemostats are well-defined, a thorough calibration process is necessary to obtain relevant data. As shown in Figure 3.12 (assembly), tubing first needs to be cut and assembled with connectors in order to build the manifold[3, 216]. In our setup, four 50ml vessels could be followed simultaneously and it was important to match all tubing dimensions in order to get an identical flow rate across multiple samples. Calibration could be carried out with water, since LB is water-based and both substances share similar viscosity, and was repeated after instruments sterilisation at least twice to make sure culture dilution rates remained stable. Once the chemostat calibration was verified and validated, all autoclavable units were sterilised and prepared for assembly immediately after autoclaving, as shown in Figure 3.12.

For this study, the benefit of using a chemostat was to obtain accurate predictions of the number of bacterial replication events and thus information regarding barcode replication cycles. Given a bacterium growth rate, it is easy to program a calibrated chemostat with settings that would comply with these specific growth requirements. Therefore, the number of replication cycles can be precisely obtained by actively diluting growing cultures at a specific rate. However, cells first need to reach exponential growth, and the number of generations between a culture inoculation and exponential growth is usually a rough estimate. As a standard method in microbiology, one can estimate the number of viable bacteria inoculated onto a plate via the number of colony forming units (CFUs) obtained after overnight growth. For the purpose of this study, we performed CFU growth curves and serially diluted cultures over time to obtain ideal dilution rates at which a minimal number of cells could be used to inoculate a chemostat. Given this minimum amount of cells, it was possible to evaluate, given a specific growth rate, the time needed for cultures to attain exponential phase (*e.g.* when the chemostat continuous flow should be turned on). Although chemostat calibration defined the exact dilution rate of growing cultures, it always took between 3 and 5 vessel volumes of fresh medium for bacteria to reach steady-state growth. Therefore, along all chemostat experiments, we recorded optical density (OD) measurements while sampling cells for barcode sequencing to ensure estimated dilution rates were accurate and cultures could reach steady-state growth.

Figure 3.13 displays an overview of the chemostat we built in this study. Panel

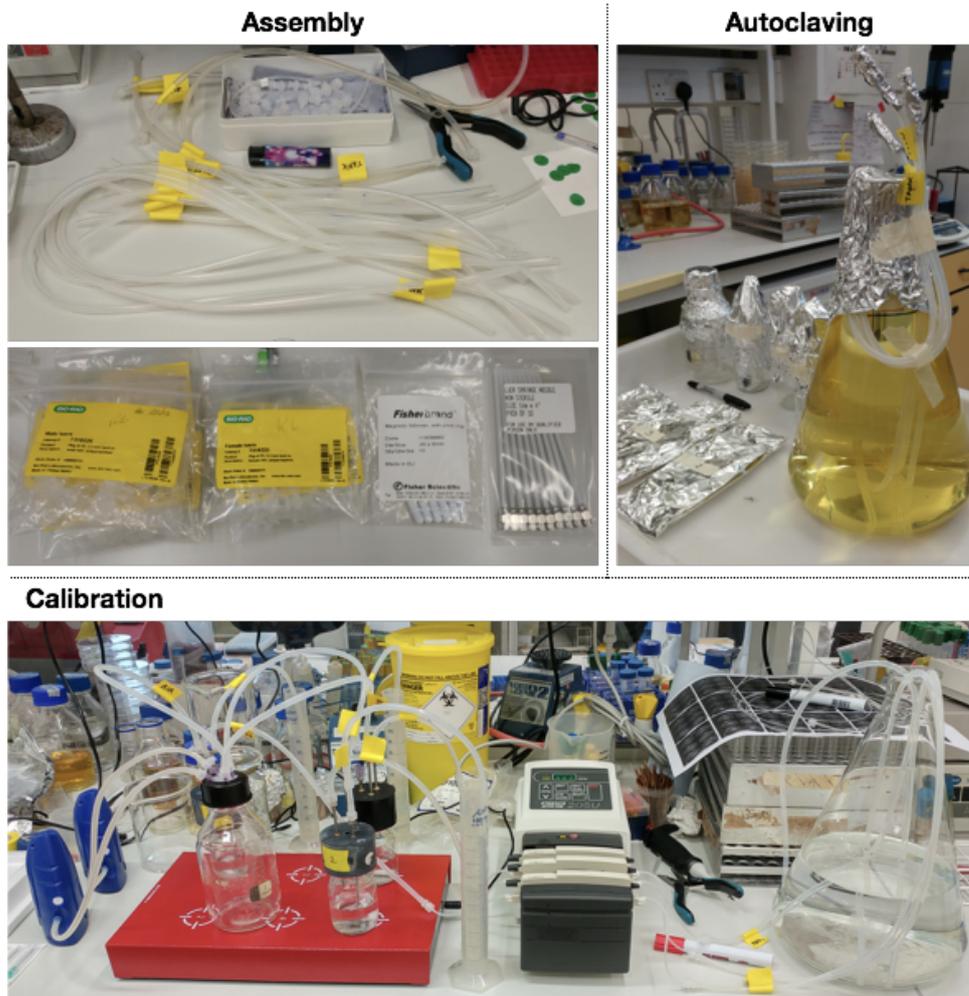


FIGURE 3.12: The setup panel (top left) outlines the tubing necessary to connect the multiple components of a chemostat with their associated plugs (male/female luers and syringes), without forgetting a magnetic stirrer to keep continuous cultures well-mixed. Autoclaving (top right panel) must be performed on all sterilisable equipment in an organised manner, since chemostats should be started as soon as their equipment has been autoclaved. Calibration (bottom panel) can be performed with water and should be timed and measured to match bacteria optimal growth rate.

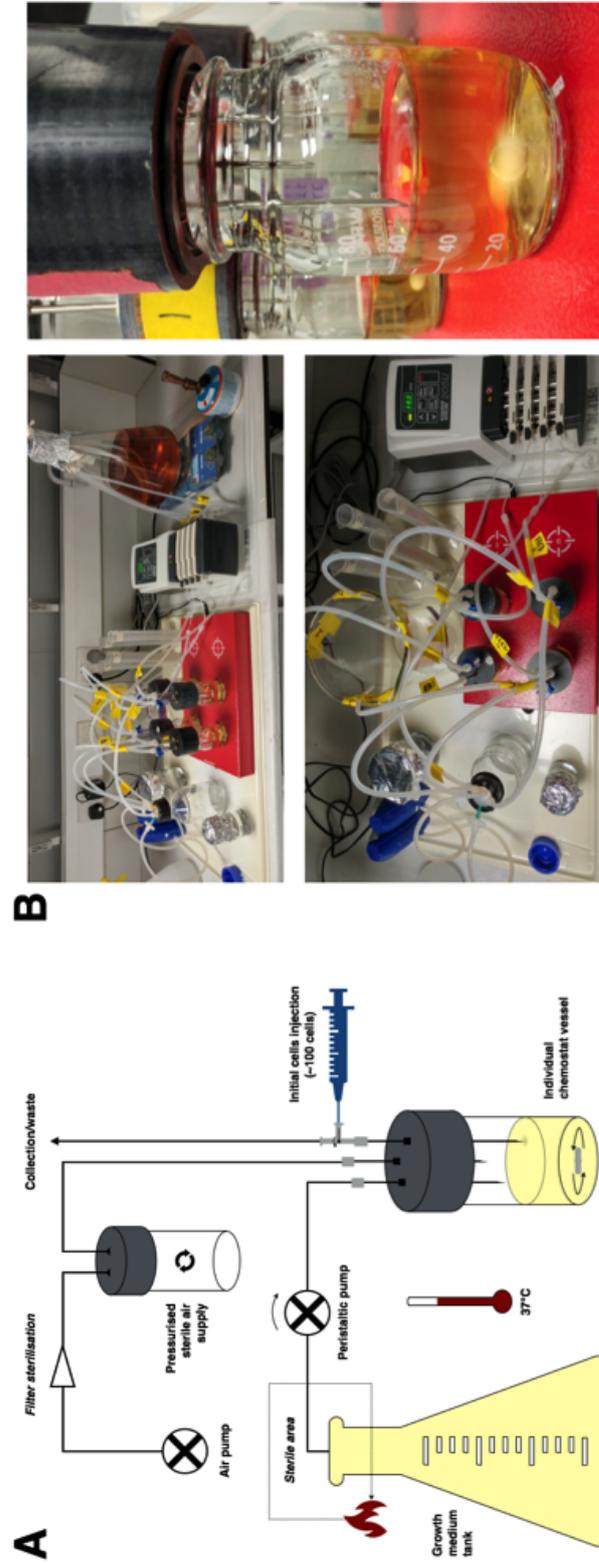


FIGURE 3.13: (A) shows simplified schematics of the chemostat manifold and components. (B) top-left picture shows the actual setup featuring 4 individual replicates. (B) right hand side picture is a zoomed in image of culture vessels, allowing to see differential needles level while (B) bottom left picture is a top view of the chemostat manifold

A details schematics of the continuous flow system: under a constant vacuum, fresh medium is pumped from a reservoir via a peristaltic pump to provide nutrients to an ongoing bacterial culture. Air pressure allows to keep the cultures dilution with fresh nutrients balanced with waste removal. This is done via a custom fusion of an Akta head to a Duran bottle cap, photographed in Figure 3.13B on the right. This picture allows to see three 14G serological needles set at different levels to adjust the bacterial culture volume with input, output and pressure ports. These caps were used to seal culture vessels in combination with a circular piece of silicon to act as a joint. We also coated silicon with vaseline to keep vessels sealed after autoclaving and to ease setup of the chemostat manifold. Pressure problems often come as the main issue in the setup of chemostats and setting up four individual vessels per experiment always allowed us to keep at least three of them under appropriate experimental conditions.

Here, we detailed the system through which we studied the replication of DNA barcodes over an exact, large number of bacterial generations. However, the experimental conditions of these assays do not represent general growth methods that are usually undertaken when growing bacteria. Therefore, we also studied DNA barcodes via a "laboratory-like" fashion, where cells would be grown and rediluted for an extensive number of times between freeze/thaw cycles. The next paragraph details the setup of this assay.

3.3.2 High-throughput screening

In large scale combinatorial or screening studies, many biological replicas or mutants can be observed simultaneously[97, 230]. Following a series of individual cultures allows eventual growth defects to be detected. These defects are unlikely to be observed on a small scale, as they may emerge in one over a hundred or more cases[14]. Extended periods of slow growth or starvation (such as stationary or decay phase conditions, for instance - Figure 3.11 phases 5 and 6) naturally appear to be stressful to bacterial cells. Therefore, redilutions from overnight cultures can propagate cells that acquired suppressor mutations to get a better fitness to stress conditions.

Automated plate handling systems now provide a strong advantage to screen large populations of bacteria. Experimental procedures such as incubation periods, growth measures and redilutions can be relatively automated to minimise human-induced errors and measurement imprecision[96]. Our high-throughput studies allowed the comparison of 384 subcultures of barcoded and control cells over a number of stationary phase redilutions, and to observe any changes represented in growth defects. Besides

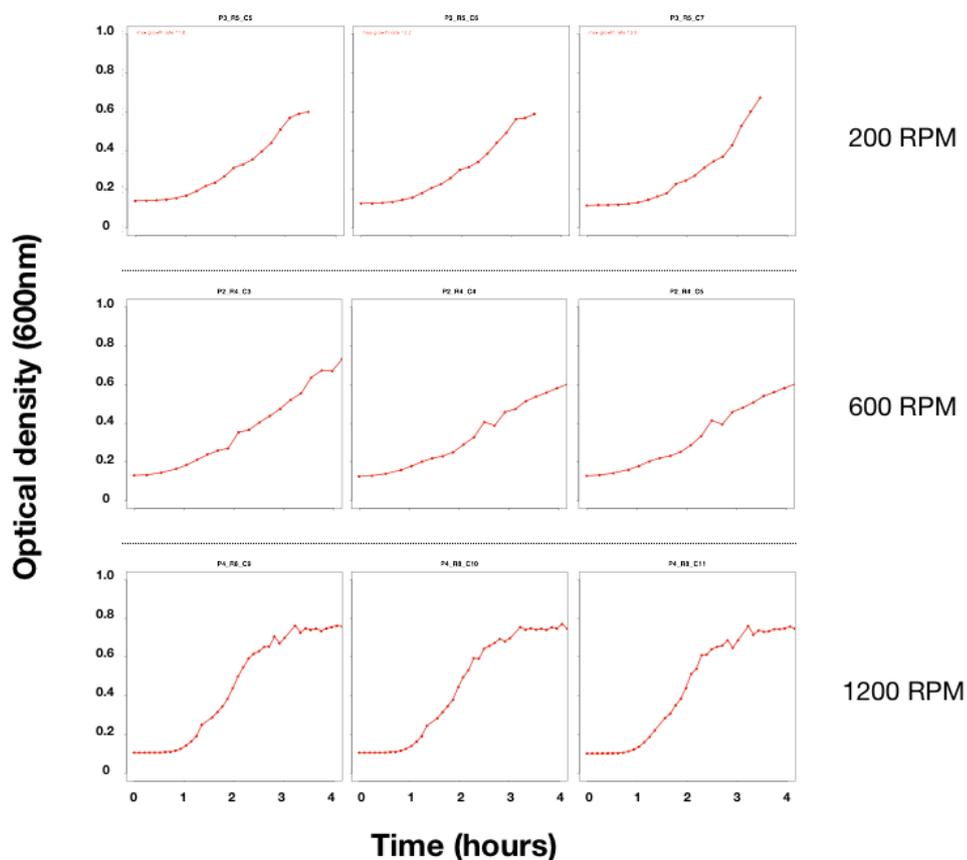


FIGURE 3.14: *B. subtilis* growth curves in LB supplemented with 0.4% glucose at 37°C with different shaking parameters: 200rpm (A), 600rpm (B) or 1200rpm (C).

growth characterisation, all barcoded samples were sequenced to uncover any potential mutations in DNA barcode sequences. These experiments were organised jointly with Dr. Peter Banks at the High Throughput Screening Facility (HTSF) at Newcastle University.

At the facility, we used a liquid handling robot (Beckman Coulter Biomek FX) and an automated plate reader to handle 8 individual 96-well plates simultaneously. Bacterial cultures needed to be conditioned for the plate reader assay via the liquid handling robot. An integrated graphical user interface on a computer set next to the equipment allows the programming of liquid handling steps one-by-one. A moving head hovers around specific locations (corresponding to media, plates and tips) and follows instructions to aspirate, dispense and mix different liquids. Before starting a growth experiment, plates were also sealed by an oxygen permeable membrane. Once all plates were ready to use, they could be placed on a rotating and shaking platform and connected to a plate reader for absorbance readings, and a growth assay could start.

Classical plate readers are generally designed to handle one plate at a time. For bacterial culture, different bacteria may require different shaking settings and handling a single plate at a time allows to shake cells continuously while the machine is not busy taking measurements. However, the equipment at the HTSF could only shake two plates in parallel while other samples would have to sit in the incubator. Realistically, this meant that plates would be shaken for 1min, absorbance readings would be taken and they would stand still for about 10min before the next shaking/measurement cycle. Although that was not a problem for *E. coli* cells, *B. subtilis* needs a thorough shaking to provide bacteria enough oxygen in order to proliferate. Therefore, growing all bacteria simultaneously needed adjustments to accommodate the growth of both model organisms. Figure 3.14 shows how higher shaking parameters could comply with a good and replicable *B. subtilis* growth. Such high-speed shaking settings would probably be detrimental if sustained for prolonged periods of time, but 1min shaking at 1200rpm every 10min appeared to fit both *E. coli* and *B. subtilis* species growth conditions.

In this assay, we followed barcoded and control strains over 10 subculture experiments. Initially, a single colony from each strain was picked from a fresh plate and grown in LB supplemented with 0.4% (w/v) glucose overnight at 37°C with regular shaking parameters (about 150rpm). In the morning, saturated cultures were spun down, resuspended in fresh LB medium, diluted 100 times and 200µl were loaded onto ThermoFischer clear 96-well microplates. For all subculture experiments, two conditions were tested: an early stop of bacterial cultures after 6h (in late exponential/start of stationary phase) and a prolonged culture in stationary phase (12h) before snap-freezing. For the first subculture, 100µl were harvested after 6h for the early sampling point, and the remaining bacterial culture was further incubated up to 12h. All subsequent cultures were diluted 100 times from frozen stocks and cultivated in a 100µl total volume for both early and late sampling points. By the end of the 10 subcultures, genomic DNA (gDNA) was extracted and screened for potential variations in barcode sequences.

3.4 DNA barcodes *in vivo* characterisation

3.4.1 Cloning method

In this study, we created a cloning method to barcode bacteria with a synthetic sequence. We applied this method to either barcode strains for the first time, or to update a strain already barcoded. As shown in Figure 3.7, the *in vivo* step firstly involves

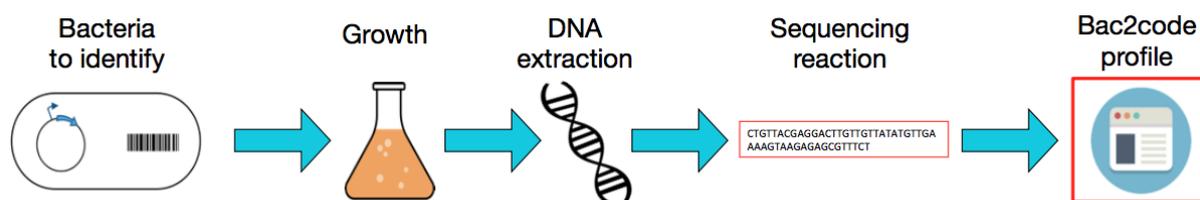


FIGURE 3.15: Process to retrieve a barcoded strain barcode information. Bacteria to be identified are grown for genome extraction and a PCR reaction is carried out to target barcode sequences.

the insertion of a barcode alongside a selection cassette. This selection marker is then removed to leave only the barcoded sequence on the genome. For biological validation, we carried out PCR screening of transformant colonies and sequencing of PCR products (Figure 3.15). In *E. coli* MG1655, 47 colonies were screened for insertion of the barcode cassette/removal of the antibiotics marker and all transformants were found to be positive by PCR and sequencing data. The cloning method in *E. coli* is based on helper plasmids that are selected at 30°C in the presence of ampicillin ($100\mu\text{g}/\text{ml}$). In all cases, plating on NA/Ampicillin plates did not allow any bacterial growth, as expected to prove the loss of helper plasmid after recombination. In the Gram-positive species *B. subtilis* 168CA, we obtained a 93% success rate for integration of the barcode cassette with a zeocin antibiotic marker ($20\mu\text{g}/\text{ml}$). The PCR screening revealed 31/32 products of the right band size, and all of the 31 expected positive clones were confirmed by sequencing. The second step of antibiotic marker removal, associated with the expression of the MazF toxin, was successful for 35 of 46 screened colonies (76%). While 35 transformants revealed the expected sequence, negative results always appeared to have developed mutations within a xylose-inducible promoter, resulting in negligible amounts of MazF toxin being produced. This frequency of suppressor mutations was observed over a range of xylose concentrations from 0.2% to 1%. The procedure detailed in Figure 3.15 approximately represents 10 replication cycles of DNA barcode sequences. After these initial 10 generations, we were able to confirm barcode sequences by Sanger sequencing for all positive PCR products. As a final validation step to prove the feasibility of our cloning method, three additional research scientists performed the cloning of DNA barcodes in two other strains of *E. coli* and *B. subtilis*, and managed to retrieve expected barcode sequences in *E. coli* BW25113 and *B. subtilis* PY79.

3.4.2 Growth characteristics

Primarily as a control to verify barcoded strain fitness compared to their wild-type, we performed a plate spotting assay to detect any replication defects visible on plates. For the preparation of chemostat experiments, we evaluated the minimal number of colony forming units (CFUs) and assessed growth rates from CFU-based growth curves.

3.4.2.1 Spot-plating assay

If bacteria show defective DNA replication, it directly affects division time and bacterial colonies grow at abnormal rates. After spotting $5\mu\text{l}$ of 10-fold dilutions of control and barcoded strains in exponential phase, we did not observe any significant changes in either bacterial strains ability to replicate (Figure 3.16). For the *Bacillus* strain, the ability to produce spores was unaffected. Spores are detectable to the naked eye as they block more light, due to their thicker cell wall, and appear as darker colonies than *E. coli* for instance.

3.4.2.2 CFUs counts

Initial plate spotting assays were followed by colony forming unit (CFU) growth curves. Over bacterial growth, we plated serial dilutions of cultures on nutrient agar (usually 10^{-6} , 10^{-7} and 10^{-8}) and captured images after 24h growth via an Epson scanner (48-bit colour and 400dpi). All images were processed via a custom protocol made with ImageJ to extract background information and detect reliable CFUs (away from edges and under a certain density). The appropriate number of CFUs detectable per plate was set from 30 to 500 CFUs/plate. Images were first transformed in 8-bit greyscale and thresholded to erase the agar background. A restricted selection, excluding petri dish edges, was further used to detect particles (individual CFUs). CFUs were detected on the basis of their sizes (from 3 to 300 pixels) and their circularity index (elliptic vs. circular structures). Table 3.4 shows barcoded and control strains estimated growth rates. These specific growth parameters were used to tweak continuous flow culture parameters and allow optimal growth in chemostats.

3.4.3 Barcodes stability over long periods of continuous flow culture

We assessed DNA barcodes integrity over a number of DNA replication cycles by sequencing samples from a chemostat. For the best approximation of barcode replication cycles, we inoculated the continuous flow culture apparatus with the minimal known

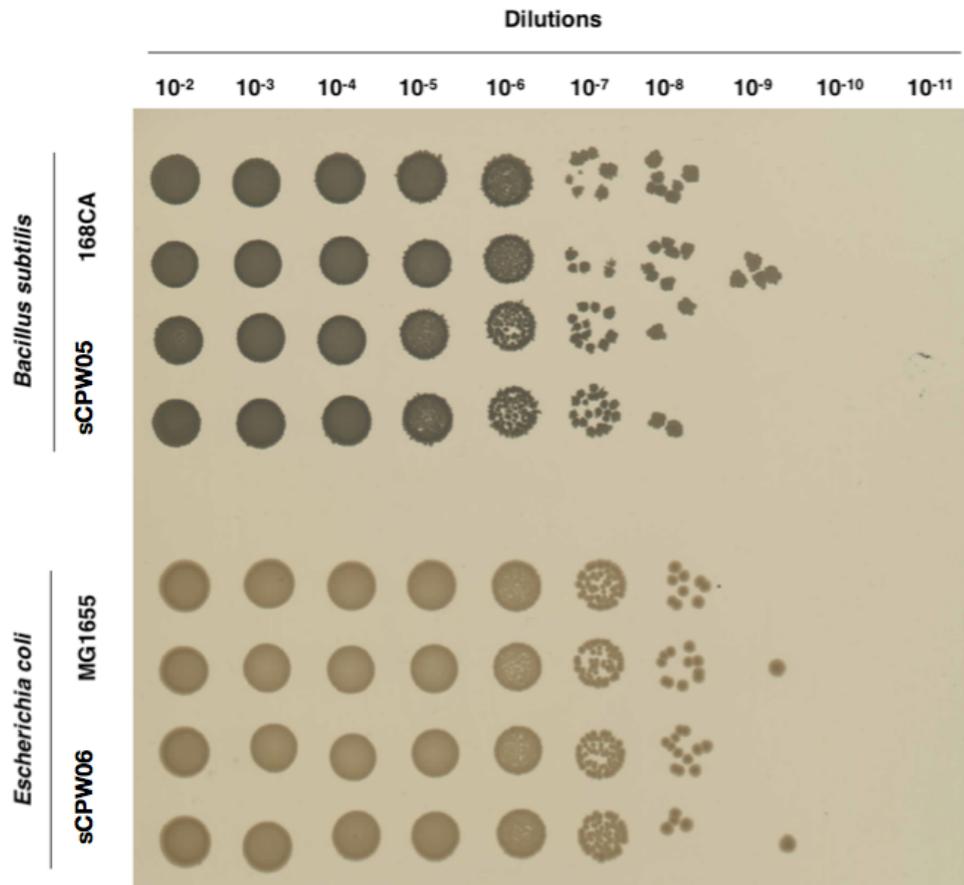


FIGURE 3.16: The photograph was taken after 24h growth and remained as such up to 72h. sCPW05 denotes the barcoded *B. subtilis* strain while sCPW06 is barcoded *E. coli* MG1655.

TABLE 3.4: Barcoded strains growth rate in *E. coli* and *B. subtilis* compared to wild-type strains.

Species	Strain	Growth rate (h^{-1})	Generation time (min/sec \pm sec)
<i>Escherichia coli</i>	MG1655 control	3.00	20m00s \pm 47s
<i>Escherichia coli</i>	barcoded MG1655	2.86	21m00s \pm 32s
<i>Bacillus subtilis</i>	168CA control	3.53	16m59s \pm 34s
<i>Bacillus subtilis</i>	barcoded 168CA	3.45	17m22s \pm 28s

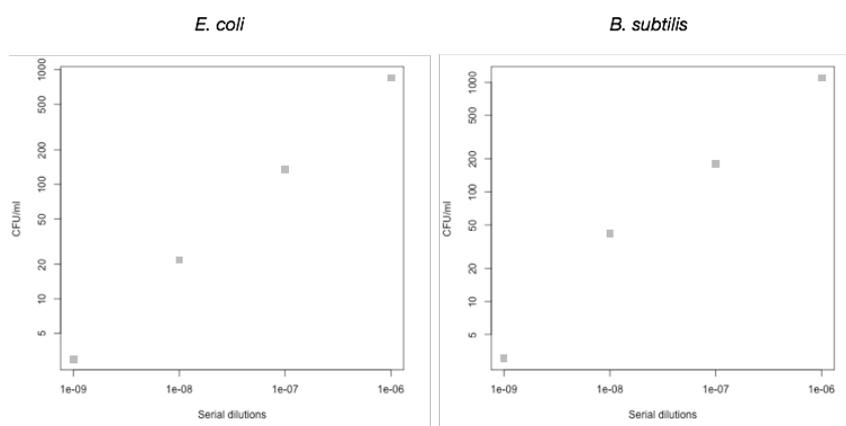


FIGURE 3.17: *E. coli* (left) and *B. subtilis* (right) strains were serially diluted from overnight cultures to assess their minimal achievable CFU dilution factor.

amount of cells that could lead to viable colonies. According to Figure 3.17, we approximated minimal achievable CFUs for each studied species (*E. coli* and *B. subtilis*) and started chemostat experiments from 1 to 3 CFUs (50-150 cells, 3.10^{-10} dilution) in triplicates. We turned the continuous flow on when bacterial strains reached exponential phase ($OD_{600} = 0.4$ after about 10h). This allowed us to provide fresh sterile medium and to remove excess culture at the right time not to overdilute bacterial cultures nor to let them slow down growth in stationary phase. As shown in Figure 3.18, bacterial populations undergo two phases of higher/lower OD before stabilising in a long-term steady state growth. We followed bacteria over 200 generations (about 4 days of culture) and retrieved barcode information at regular time points (every 15-25 generations). For *E. coli* the same cultures were followed over 200 consecutive generations, while in *B. subtilis*, cells were followed for 100 generations, induced to stress and sporulation by ethanol treatment and regrown from spores for a further 100 generations. Barcode DNA sequences were obtained by Sanger sequencing of barcode PCR products of specific genomic DNA (gDNA) amplification. We analysed 128 sequencing reactions, including 24 controls to compare the evolution of barcoded vs. wild-type strains. We confirmed that control wild-type sequences remained unchanged and found no variation in barcode sequences over 200 generations.

3.4.4 Barcodes stability in high-throughput subcultures

We evaluated the potential for spontaneous mutations in DNA barcodes in a high-throughput assay, where barcoded and control strains were subcultured in 96-well

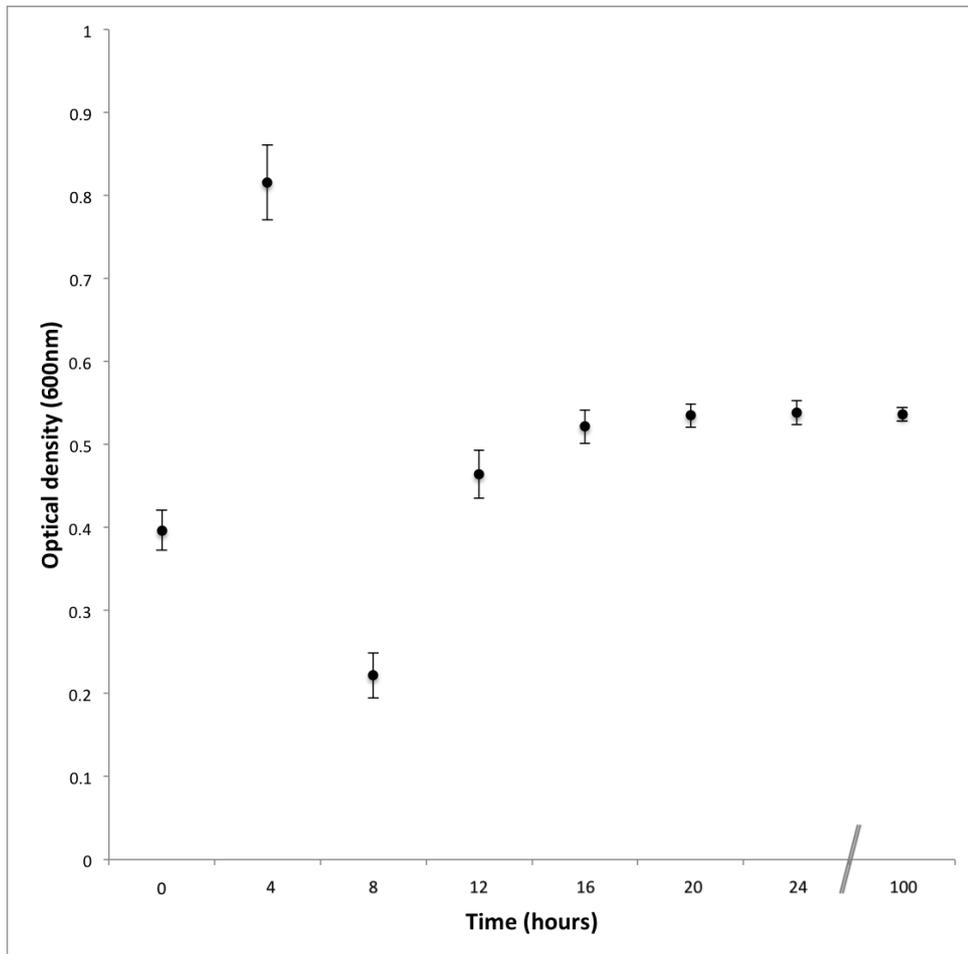


FIGURE 3.18: Optical density (OD_{600}) measurements initially vary in continuous flow bacterial cultures. Cultures adjust their growth to nutrients limitation and rapidly reach steady state, where same optical densities can be observed up to hundred hours (or more). OD measurements were averaged from 4 individual experiments, each performed in triplicate. Error bars show standard deviation across samples.

plates. Over 10 subcultures, we estimated from 100 fold dilutions of previous subcultures that final samples reached about 100 generations. Figures 3.19 and 3.20 shows the overview of the high-throughput method. From an initial subculture sampled at early and late time points for 2x96 (*E. coli*/*B. subtilis*) barcoded strains, serial subcultures were performed before sampling for sequencing. We extracted gDNA, amplified the genome region of interest (barcode integration locus) by PCR and sequenced all reactions. In our assay, sequencing of 384 barcoded strains tested in a normal vs. stress conditions revealed intact DNA barcode sequences. As shown in Figures 3.19 and 3.20, no major difference in growth rates was observed for different samples. The majority of sequencing reads left a 26-27 nucleotide gap downstream of the universal primer binding site and then showed a perfect match with the expected alignment. In less than 5% of cases, sequencing data quality was noisy but a second complementary read would always manage to recover the integrity of a barcode sequence. This, in particular, is the reason why we developed DNA barcodes with an algorithm including an error correction feature.

3.4.5 Barcodes readability

Due to the technicalities of Sanger sequencing, it would not be expected to obtain fully complementary sequences from the first nucleotide after barcodes universal primer binding site. As detailed above, most sequencing reads show a gap in their 5'-end that is formed at the start of the sequencing reaction. Sequences obtained from Sanger sequencing and chromatogram data (Figure 3.21) generally show a perfect match with barcode sequencing after 20-30 nucleotides. This was important to retrieve DNA barcodes 3'-ends containing checksum elements. In order to map retrieved DNA barcodes, our server recognise an encoded synchronisation sequence and looks for upstream (barcode UUID) and downstream (checksum) DNA. The checksum accounts for mismatches in the 96bp encoding a barcode UUID, and always allowed us to retrieve the correct barcoded strain information, even with gaps of variable length in sequencing reads. To verify that missing 5'-ends would indeed match the expected barcode, we repeated a sequencing reaction with an alternative universal primer (*E. coli* and *B. subtilis* specific). A second read always allowed us to retrieve the integrity of correct barcode sequences.

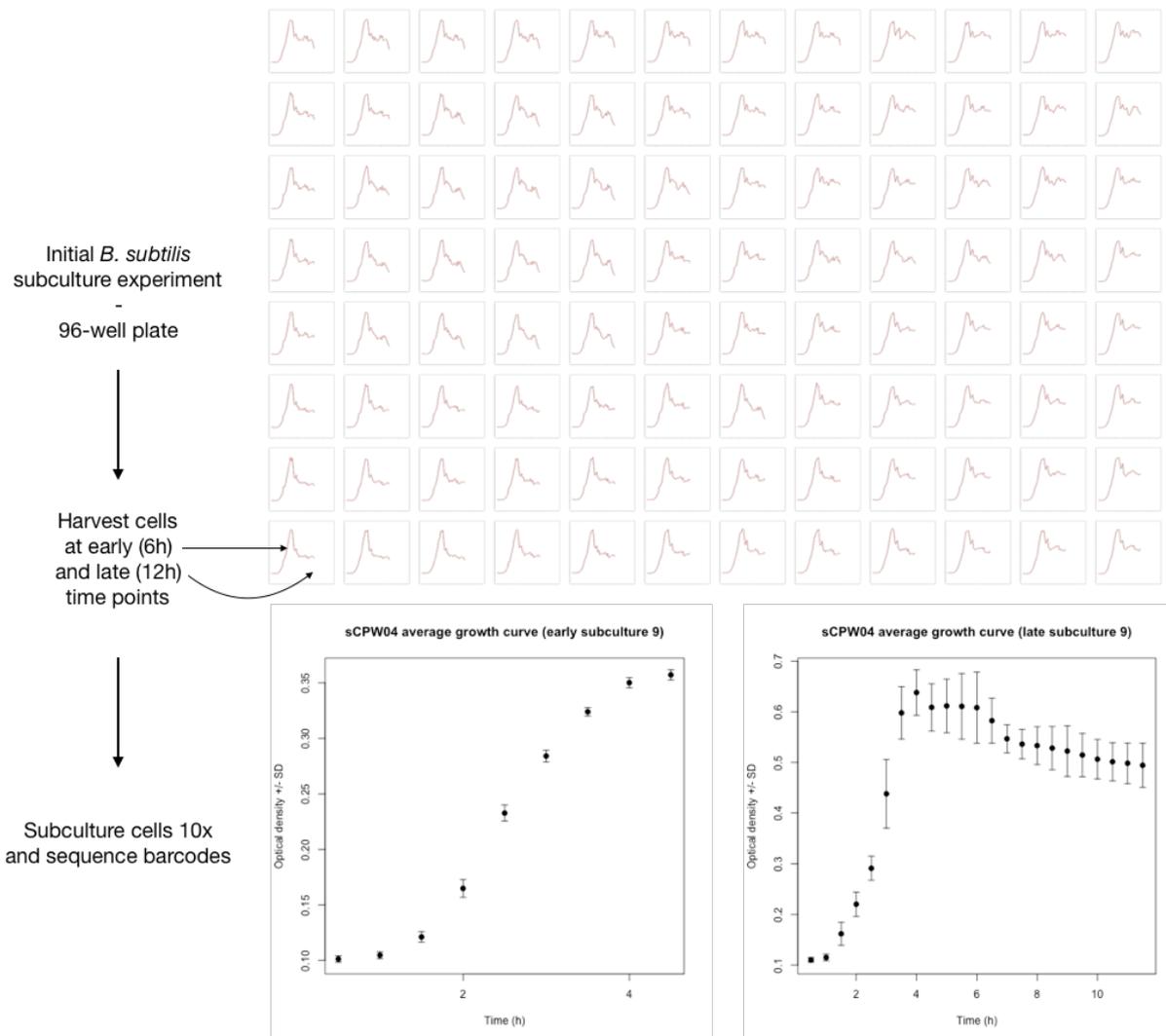


FIGURE 3.19: Top and bottom panels describe the high-throughput assay for *B. subtilis*. Top panel shows growth of the initial 96-well plate that was used throughout this assay, sampled after a 6h incubation period or a prolonged 12h period where cells are left in stationary phase. Nine subsequent cultures – following early/late samples separately – were carried out before extracting DNA for sequencing. In the bottom panel, growth curves are averaged for the 96 samples of the same plate after the ninth subculture experiment, and error bars show standard deviation.

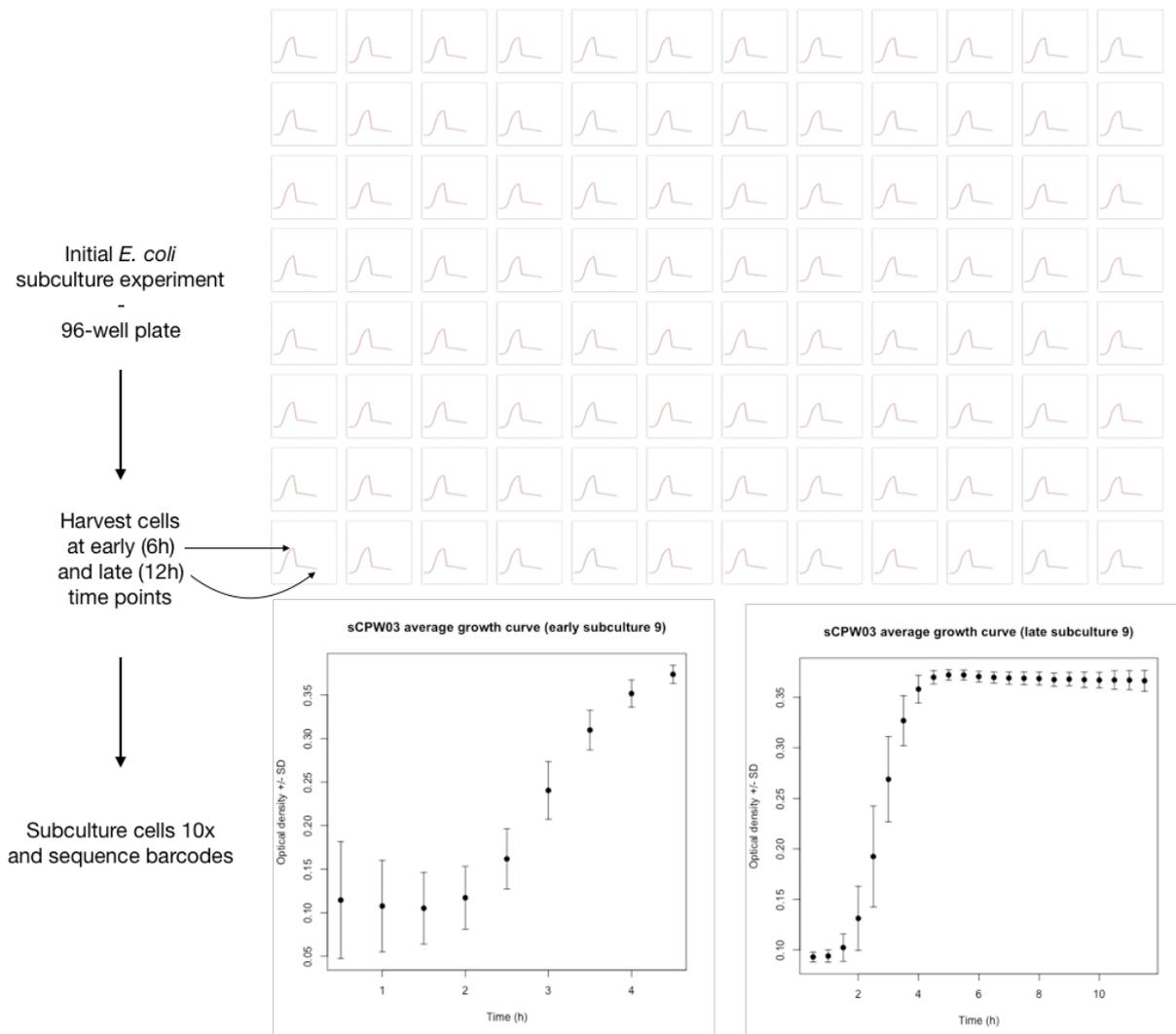


FIGURE 3.20: Top and bottom panels describe the high-throughput assay for *E. coli*. Top panel shows growth of the initial 96-well plate that was used throughout this assay, sampled after a 6h incubation period or a prolonged 12h period where cells are left in stationary phase. Nine subsequent cultures – following early/late samples separately – were carried out before extracting DNA for sequencing. In the bottom panel, growth curves are averaged for the 96 samples of the same plate after the ninth subculture experiment, and error bars show standard deviation.

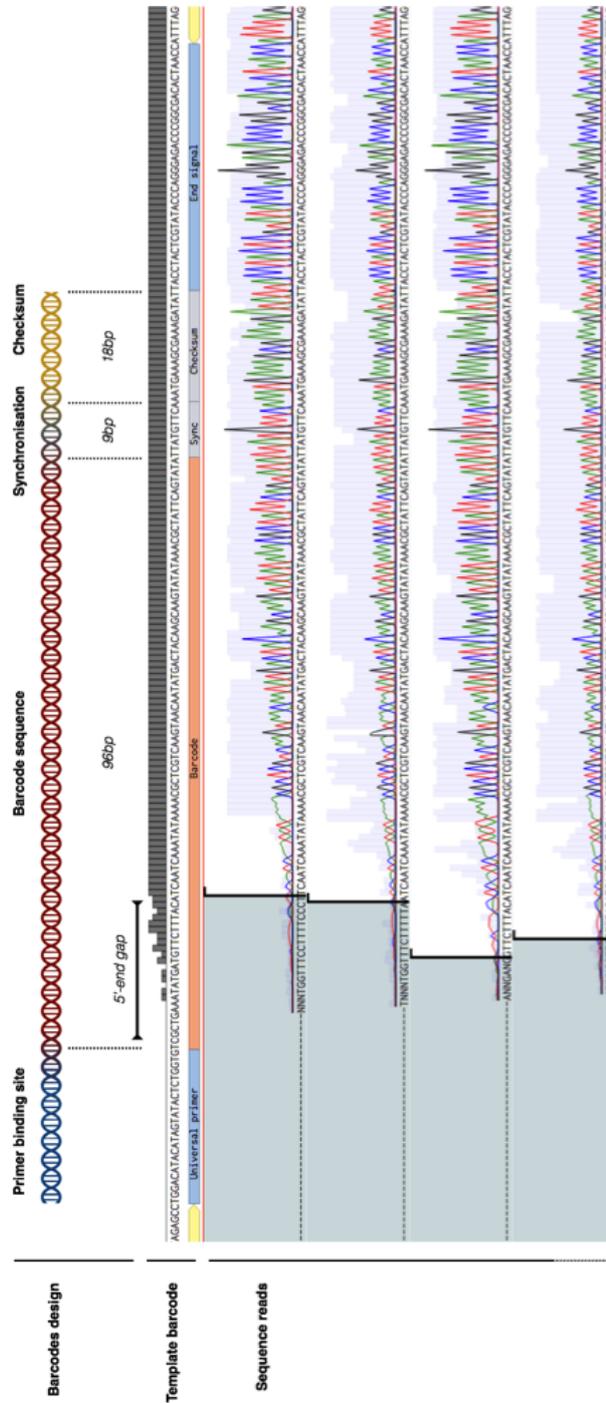


FIGURE 3.21: Typical sequencing reads alignment for barcode sequence retrieval. A universal primer is used to read barcode DNA sequences, which leaves a 5' gap in reads, then shows perfect alignment with expected template sequence. Barcodes 3'-end alignment matches their checksum, allowing recovery of missing/inaccurate 5'-end fragments. Here, four individual reads of the same DNA barcode are aligned to a template sequence and covered in grey at 5'-end mismatches. Height of the different peaks in sequencing reads describe nucleotide incorporation quality: higher peaks are associated with good quality data while lower peaks are of poor sequencing quality.

3.5 Considerations in the study of DNA barcodes

Living cells rely on DNA replication to proliferate in a given environment. In *E. coli* and *B. subtilis*, DNA polymerase enzymes initiate bidirectional replication of the chromosome at the *oriC* region[166, 192, 130]. In normal growth conditions, both halves of the chromosome are simultaneously replicated and a second copy of the chromosome is segregated to a daughter cell, subsequently created by cellular division[21]. DNA polymerases are high-fidelity enzymes that misincorporate or skip nucleotides at a 10^{-6} to 10^{-8} error rate[130]. In bacteria, they are connected to other families of repair proteins and DNA replication exhibits an overall error rate of 10^{-9} to 10^{-10} per nucleotide. In practice, one may observe up to 10 mutations at every cycle of DNA replication in a ~ 5 Mb genome. We have developed a method to tag bacterial strains with DNA barcodes and had to ensure that these sequences would remain stable over time. To test this, we followed the proliferation of barcoded cells over hundreds of replication cycles and sequenced barcode nucleotide sequences. In the following, we discuss different aspects of the setup of DNA barcoding protocols and their optimisation for the study of barcode stability.

3.5.1 Barcode integration

3.5.1.1 Setup of a universal cloning method

We presented evidence for successful integration of DNA barcodes in *E. coli* and *B. subtilis* strains, at two respective loci. Although many cloning techniques may have been applicable to barcode insertion, homologous recombination was the best compromise to obtain a consensus protocol between bacterial species. The main advantage of using homologous fragments was that the procedure to barcode strains would remain the same, regardless of inserting a new barcode sequence or replacing an existing one by a updated version. Moreover, the adjustment of standard homology arm length provided the ability to control recombination specificity in both species. We started from $1kb$ homology arms upstream and downstream of barcode sequences, and observed identical cloning results with the use of smaller homology fragments ($\sim 500bp$). As a rule of thumb, $300bp$ or more are generally used for cloning via homologous recombination. Therefore, we could reduce barcode flanking homology regions even further, but this may also result in a decrease in the efficiency of the cloning method (*e.g.* total number of colonies obtained).

During cloning verification steps, we always identified correct transformant colonies and obtained an optimal ratio of true vs. false positives ($\geq 90\%$). However, we report

that the mini-MazF cassette is susceptible to develop mutations rendering the P_{XYL} promoter obsolete. This issue could be tackled by substituting the xylose-inducible system for a more stringent device, such as the P_{SPAC} or $P_{HYPERSPANK}$ IPTG-regulated promoters. Nevertheless, screening for the loopout of antibiotic resistance marker in *B. subtilis* still displayed a 76% success rate, and given the easiness of the cloning validation technique (plating transformants on +/- antibiotic plate), it remained trivial to find positive transformants. Altogether, we demonstrated the setup of a robust cloning method, analogous for Gram-negative/-positive species, and showed its applicability via multiple barcoding procedures that were proof-tested by three additional scientists.

3.5.1.2 Extending the barcoding technique

The current methodology to integrate barcodes in the bacterial chromosome uses antibiotic selection markers. By default, we used chloramphenicol and zeocin for Gram-negative and Gram-positive species respectively. Since a given design may already use one of these antibiotics, we provided the possibility for recombinant DNA assembly to be performed with alternative selection cassettes (kanamycin, tetracycline, erythromycin, *etcetera*). Although the barcoding method remains a two-step process with (i) insertion of barcode sequence and (ii) curation of selection marker, it accommodates both species and displays good efficiency at barcoding strains.

In order to streamline our initial cloning method, Jonathan Tellechea is now working on adapting this technique to the CrispR/Cas system. With this type of genome-editing tool, there is no need for a selection marker to be propagated along with insert DNA[134]. Therefore, this reduces the integration process to a single transformation step and alleviates the need for selection marker curation. However, this approach is slightly less flexible for the replacement of existing DNA barcodes. This is because to promote its activity, the CrispR/Cas system recognises specific DNA motifs flanking a genome region to be edited. For the updating of barcode sequences, this would require two different sites to be used, thereby the development of two genome editing tools. Yet, the CrispR/Cas system is a very powerful tool and an increasing number of applications use this method for genome editing. With complementary cloning approaches, we aim to expand the range of techniques that could be used to barcode strains and to facilitate a high-throughput identification of bacterial species.

3.5.2 Large scale characterisation studies

We demonstrated in the chemostat and high-throughput assays the stability of barcode DNA sequences over a large number of bacterial generations. While we successfully derived that DNA barcodes were stable over time, we also encountered specific equipment challenges during these experiments. In the following subsections, we present the main limitations of chemostat and high-throughput assays for the study of bar-coded strain stability.

3.5.2.1 Continuous flow system

Chemostats are a type of continuous flow system that allow constant growth of bacterial species in a well-controlled environment. Their setup and calibration process is of great importance; a thorough characterisation of flow rates, for a given device assembly, is necessary in order to exert an adequate control over the studied strains. Although this process should be repeated between autoclave cycles, it is relatively straightforward to obtain stable parameters for bacterial growth. The main limitations of chemostats are the amount of medium consumed and the waste produced by growing cultures. *Per se*, preparing large volumes of medium is not an issue. The trickier step was to procure this fresh medium to the chemostat input reservoir, and to keep sterile conditions over several iterations of this process. Since the chemostat apparatus was placed in a walk-in incubator set at 37°C , the chances for contamination due to other people's nearby cultures were higher. In order to limit the likeliness of this scenario, we always worked in proximity of a bunsen burner to create a sterile environment nearby the input reservoir whilst topping it up with fresh medium. However, a strong air-flow needed to maintain ambient temperature in the incubator may sometimes have impaired with the use of bunsen burners. As one would expect, we report that sample contamination was more likely to take place after prolonged period of culture for the same experiment ($\geq 100\text{h}$). If this situation were to occur, we would end the assay and reiterate it.

As we have detailed earlier, cultures were followed over hundreds of hours. The continuous flow of the chemostat was based on a peristaltic pump that could be programmed to run at different speeds. The tubing used with this pump had to be of thinner diameter than normal wiring of the manifold. In general, after 100h of experiment, we started observing silicon tubing degradation due to repeated pressure of the peristaltic pump (cf. Figure 3.22). In a couple of instances, chemostat assays failed because of silicon wire bursting. Such situations would normally only affect a single



FIGURE 3.22: Schematics of a peristaltic pump. Tubing passing by the pump is compressed at specific locations, which creates a flow movement from input to output ports. The pump rotation speed is usually digitally controlled and is expressed in number of rotations per minute.

bioreactor, but this would impede on the number of biological replicates used per assay. In order to prevent this type of occurrence, we preferred using new tubing for the peristaltic pump for each experiment. The rest of the manifold would otherwise be cleaned with 2L of 70% ethanol and 5L of water before being sent for autoclave instrument sterilisation. In order to preserve chemostats in a sterile working state, we recommend taking most care in the storage and handling of manifold pieces, and to perform a thorough examination of individual parts prior to the start of an assay.

3.5.2.2 High-throughput assays

Although we showed the successful retrieval of DNA barcode sequences in high-throughput settings at the HTSF (Newcastle University), we first needed to adjust growth parameters to accommodate both *E. coli* and *B. subtilis* strains growth profile. This procedure, as outlined in Figure 3.14, consisted of increasing shaking parameters to observe a good *B. subtilis* growth fitness. For both bacterial strains, we harvested cells from subcultures at early and late stationary phases, and snap-froze them in 50% glycerol. These two sampling points represented normal vs. stress conditions, where cells harvested at a later time would have more potential to have developed mutations. Nevertheless, we did not observe any difference for any of the screened barcode sequences. As a complementary approach to test the viability of barcoded cells, we also exposed bacterial strains to other stress conditions such as ethanol treatment or UV-exposure, and could nonetheless retrieve correct DNA barcodes after subgrowth of bacteria. The screen of a large amount of biological replicates helped us to assess the robustness of barcode sequence insertion in the bacterial genome and demonstrated artificial sequence stability over time.

3.5.3 Barcode design and sequencing data

The design of DNA barcodes was driven by the creation of bio-orthogonal artificial sequences that could be used as bacterial identifiers. During the development of this project, we set up a minimal architecture of barcodes that would allow for their retrieval via universal primers. Nevertheless, it is typical for Sanger sequencing to leave gaps at the 5'-end of sequencing reads. Figure 3.21 exemplifies this instance, where barcode DNA located directly downstream of the primer binding site could not be retrieved in a single reaction. Although a second sequencing read from the primer located downstream of the barcode insertion site could always help recover the entire barcode sequence, an alternative would have been to include a 30bp spacer between universal primer and start of barcode sequences. This method would have ensured the entire coverage of DNA barcodes in a single sequencing reaction, and would not have impaired with any of the cloning features of the barcoding protocol. Finally, it is worth noting that sequencing results from barcode stability studies were consistent with the expectations one could have prior to the assays. Since bacterial DNA polymerases are high-fidelity enzymes and only misincorporate a few - if any - nucleotides at every replication cycle[67], DNA barcode sequences would not be expected to be mutated even after 200 bacterial generations.

3.6 Summary

Here, we propose a unified platform for microbial tracking and bacterial strain identification through the use of inert synthetic DNA barcodes. The system works on the gathering of a strain experimental data together with *in silico* resources, and the storage of this information as short DNA links within barcoded strains. We discussed in this chapter the considerations involved in barcode design to efficiently tag two model microorganisms, and how to retrieve strain documentation by a simple sequencing reaction. Then, we demonstrated the efficiency of our barcode DNA cloning method. For the development of our barcoding platform, we adapted the experimental protocols that were used to manipulate barcoded cells to a barcoding kit, presented in Appendix C. Finally, we showed the stability of DNA barcodes *in vivo* over a large number of bacterial generations with the use of continuous flow and high-throughput culture equipment. Altogether, we set the foreground for the development of DNA barcoding technologies. To open towards barcodes scalability, we discuss in the next

chapter combinatorial genetic libraries that were characterised to populate our platform with bacterial signaling example resources, and discuss further aspects of this work in Chapter 6.

Chapter 4

Genetic circuits engineering

In Chapter 3, we introduced the Bac2code platform with the creation and characterisation of DNA barcode sequences. These sequences are used as identifiers to document the information about specific bacterial strains. After setting up the barcoding framework, we moved on to building a biorepository of genetic circuits to populate our platform with documented examples of barcoded strains. This chapter provides an overview of the construction and characterisation of genetic circuits that could be found on the Bac2code server.

4.1 Bacterial communication via quorum sensing

Cell-to-cell signaling requires the intervention of specific metabolites that act as activators/repressors over genetic subunits[40, 140]. Widely spread amongst the bacterial kingdom, quorum sensing is a universal way for bacteria to synchronise differential growth behaviour. Quorum compounds are signaling molecules that accumulate within a bacterial population and trigger a coordinated response over a certain threshold, dependent on cell density. Initially discovered in *Vibrio fischeri* by the emission of bioluminescence in the dark[120], quorum sensing has now been characterised in many Gram-negative and Gram-positive organisms as a mean of bacterial communication[155, 75, 206, 232]. For instance, *Vibrio fischeri* swims freely in seawater at concentrations of 10 cells for liter and does not produce any light in these conditions. However, bacteria can stick and develop with a variety of sea organisms in a symbiotic relationship. In such environments, bacteria can grow to concentrations of 10^{10} cells per liter and coordinate cellular responses, coupled to a process of luminescence emission[198, 197]. As an example, the Hawaiian squid *Euprymna scolopes* uses *Vibrio fischeri* as a light organ, where fluctuating bacterial cell densities allow it to be practically invisible, or to hunt in deep sea water whilst using bacteria as a light source.

In this chapter, we deconvoluted a quorum sensing system into sender and receiver components in order to characterise signal transduction in bacteria.

4.1.1 Quorum signaling molecules: the *lux* model

Over the years, bacterial behaviour changes based on the detection of specific cell densities has been referred to as quorum sensing[71, 31]. Quorum compounds are very specific signaling molecules mediating intercellular communication via autoinduction[99, 94, 143, 78]. In simple terms, a chemical species is produced by a protein that positively regulates its own expression. This mechanism, called a positive feedback loop, allows the amplification of signals as cellular density increases. There are two known types of quorum molecules in bacteria: Gram-negative species use N-acetylated homoserine lactones while Gram-positives use processed oligo-peptides[148, 254, 74]. We focussed on the engineering of *E. coli* species and thus specialised in a specific type of homoserine lactone (AHL).

We based this work on the use of *Vibrio fischeri lux* system. In this organism, the quorum sensing machinery is controlled by the bipartite LuxR/LuxI operon[217]. As shown in Figure 2.7, the left part of the operon drives the expression of the LuxR protein regulator, while the right part produces the LuxICDABEG transcript. The latter encodes the AHL autoinducer (LuxI) and a cassette containing bioluminescence genes (LuxCDABEG). In combination with AHL, LuxR activates transcription of LuxI and more of the quorum molecule is produced. In the absence of the autoinducer, LuxR represses the production of quorum sensing molecules.

4.1.1.1 LuxR regulator

Low constitutive expression of LuxR via the promoter P_{L-lux} produces dimer proteins that bind to the -35 unit of P_{R-lux} . The P_{R-lux} promoter uses a *lux* box as its -35 element. This box is a 20bp inverted palindromic repeat ('ACCTGTAGGATCGTACAGGT') and allows dimerisation of active LuxR proteins[84]. However, in the absence of AHL (quorum compound), LuxR is inactive and does not allow the RNAP to initiate transcription at this site. Binding of AHL to the N-terminal domain of LuxR allows its C-terminal DNA binding domain to become active. Therefore, the RNAP holoenzyme can only initiate its activity in the presence of AHL, which facilitates LuxR dimerisation on the P_{R-lux} -35 to act as an activator of the RNAP open-complex formation. There is a nonlinear relationship between concentration and response behaviour in quorum sensing processes[124, 242]. The response exhibited by LuxR on the P_{R-lux} promoter is

typical of DNA binding proteins, and is essential for signal recording in *in vivo* expression studies.

4.1.1.2 LuxI autoinducer

In *Vibrio fischeri*, the AHL autoinducer (3-oxo-C6 HSL) is the product of the LuxI protein catalytic activity[52, 70, 82]. Precursors that are found within the cytosol (acyl-ACP and S-adenosylmethionine or SAM) are converted by LuxI into AHL and become available to ease LuxR binding to P_{R-lux} . This, in turn, activates transcription of more LuxI protein resulting in higher autoinducer levels[219]. As we described earlier, LuxI is followed by the LuxCDABEG proteins in its native host and these are responsible for the emission of luminescence. Nevertheless, for the sake of this study, we replaced this large operon by a single fluorescent protein, offering better capabilities for *in vivo* characterisation. This positive feedback loop is very important for hysteresis of transfer curves in modelling studies. Here, hysteresis is shown as a lag in fluorescence emission created by a change in inducer concentration. This phenomenon is further described in the next sections.

4.1.2 Inducible quorum devices

As we detailed above, quorum sensing systems are found in a range of bacteria, and the *V. fischeri* Lux operon was cloned in *E. coli* and shown to display similar functions as in its native host[57]. In this Chapter, we describe the deconvolution of the *lux* operon into sender and receiver genetic circuits in *E. coli*[152]. Figure 4.1 describes the rationale behind building these circuits. Using generic methods for bioengineering described in Chapter 2 and the next section, we constructed a series of bacterial devices (plasmids) and characterised their behaviour *in vivo*. We replaced the *lux* emission of the luminescence system by the emission of fluorescence, and followed signalling molecules over a range of conditions using characterisation techniques providing single-cell to population scale resolution. The following sections describe the genetic assemblies built in this study to re-orchestrate quorum behaviour in *E. coli*. Bacterial systems were created, characterised, modelled and re-evaluated in order to provide the most efficient parameters for signal emission, detection and amplification in genetic circuits.

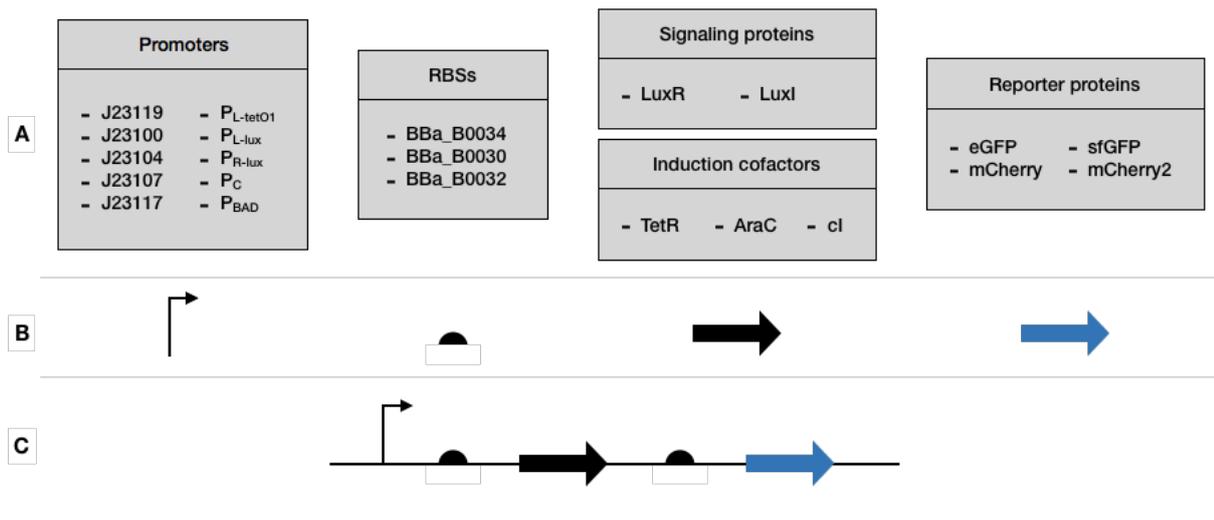


FIGURE 4.1: Assembly of genetic circuits in synthetic biology. (A) describes the different genetic parts that were combined to build genetic circuits. (B) shows a standard visual representation of these parts. (C) displays an example of genetic circuit functional assembly producing a protein of interest along a fluorescent reporter.

4.2 Construction of genetic circuits

As suggested in Figure 1.1, the construction of genetic circuits is an incremental process in which devices of similar function are optimised based on previously obtained characterisation results. In this section, we walk through the process of engineering quorum devices in *E. coli* for the study of intercellular communication. In particular, we explain the key features that allowed for the conversion of the LuxRI unit into a tightly controlled sender/receiver system.

4.2.1 Engineering the *lux* operon

In this study, we isolated the two main components of the LuxR/LuxI system. As shown in Figure 4.2, the *V. fischeri* Lux operon was divided into: (i) sender bacterial units controlling production of the quorum molecule and (ii) receiver cells detecting environmental quorum elements that would diffuse through their membrane. A third type of device was also created where detection of AHL allowed P_R autoinduction and formation of a feed forward loop. Experiments were performed in four strains of *E. coli*: *DH5 α* , *DH5 α λ pir*, *MG1655* and *BW25113*. We used *E. coli* *DH5 α* or *DH5 α λ pir* for cloning and *MG1655* and *BW25113* for expression of genetic constructs. Plasmids that were built in this study are detailed in Table 2.1 and followed two main design

principles: (i) optimisation of the inducible promoter to produce quorum sensing compounds and (ii) reduction of the background noise/response maximisation in receiver cells. This process involved basic modifications of genetic structures regulation. In further subsections, we describe genetic circuit variants that were created for *in vivo* characterisation.

4.2.2 Genetic circuits library

All plasmids were isolated and confirmed for correct insertion by colony PCR or restriction enzymes digestion. Positive isolates were then sequenced by Sanger sequencing to obtain exact sequence information. In cases of unexpected sequence mutations, plasmids were corrected by amplification from overlapping primers and selected by product recircularisation via *in vitro* mutagenesis.

4.2.2.1 Setup of bacterial terminators

The LuxR/LuxI system is natively controlled by a bipartite promoter. In order to separate the different biological entities for specification in sender/receiver circuits, we placed terminators in template plasmids to isolate target products from cofactor regulation. As detailed in Chapter 2, we used a fusion of the *rrnB* T1 and bacteriophage TE terminators to disrupt transcription of mRNA molecules between regulatory subunits. If short product PCR amplification fails due to secondary structures, we recommend using a custom synthesised terminator template for cloning. We built genetic circuits in two main template plasmids: pACYC184 for low copy plasmids and pUC19 for high copy replicons (Figure 4.3). For pACYC184 circuits, we reduced the plasmid size by knocking out the tetracycline resistance gene and replaced it by our terminators fusion. For pUC19 (and pSB1AK3), we inserted terminators in the standard cloning site. We used these plasmids as the basis for further genetic constructs assembly.

4.2.2.2 Choosing adequate ribosome binding sites

For the construction of a plasmid library and fair comparison of genetic constructs, we chose to place all expressed proteins under the control of the same ribosome binding site (RBS). We screened three different RBSs widely used in the synthetic biology community, which represented a set of high-, middle- and low-expression profiles (respectively B34, B30 and B32)[20]. We studied these sites in the context of the expression of a single, or two individual proteins as shown in Figure 4.4. We found that co-expressing proteins usually lowered the fluorescence profile of genetic constructs, as opposed to

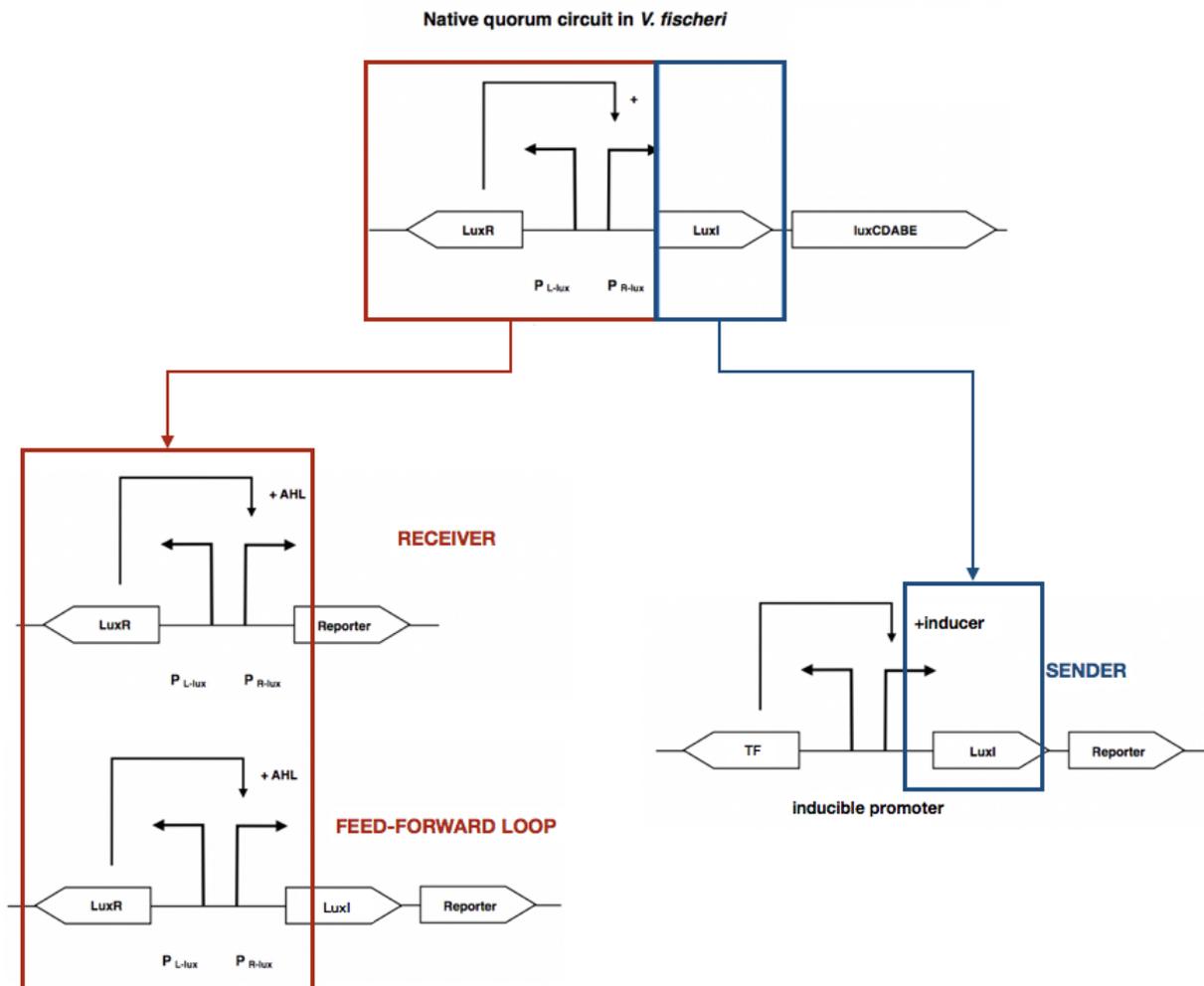


FIGURE 4.2: Schematics of the *V. fischeri* Lux operon repurposing. Sender and receiver units are created from subparts of the native operon, and remotely controlled by the use of a specific inducer or by the detection/amplification of AHL. The feed-forward loop model includes LuxI upstream the fluorescent reporter, providing more autoinducer to LuxR. For the sender unit, TF denotes a transcription factor that represses or activates the transcription of the autoinducer.

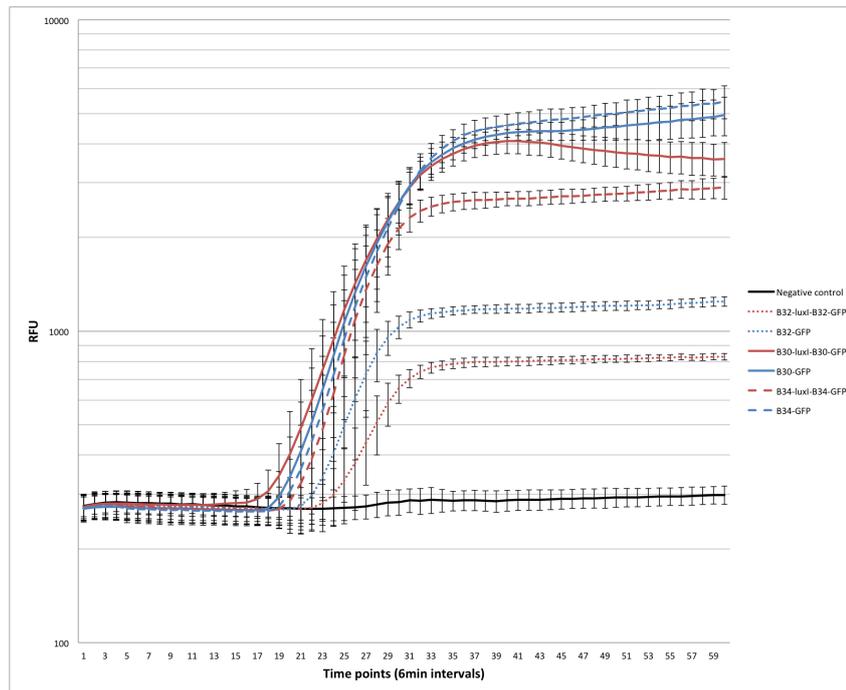


FIGURE 4.4: Ribosome binding sites characterisation: output fluorescence (RFU) over time for constructs with BBaB0034 (B34), BBaB0030 (B30) and BBaB0032 (B32) driving the expression of a single protein (GFP) or two individual proteins simultaneously (GFP and LuxI). Error bars show standard deviation over three repeats of the assay.

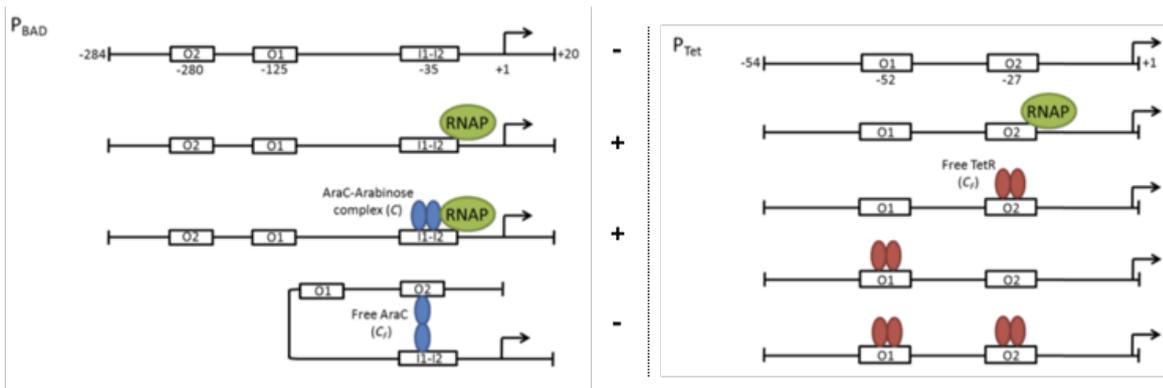


FIGURE 4.5: Inducible promoters used in this study. P_{BAD} is positively regulated by AraC in the presence of arabinose. P_{TET} is negatively regulated by its cognate protein TetR. Plus and minus symbols indicate the state of transcription for each of the different promoter states. Adapted from [223]

TABLE 4.1: Fluorescent proteins used as genetic circuits reporter.

Fluorescence	Multimer ($\lambda_{ex}/\lambda_{em}$, nm)	Monomer ($\lambda_{ex}/\lambda_{em}$, nm)
Green	eGFP (489/508)	sfGFP (485/510)
Red	mCherry (587/610)	mCherry2 (587/610)

4.2.2.4 Detecting quorum molecules

The recognition of AHL is carried out by LuxR and triggers subsequent activation of the P_{R-lux} promoter. We created plasmids with different promoters for the expression of LuxR in order to reduce the background signal observed in the native *V. fischeri* genetic context. We first separated P_{L-lux} and P_{R-lux} units of the *lux* promoter by a terminator to avoid P_{L-lux} leakiness[145]. Then we replaced TetR P_{L-lux} promoter for promoters from Anderson’s library to fluctuate available amounts of cofactors to trigger cellular response. In these different setups, we also tested the effect of removing degradation tags from LuxR in translational regulation. Therefore, we screened for combinations of adequate protein degradation time, promoter and RBS in order to obtain levels of LuxR that can sustain robust induction of P_{R-lux} .

In our circuits, activation of the P_{R-lux} promoter leads to the production of an output molecule (fluorescence). It proves that environmental AHL was detected at a permissive concentration and allowed binding of LuxR to P_{R-lux} and recruitment of the RNAP. We derived several options: circuits may simply be sensors, which provide a yes/no answer to the detection of a molecule, or they may be amplifiers that reinforce signal production[170]. We built both types of circuits and the latter was obtained by placing LuxI under the control of a low strength RBS, upstream of the fluorescent reporter controlled by P_{R-lux} (cf. Figure 4.2). As a final step for further characterisation studies, we also included a modified version of the P_{R-lux} promoter, where a cI protein binding site interferes with LuxR binding ($P_{RcI-lux}$ promoter)[48]. In this scenario, a sensor strain harbouring a plasmid driving the inducible expression of the cI protein would be able to actively repress the activation of the P_{R-lux} promoter. For the chaining of sender and receiver strains, this would guarantee a perfect control over signal production, detection and propagation (further details are given in Chapter 5).

4.2.2.5 Cotranscribed fluorophores

The construction of tunable genetic circuits in synthetic biology often involves the use of fluorescent reporter proteins. In our genetic devices design, we placed genes encoding reporter protein downstream of the LuxI coding sequence for sender plasmids and

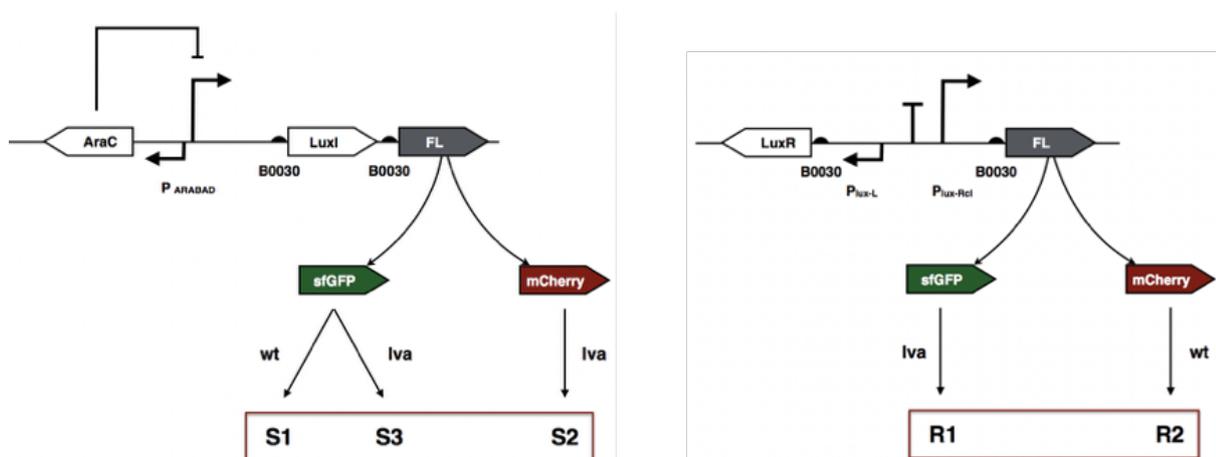


FIGURE 4.6: Genetic circuit variations in the use of fluorescent proteins. FL denotes the location of the fluorescent protein. Both sender and receiver circuits were propagated with monomeric fluorescent proteins (sfGFP and mCherry2). On the left, sender circuits were assembled with either wild-type (wt), ssrA-tagged sfGFP and ssrA-tagged mCherry2. On the right panel, receiver devices were assembled with either ssrA-tagged sfGFP or wild-type mCherry2. S1/S2/S3 and R1/R2 respectively represent different versions of the sender and receiver devices.

downstream of the $P_{RcI-lux}$ promoter in receiver devices. We compared a set of fluorescent proteins to obtain minimal noise/maximum signal emission in genetic circuits. Table 4.1 references the proteins we used for respective monomeric/multimeric fluorophores. Individual strain characterisation was mostly followed by the production of green fluorescence. Red light detection channels were characterised as an alternative for more complex joint systems (mixed populations of sender and receiver devices). Our optimised variations of fluorophores were based on the genetic circuits outlined in Figure 4.6.

In this study, we developed series of genetic variants aimed at sending and detecting/amplifying quorum signals. Optimisation of genetic circuit behaviour depends over a set of factors. Unless high-throughput studies are endeavoured, considerations should be taken to rationally restrict the number of combinatorial changes to be carried out in a single device. This is because as this number of changes increase, the number of controls required to study these changes goes up too, and this may add a significant amount of labour to obtain results that may not be concordant with one another. In this section, we have shown our model and variations of quorum devices. We assembled these constructs in cloning strains (*E. coli DH5 α*) and later expressed them in wild-type or tailored strains (BW25113). The next section describes the results that were obtained

for plasmids characterisation in *in vivo* studies.

4.3 Single cell and population scale *in vivo* characterisation

In this study, we built a series of quorum devices that were tailored to producing a quorum signal, detecting this signal and amplifying it. Here, we describe the main results that were observed from the construction of genetic circuits detailed in the previous section.

4.3.1 Production of quorum compound

4.3.1.1 P_{TET} and P_{BAD} promoters

As aforementioned, we tested two genetic constructs for the production of AHL in *E. coli*: the P_{TET} and P_{BAD} inducible systems[144]. Tailoring the expression of quorum signal was coordinated with the emission of green or red fluorescent molecules. Therefore, we optimised the inducible promoters expression for the highest production of target molecules by the detection of fluorescence. For the P_{TET} promoter, we tested a range of promoter strengths and only strong expression of TetR would allow minimal levels of background fluorescence in the OFF state of the device. With P_{BAD} , the pBAD33 plasmid backbone provided by itself minimal leakiness of fluorescence expression levels. Figure 4.7 shows different basal levels of background fluorescence observed in the absence of inducer in bacterial cultures. Under the microscope, although P_{BAD} would occasionally show one bright cell, the vast majority of bacteria kept a homogeneous and low level of background fluorescence. P_{TET} , however, displayed a more heterogeneous response in fluorescence levels, varying from cells with no fluorescence to cells with robust and significant signals. In genetic constructs with weaker promoters driving the expression of TetR, more background fluorescence was observed in all cells for the same heterogeneity within bacterial populations. Under uninduced conditions, we thus report a better signal/noise ratio via the use of the activable arabinose promoter.

To activate the production of quorum compound, cells were either provided anhydrotetracycline (aTc) or L-arabinose, respectively for P_{TET} and P_{BAD} systems. We tested a range of inducer concentrations and compared bacterial strains in maximal levels of

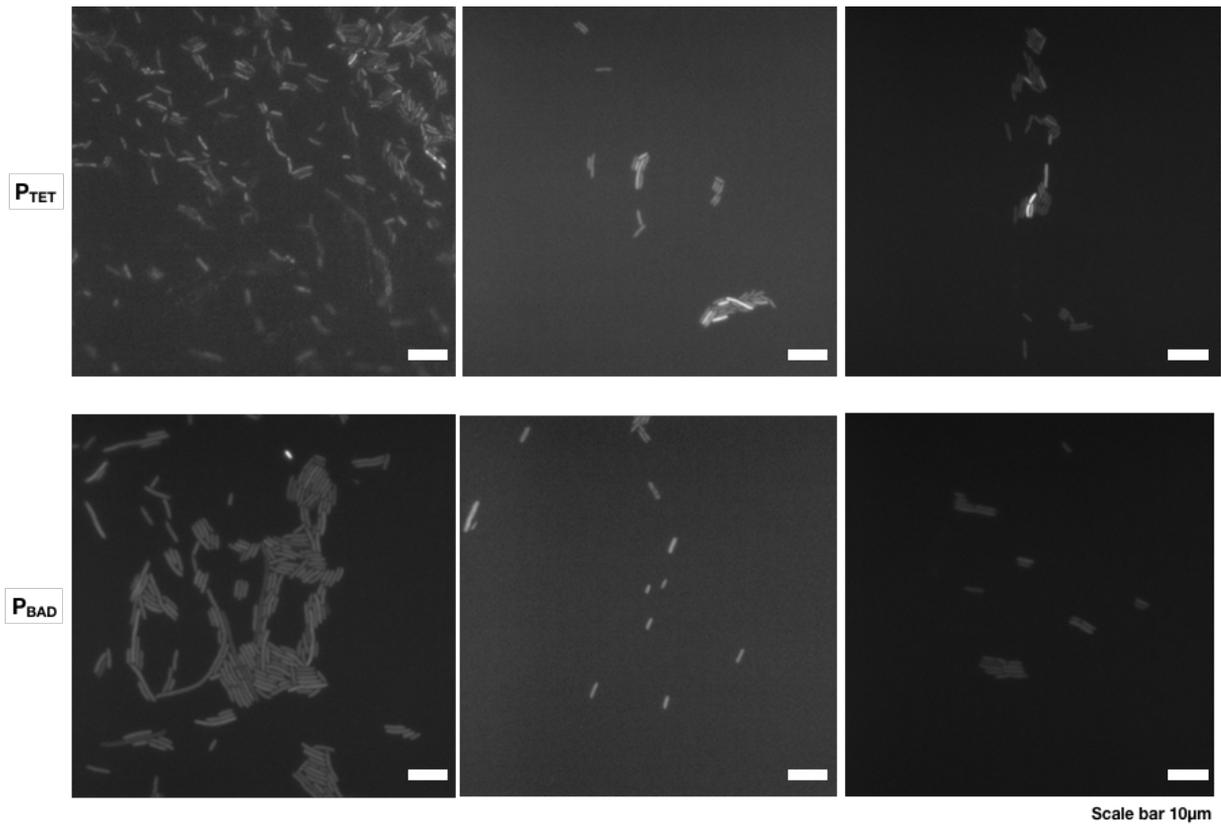


FIGURE 4.7: Fluorescence images of uninduced sender genetic circuits. The Tet repressible promoter basal fluorescence levels for strong TetR expression are shown in the top panel. Bottom panel shows the arabinose system behaviour in comparable conditions (absence of inducer) from the pBAD33 plasmid backbone. Both circuits are shown for the emission of green fluorescence with 300ms UV exposure.

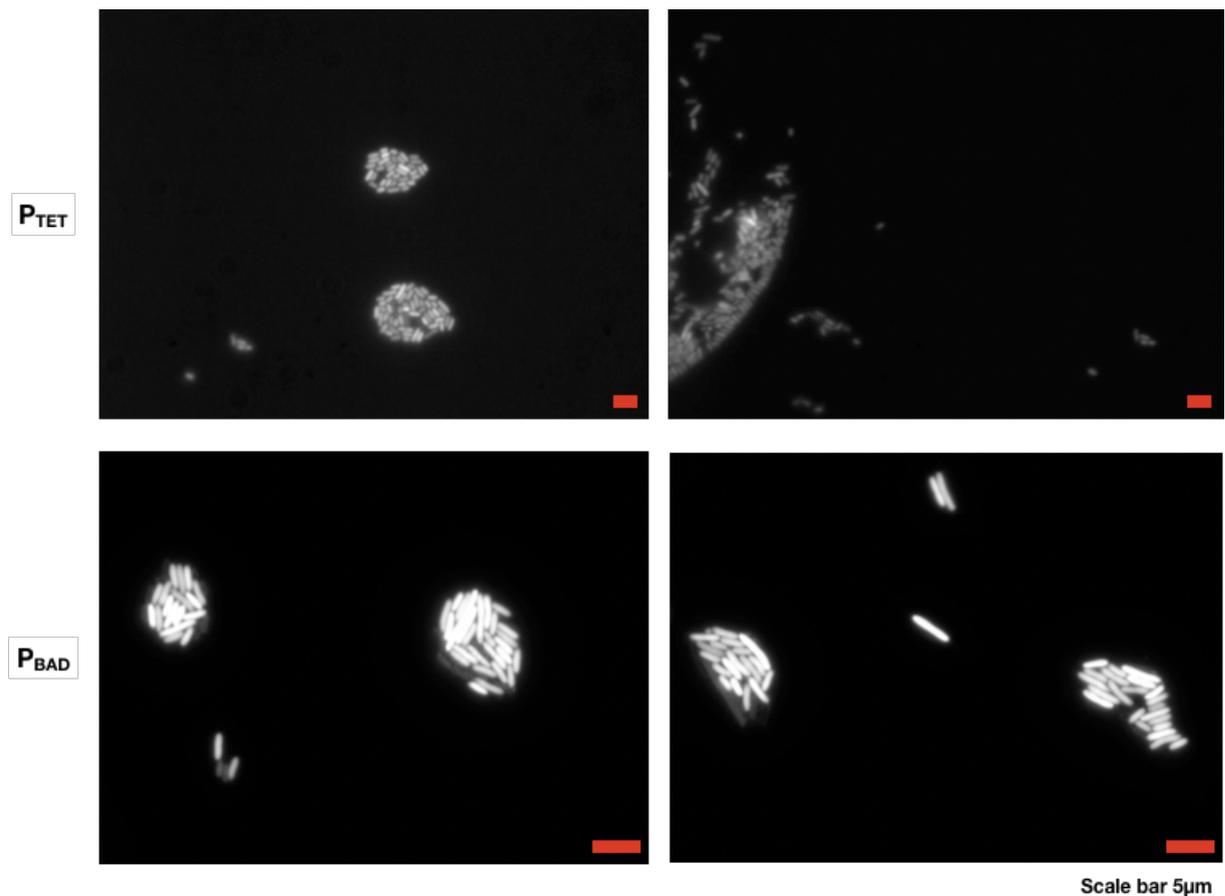


FIGURE 4.8: Fluorescence images of induced sender genetic circuits. The P_{TET} system is displayed at the top, while bottom panel shows the arabinose-regulated genetic circuits. Both systems were screened for fluorescence emission of the same green fluorophore with 300ms UV exposure.

fluorescence induction by microscopy. As shown in Figure 4.8, both promoters produced a homogeneous response to external induction. However, we constantly observed the P_{BAD} promoter to display higher fluorescence signals than maximal achievable fluorescence observed with P_{TET} . Therefore, to test whether these differences wouldn't be an imaging artefact, we explored higher level characterisation methods to get a population scale summary of bacterial behaviour.

Analysis via fluorescence microscopy allowed us to perform single-cell resolution imaging and to retrieve information about fluorescence levels fluctuation within cells. We used flow cytometry as a complementary approach to represent fluorescence behaviour changes at the population scale. Figure 4.9 displays the fluorescence profiles obtained in uninduced/induced conditions for both P_{TET} and P_{BAD} circuits. For a

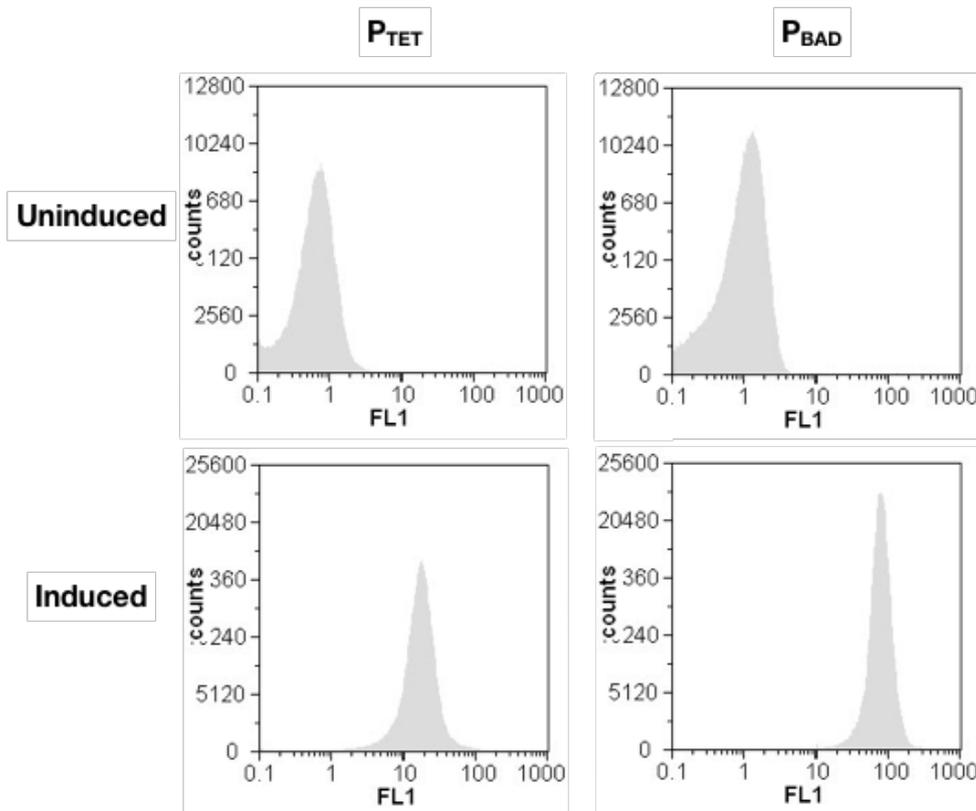


FIGURE 4.9: Sender genetic circuits were characterised via flow cytometry in uninduced/induced conditions in comparable cell counts. Top panel images display uninduced sample results, while bottom panel shows induced conditions. P_{TET} data is shown on the left and P_{BAD} on the right. In all graphs, the y-axis represents a number of cells screened per population and the x-axis displays fluorescence fold changes (FL1, blue laser).

comparable amount of uninduced or induced cells and identical equipment parameters, the P_{BAD} promoter always outperformed P_{TET} by producing a higher and more homogeneous fluorescence intensity signal. All-in-all, with the combination of single cell imaging and fluorescence cytometry, the arabinose inducible promoter appeared to provide more robust cellular behaviour than the repressible Tet system, for both minimising uninduced background noise and optimising fluorescence fold change in induced conditions.

4.3.1.2 *E. coli* expression strains

For expression studies, we used *E. coli* MG1655 and BW25113. Both strains correspond to common wild-type laboratory *E. coli*, but BW25113 was built from MG1655 and features the knockout of inducible rhamnose and arabinose regulatory pathways (Rha^-

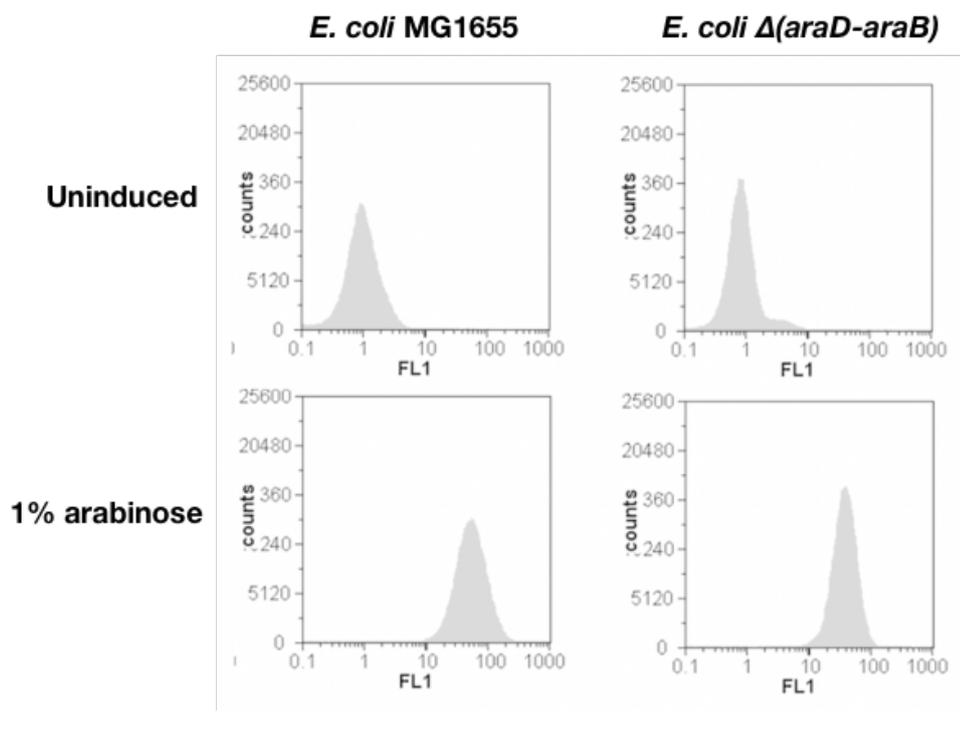


FIGURE 4.10: Flow cytometry results for induction of the sender quorum circuit propagated in two strains of *E. coli*: MG1655 (left) and BW25113 (right). Colored peaks represent the count of individual cells (y-axis) for certain fluorescence levels (FL1, x-axis). Samples were analysed in the absence of inducer (no arabinose, top panel) and in the presence of 1% arabinose (bottom panel) in similar cell counts and in the same equipment settings.

and Ara^-][85]. We used BW25113 for the use of P_{BAD} in sender genetic circuits in order to reduce the amount of cellular resources using arabinose. Figure 4.10 shows variable fluorescence responses that were observed for the same genetic circuits propagated in MG1655 and BW25113. Using flow cytometry, we observed a more homogeneous fluorescence (width of fluorescence peaks) for genetic constructs in BW25113. This is concordant with this specific strain genotype, since the inducer molecule is only used by the synthetic circuit, and not by any other cellular processes.

In a complementary assay, we tested an additional inducer concentration and observed cells growing in exponential phase under a fluorescence microscope for both *E. coli* MG1655 and BW25113. We took pictures, averaged data from 100 individual cells and show differential induction profiles in respective strains in Figure 4.11. Constructs propagated in *E. coli* MG1655 could reach slightly higher fluorescence levels but always showed bigger background fluorescence in uninduced conditions, thus showing a smaller fluorescence fold change compared to BW25113. In *E. coli* BW25113, cells

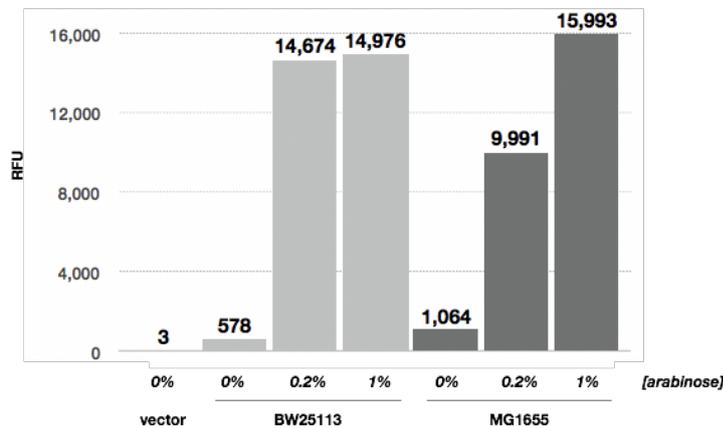


FIGURE 4.11: Single cell fluorescence intensities for different inducer concentrations in *E. coli* BW25113 and MG1655. Vector corresponds to an empty pBAD33 plasmid in uninduced conditions. Light grey bars correspond to BW25113 measurements and dark grey bars to MG1655. RFU represents relative fluorescence units averaged from 100 cells imaged at $OD_{600nm} = 0.4$.

displayed a full OFF/ON fluorescence response for various inducer levels. This is concordant with the availability of arabinose in wild-type vs. *Ara*⁻ conditions. At 0.2% arabinose in *E. coli* MG1655, cellular resources compete for inducer molecules: this is visible by a lower average fluorescence (9,991 RFU, 9x fold induction) compared to a 25x fold induction (14,674 RFU) in BW25113 for identical conditions. In order to obtain repeatable results in a range of conditions, we chose *E. coli* BW25113 as a model for the study of sender devices in the P_{BAD} context. This, in particular, was a better choice to show transfer curves hysteresis in experimental modelling studies discussed in the next section.

4.3.1.3 Continuous time-lapse of quorum circuit activation

The results described so far were focussed on discrete time points to analyse genetic circuits by fluorescence. However, we performed an extensive continuous *in vivo* characterisation of sender devices in automated plate readers (BMG Fluostar and Clariostar plate readers). This allowed us to gather kinetic information over the state of bacterial cells. In this setup, we studied cellular response as a function of variable fluorophores. In all cases, monomeric proteins were preferred as fluorescent reporters rather than

multimeric proteins. For the simple system of co-transcribing fluorophores with a target gene, monomeric proteins provide a more homogeneous fluorescent signal at single cell resolutions. As a rule of thumb in synthetic biology, green fluorescent proteins are often chosen instead of their red fluorescent homologues. In Figure 4.12 and 4.13, we show individual fluorescence induction profiles and their mean/median values for the emission of mCherry2 coupled with LuxI in panel A. Compared to measurements obtained with green fluorescence (Figure 4.13), red fluorescence detection provided much noisier data (Figure 4.12). Figures 4.12 and 4.13 show that AHL sender circuits had an earlier onset of induction by arabinose when the sfGFP was used as reporter. Fluorescence was generally detectable from 3h after induction for mCherry2 constructs compared to 2h for sfGFP devices. This can be explained by different maturation times of these fluorophores, which delay the onset of fluorescence for longer maturation time proteins. In terms of stability, green fluorescence again outperformed mCherry expression because the signal/noise ratio in mCherry constructs significantly varied over time. This was due to the harder calibration of plate readers for detection of red fluorescence, and explained by the kinetic properties of red fluorescent molecules and sensitivity of the equipment. Although signal quality varied between green and red fluorophores, synthetic circuits we created always displayed a quantifiable and significant OFF/ON behaviour for the activation of target molecules production. However, due to better kinetics observed with green fluorescence, sfGFP was our fluorophore of choice for further studies.

4.3.1.4 Testing for AHL production

We demonstrated in the previous sections how we could obtain a robust fluorescent signal coupled to the production of AHL. However, we did not present direct evidence for production of these chemical species. To verify functionality of sender circuits, we used an *E. coli* strain harboring the pSB401 plasmid for detection of LuxI catalytic activity. This plasmid encodes the native LuxR/LuxI operon from *V. fischeri* and is thus an AHL sensor. We first tested for the functionality of this control sensor strain by external AHL induction. As shown in Figure 4.14, in the absence of AHL, a 5 minute exposure for detection of bioluminescence from plates did not show any signal for the pSB401 plasmid alone, and a clear detectable luminescence in the presence of 1mM AHL. This proved that pSB401 could be used as a control for specific production of AHL.

In order to assess the capability of our sender genetic circuits to produce AHL, we cultivated bacterial cultures in the presence of inducer and fed their supernatant to

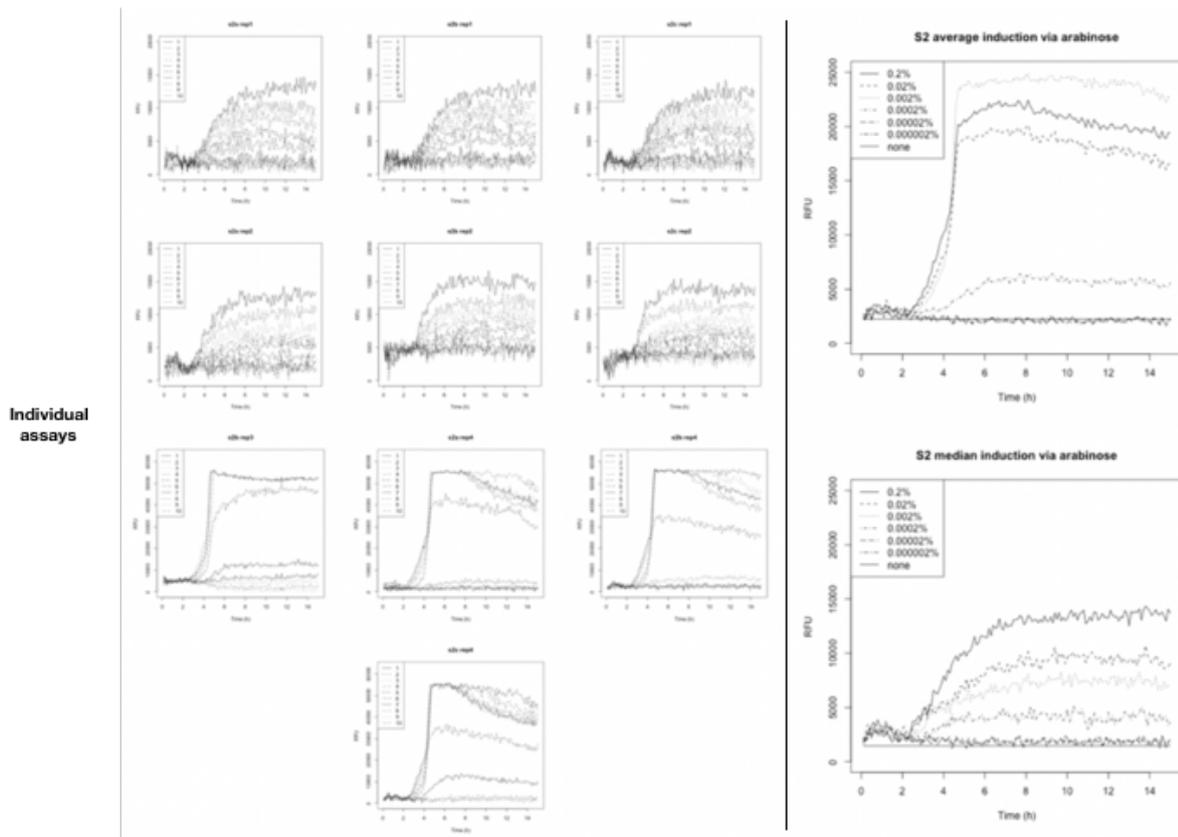


FIGURE 4.12: Plate reader assays results for the induction of sender genetic circuits by arabinose over red channel luorescence. In all graphs, the y-axis represents variation of fluorescence intensity in a range of inducer concentrations over time (x-axis). On the left, different assays fluorescence profiles are presented while their average are displayed on the right. The top right corner graph displays the mean of all fluorescence measurements while the bottom right corner graph shows median values.

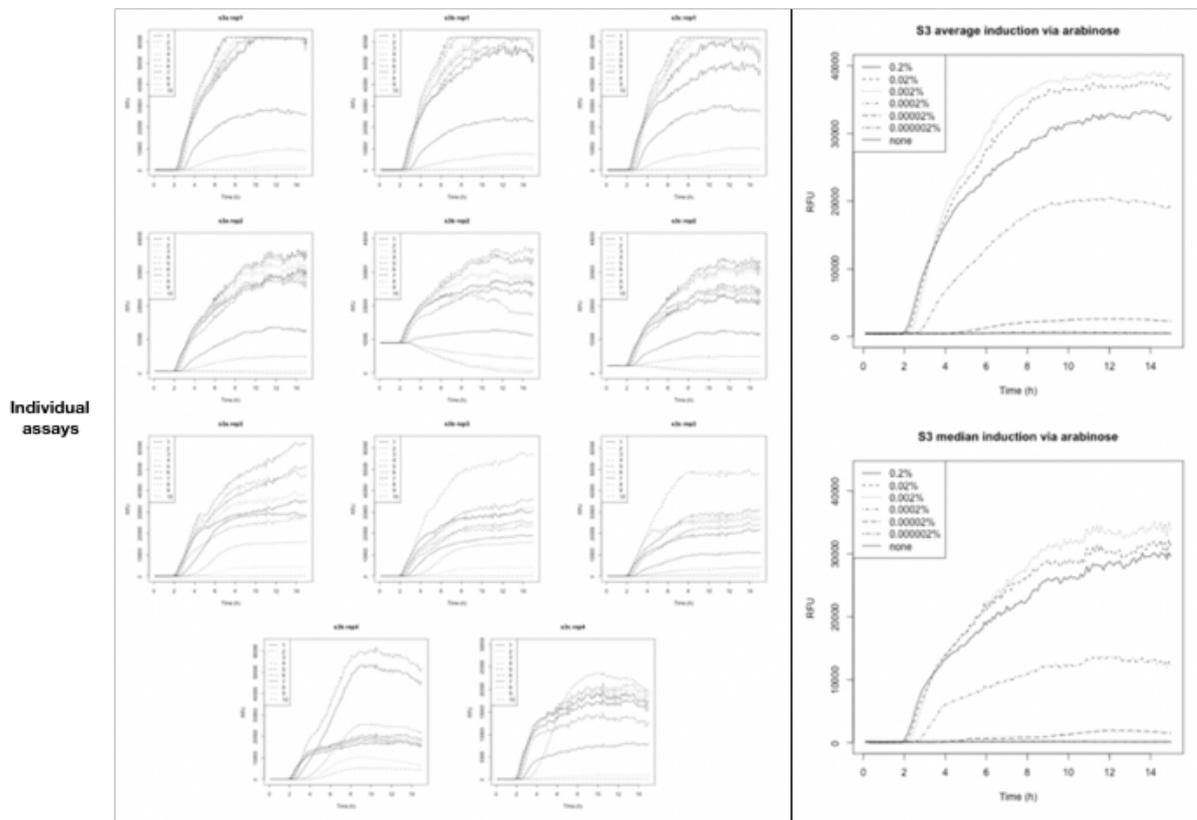


FIGURE 4.13: Plate reader assays results for the induction of sender genetic circuits by arabinose over green channel luorescence. In all graphs, the y-axis represents variation of fluorescence intensity in a range of inducer concentrations over time (x-axis). On the left, different assays fluorescence profiles are presented while their average are displayed on the right. The top right corner graph displays the mean of all fluorescence measurements while the bottom right corner graph shows median values.



FIGURE 4.14: Plasmid pSB401 used as control for specific production of AHL. No luminescence can be observed for the pSB401 sensor alone (left) and a robust signal can be detected in the presence of $1mM$ AHL (right).

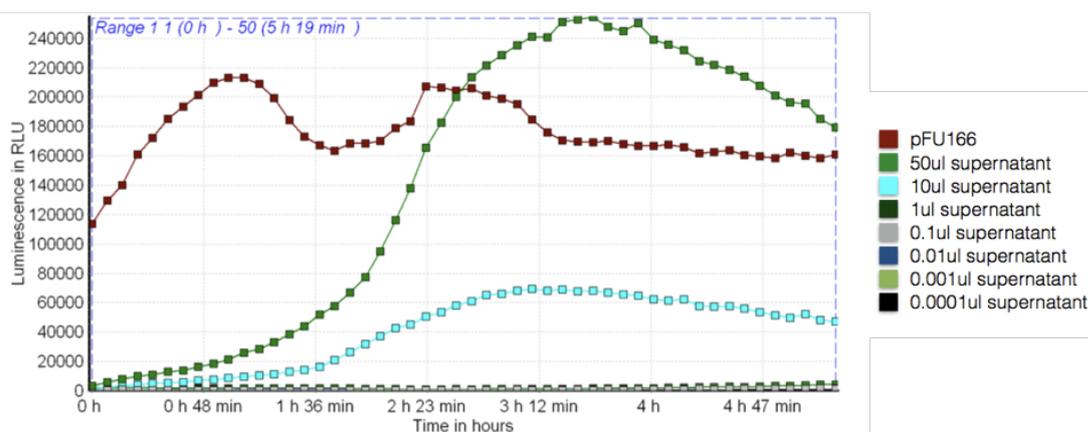


FIGURE 4.15: Bioluminescence induction profiles of pSB401 AHL sensor strains in the presence of variable levels of induced sender cells supernatant. The pFU166 plasmid is used as a positive control for constitutive luminescence expression. Luminescence values are plotted on the y-axis over time (x-axis).

pSB401 sensor cells. With plate reader assays for the detection of bioluminescence, we observed variable induction of sensor strains for different volumes of sender supernatant (Figure 4.15). These results were always compared to basal light levels emitted by the pFU166 plasmid constitutively expressing the LuxCDABEG operon[248]. For instance, results displayed in Figure 4.15 show that a 1:4000 dilution of the supernatant of an induced sender cells overnight culture ($50\mu\text{l}$) allowed to reach relatively high levels of bioluminescence, thereby production of AHL. For the different sender genetic circuits engineered in this study, we always verified AHL detection with pSB401 as a control besides appropriate fluorescence emission.

Here, we described the results obtained for a few variants of a genetic circuit library: sender devices. We demonstrated their ability to produce a target quorum molecule (AHL), and optimised their design characteristics based on *in vivo* measurements. For best signal emission towards sensor colonies, we used induction of LuxI via the P_{BAD} promoter coupled with emission of green fluorescence with sfGFP.

4.3.2 Quorum sensing response

As we engineered a sender genetic circuit, we also created receiver devices that detected the presence of quorum sensing compounds. Starting from the native *V. fischeri* LuxR/LuxI operon, we optimised the detection of AHL in *E. coli* using differential

LuxR regulation and circuits that would only detect, or detect and amplify received signals.

4.3.2.1 LuxR regulation

In *V. fischeri*, the P_{L-lux} promoter of the *lux* operon provides weak LuxR constitutive expression. However, the binding of LuxR to *lux* boxes along the operon may reverse the transcriptional state of this promoter, and it may exhibit leaky opposite direction activity. Since LuxR is expressed from the complementary DNA strand, we placed a double-terminator directly downstream of its promoter sequence to reduce background noise observed in genetic circuits. In continuous plate reader assays, we tested the induction of the P_{R-lux} promoter by detecting mCherry (replacing the native LuxICDABEG polycistronic mRNA). In Figure 4.16, we observe a strong variation between induction profiles of individual experiments and AHL detection could only be reported for $100nM+$ inducer concentrations. Although individual profiles showed variable response over different assays, we demonstrated that we could detect environmental levels of AHL in *E. coli* for this basic synthetic circuit using red fluorescence.

As we detailed for sender genetic circuits, different fluorophores were tested for the induction of bacterial response. We have shown previously that a significant signal was detected with red fluorescence, and attempted to improve this system by using the sfGFP instead of mCherry. As shown in Figure 4.17, we detected a stronger fluorescent response when using sfGFP as a reporter. Besides showing brighter signal, we also observed, for sfGFP constructs, a more homogeneous signal within bacterial populations. In induced conditions, mCherry constructs produced an average 85x induction varying from 50x - 500x, whereas sfGFP devices were very centered around a 200x induction level. These results show that - based on discrete time points - the sfGFP was brighter and provided a better signal robustness for the detection of AHL.

Reporter devices presented so far were based on the constitutive P_{L-lux} promoter to drive low expression of LuxR. To improve the AHL detection threshold we obtained with previous constructs, we upregulated the expression of LuxR and controlled its appropriate degradation via an *ssrA*-tag. We tested a range of promoters varying from low to very high constitutive expression and obtained best fluorescence induction profiles for relatively high LuxR levels. Figure 4.18 displays the differential response that was observed between the use of the native P_{L-lux} and *J23104* to drive LuxR expression. With higher levels of environmental LuxR proteins, we observed fluorescence induction for a few nanomolars of AHL, making the system 100 times more sensitive than with the use of the native *V. fischeri* promoter. The bacterial fluorescent response

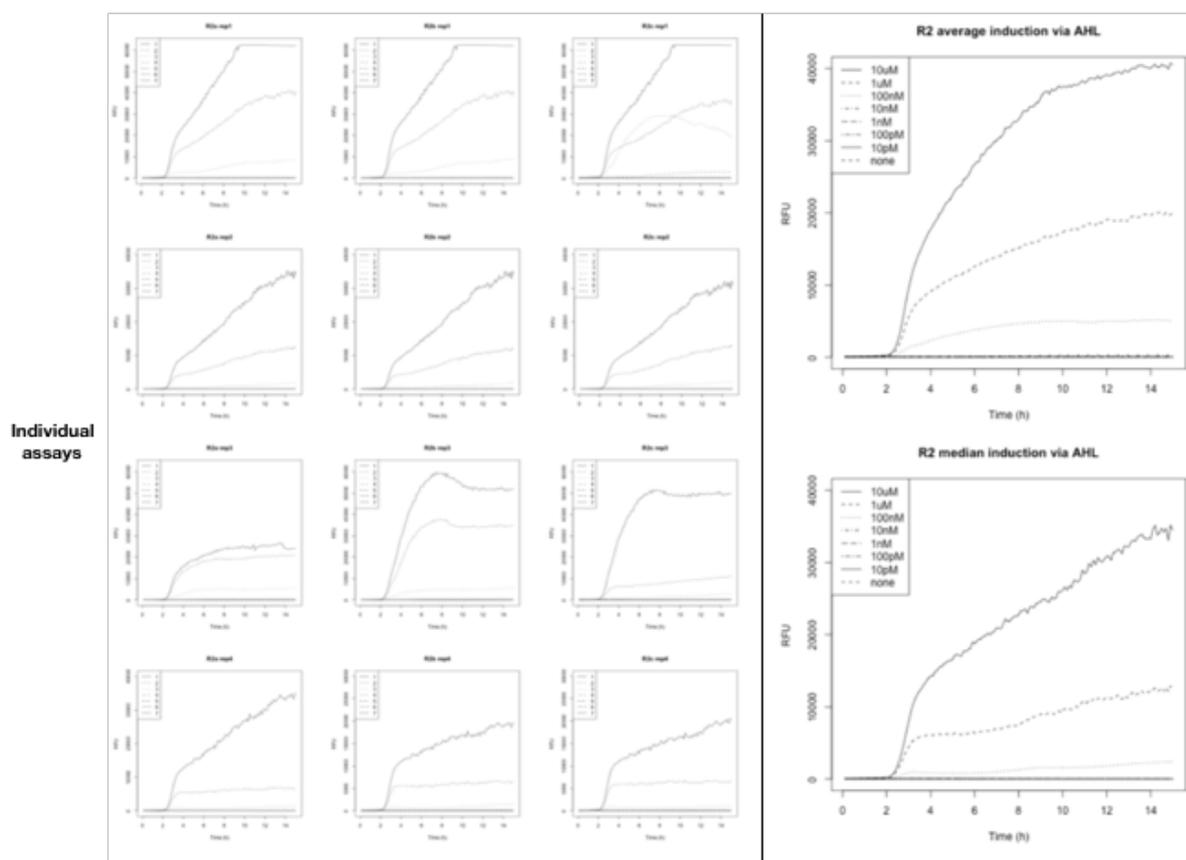


FIGURE 4.16: Plate reader assays results for the induction of receiver genetic circuits by AHL over red channel channel fluorescence. In all graphs, the y-axis represents variation of fluorescence intensity in a range of inducer concentrations over time (x-axis). On the left, fluorescence profiles of individual assays performed on different days are presented, while their average are displayed on the right. The top right corner graph displays the mean of all fluorescence measurements while the bottom right corner graph shows median values.

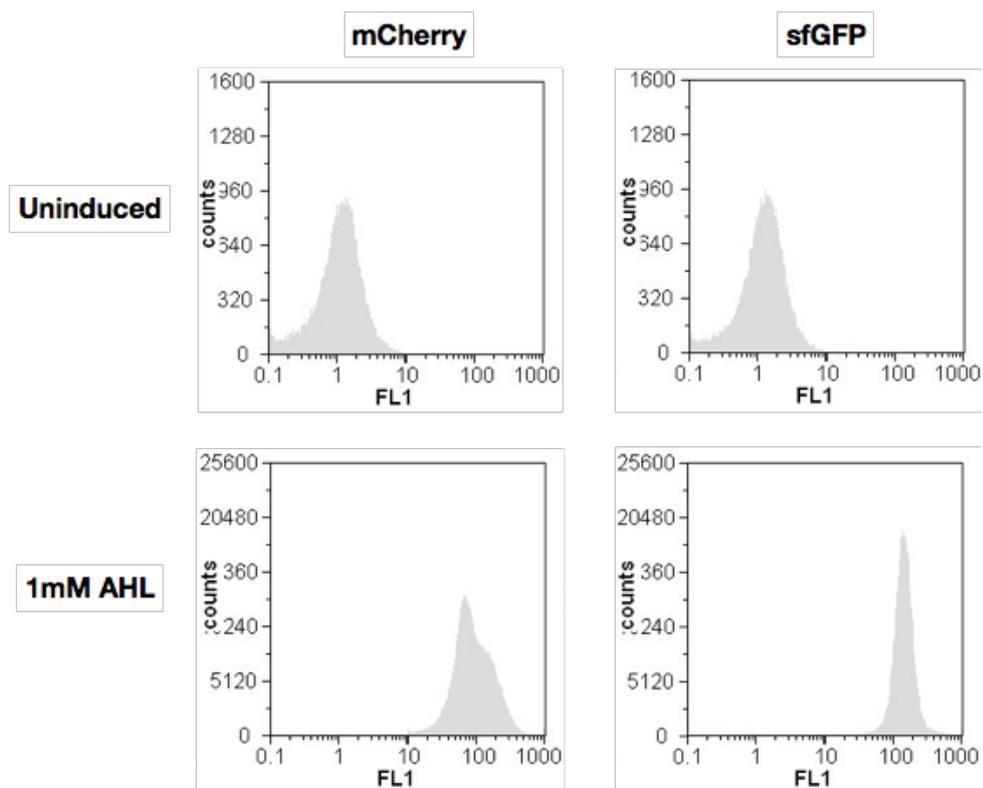


FIGURE 4.17: Flow cytometry results for the induction of receiver quorum circuits. Colored peaks represent the count of individual cells (y-axis) for certain fluorescence levels (FL1, x-axis). Samples were analysed in the absence of inducer (no AHL, top panel) and in the presence of 1mM AHL (bottom panel) in similar cell counts and system settings.

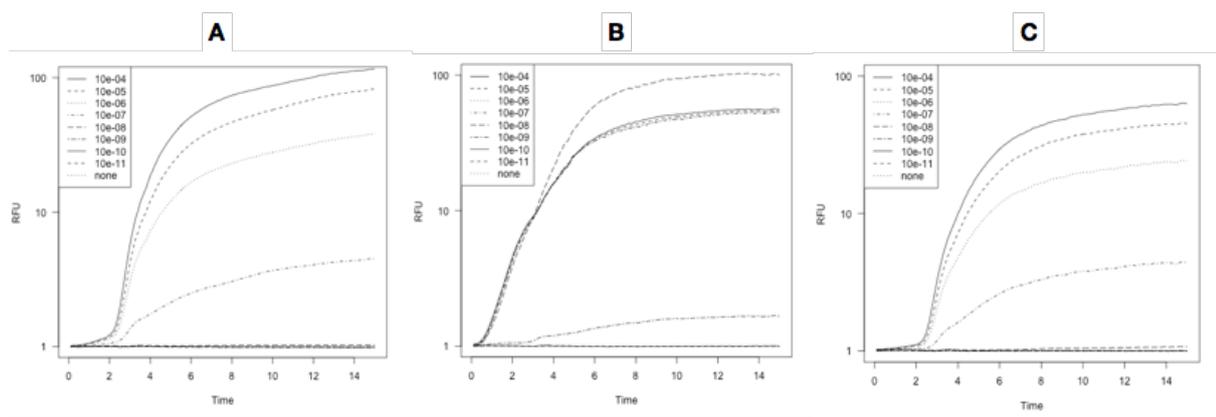


FIGURE 4.18: Effect of modifying LuxR regulation over receiver circuits fluorescence induction. For different AHL concentrations, both graphs show the induction of a green fluorescence signal (sfGFP, y-axis) over time (x-axis). In (A) and (C), LuxR is regulated via *V. fischeri* P_{L-lux} promoter and in (B) via the strong *E. coli* $J23104$ promoter. In (C), LuxI was cotranscribed with the fluorophore.

was reduced for stronger constitutive promoters, probably due to a LuxR crowding effect, where overnumerous LuxR molecules compete for AHL and binding to P_{R-lux} . For low synthetic LuxR expression, we also observed reduced signals due to a shortage of LuxR proteins compared to the available inducer. Driven by the expression of the $J23104$ promoter, we obtained robust and identical induction profiles for several inducer dilutions. One induction rate in particular produced a better response than all higher AHL concentrations, which is the closest to the theoretical P_{R-lux} promoter K_D . Moreover, while it would take about 2h for cells to peak in fluorescence in the native P_{L-lux} context, $J23104$ regulation allowed a quasi instant peak of fluorescence after starting time-lapse experiments. Here, we showed how adjusting cofactor promoter strength provided a greater control over the inducible P_{R-lux} in *E. coli*. Based on the optimisation of the response fluorophore and LuxR regulatory mechanism, we engineered receiver genetic circuits that coordinate a robust emission of fluorescence when sensing environmental AHL.

In Figure 4.18, we can observe a reduction of the overall fluorescence profile in the amplifier circuit (C) as opposed to the native *V. fischeri* context (A). One would naturally expect the opposite situation, where a rise in fluorescence levels should be created by an autoinduction loop. However, there are a few possible explanations: cells may be undertaking a certain metabolic load that, over a certain induction threshold, becomes toxic to bacteria, or the lux promoter may also display some properties such that higher inducer concentrations become less active at starting transcription (not uncommon for

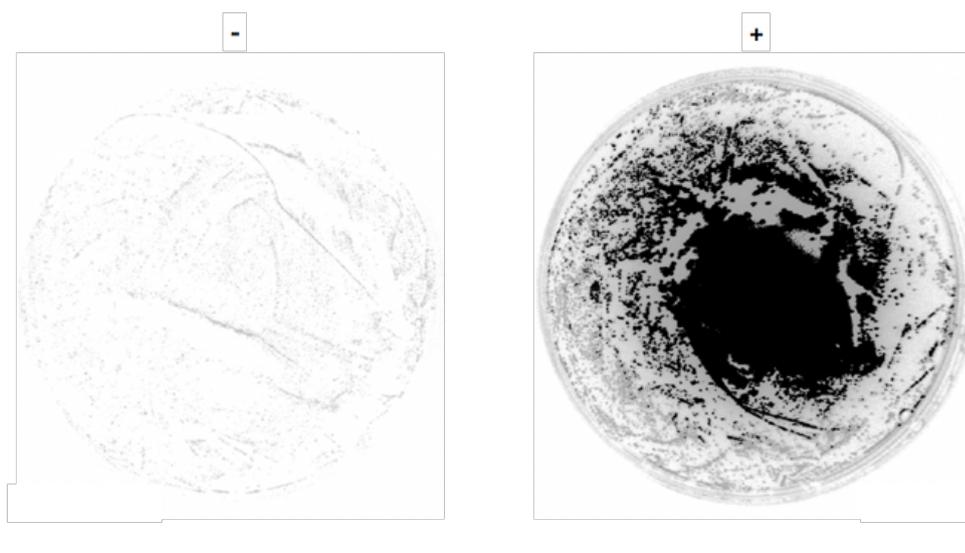


FIGURE 4.19: AHL sensor induction by sender genetic devices supernatant. On the left, receiver cultures were grown after plating $50\mu\text{l}$ of uninduced sender circuit supernatant. On the right, the sensor was imaged after induction by $50\mu\text{l}$ of induced sender circuit supernatant.

inducible promoters, as exemplified in Figure 4.13). Finally, it is also worth keeping in mind that the response detected via these experiments is solely based on fluorescence, which is only a reporter for the activity of the promoter and an approximation of both transcription and translation.

4.3.2.2 Induction via sender genetic circuits

As a control for the compatibility of our genetic constructs, we verified that AHL produced by sender devices could indeed induce fluorescence of sensor strains. In Figure 4.19, we detected the emission of fluorescence of receiver cells via a Typhoon scanner set in fluorescence settings. Supernatant from uninduced/induced sender cells was spread on plate over a high cellular density of sensor cells (necessary for the resolution of this technique). A thin basal level of fluorescence was observed for uninduced samples, while induction via sender cells supernatant produced a clear saturated signal over the bacterial lawn. This method had low resolution but allowed us to check the state of fluorescent cells from plates fairly easily.

For further validation of sender and receiver cells activity, we grouped their circuits together within the same bacterial strain. For these means, we used different fluorophores for each circuit (mCherry/sfGFP for sender/receiver cells respectively) and observed whether induction of AHL production from sender circuits would allow colocalised receiver plasmids to induce a subsequent response. As shown in Figure

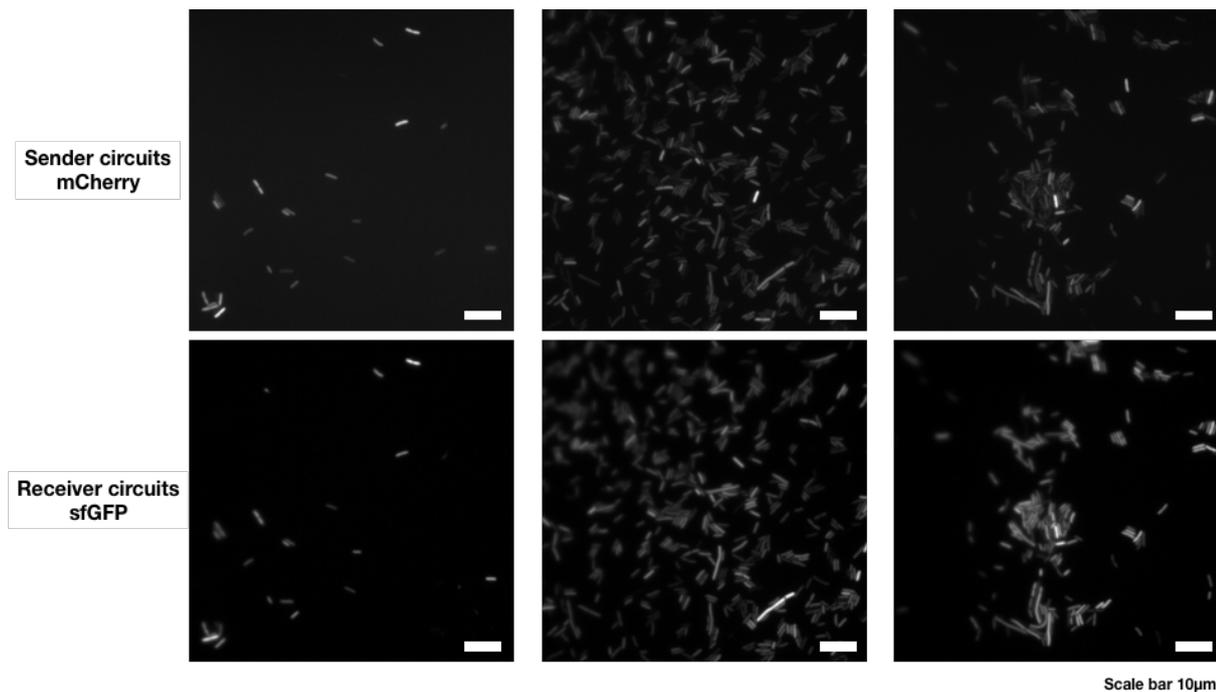


FIGURE 4.20: *E. coli* BW25113 was cotransformed with sender and receiver circuits. Cells were induced by arabinose to produce a red fluorescent signal and AHL, subsequently used by receiver circuits to report their activity by green fluorescence. Red channel fluorescence was imaged with 500ms exposure and green fluorescence at 200ms exposure.

4.20, fluorescence microscopy results always colocalised the expression of green fluorescence (functional receiver circuit) with cells expressing red fluorescence (activated sender device). Relative variations in sender induction was also observed to differentially induce receiver circuits activity within the same cell. Altogether, these controls allowed us to match sender and receiver circuits, and to verify that both circuits were behaving according to their expected behaviour.

4.3.2.3 Detector/amplifier circuits

For the building of receiver genetic devices, we undertook two approaches: (i) making sensor-only strains that detect and signal the presence of AHL, and (ii) amplifying any detected AHL signal. For the latter, we cotranscribed LuxI with the fluorescence emission unit in order to actively maintain bacterial induction response. In Figure 4.18 A and C, we compared genetic circuits for the same regulation of LuxR (P_{L-lux} expression) in both conditions. In these settings, coupling of LuxI downstream of the P_{R-lux} promoter mimics the regulatory situation as it happens in *V. fischeri*. With the

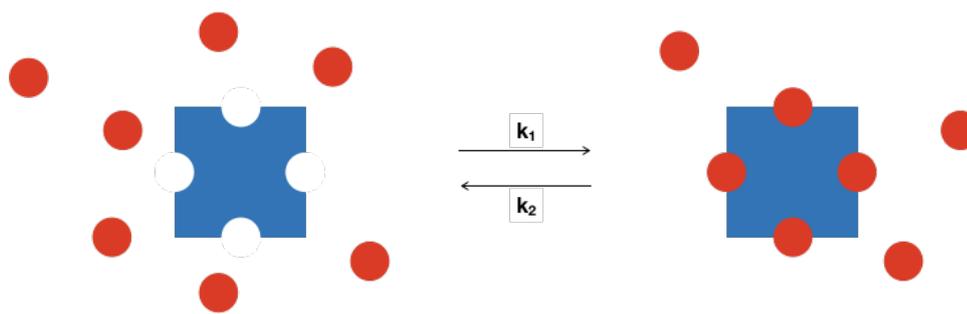


FIGURE 4.21: Level of abstraction used to model genetic circuits. Ligands (red circles) can bind a compound (blue square) on specific binding sites. For modelling studies, we abstracted a Hill formalism by considering two states of biological complexes: bound and unbound, dependent over constants k_1 and k_2 .

expression of this additional protein, genetic circuits generally showed a lower maximum induction level but identical or higher induction for smaller AHL concentration samples. This is the logical effect created by setting up a feed-forward loop in AHL sensor devices (cf. Figure 4.2). Therefore, we built two kinds of receiver circuits that may be used in different biological contexts for intercellular communication between bacteria. In the next section, we discuss modelling studies that helped us to draw these conclusions.

4.4 Mathematical modelling

4.4.1 Non-linearity in biology

In cells, molecular interactions have a cooperative effect: the coordinated action of representative chemical species over a substrate has a larger effect than the sum of these species only. This process is called non-linearity[171]. In biology, enzymes can have multiple sites where ligands can bind, and full enzymatic activity can only be reached when all these allosteric centers are occupied. In theory, one could model every individual binding state and characterise this systems intermediate behavioural changes. However, we do not exert enough control over genetic devices to be able to measure each condition precisely. Therefore, it is often easier to abstract biological complexes as an all-in/all-out situation, as described in Figure 4.21. Here, we adopted a Hill formalism and chose to group all "unknown" biochemical reactions (e.g. transcription/translation intermediate steps) under non-linear terms of mathematical equations.

4.4.2 Modelling via the Hill equation

For the study of our genetic devices, we were interested in the production of specific output molecules (fluorophores). We abstracted circuits transcriptional and translational regulation and derived quantities of signal (S) over time by the equation:

$$\frac{dS}{dt} = \frac{K_i^n}{K_i^n + L^n}$$

where L^n is a ligand concentration and K_i the inhibition constant between the molecule of interest and its ligand. In the Hill formalism, n represents the non-linearity term that explains underlying molecular processes abstracted in the model. This equation is often used to model the behaviour of promoters, where the RNAP competes with additional factors to start transcription. In this kind of kinetic processes, we model by the equation $C + L \rightleftharpoons CL$ the coupling of a compound C (RNAP) to its binding molecule L (ligand). Compounds thereby have two states: bound and unbound. At any time, the amount of available compound is equal to the ratio between the unbound compound and the overall number of chemical species it is involved with:

$$C_{available} = \frac{C}{C + CL}$$

The amount of bound compound CL is subject to fluctuate over time, given different ligand concentrations and its reversible state into individual subunits. It is thus common to use a dissociation constant that represents the rate at which CL complexes break down into separate chemical complexes (as it happens in cells). Therefore, in a steady-state system, we can transform the previous equation into the following:

$$\frac{C}{C + CL} = \frac{C}{C + \frac{C \cdot L}{K_D}} = \frac{K_D}{K_D + L}$$

where K_D is the dissociation constant between polymerase and inhibitor complexes (AraC or TetR cofactors). The term $\frac{K_D}{K_D + L}$ is the simplest case of Hill equation with $n = 1$ and $K_i = K_D$. Here, only one molecule of ligand can bind to the polymerase and affect its state at a given rate. However, as we described in previous sections, there are several cofactors that affect the production of a target molecule in biological systems, not a transcription rate only. This is represented by an added complexity for molecules production with a Hill coefficient (or cooperativity) index $n > 1$. This allows us to

revise a basic biochemical reaction description to:



and implies changing the subsequent equation as:

$$\frac{C}{C + CL_n} = \frac{C}{C + \frac{C^n L^n}{K_D}} = \frac{K_D}{K_D + L^n}$$

To comply with $\frac{dS}{dt} = \frac{K_i^n}{K_i^n + L^n}$, the condition $K_i = \sqrt[n]{K_D}$ must be observed. Related to the dissociation constant of a certain ligands, K_i is also referred to as the inhibition constant that describes half of a target enzyme activity. Such Hill formalism is commonly used for the modelling of biological species. Overall, it makes it easy to derive precise chemical parameters and to observe any strong variations between experimental data and simulated models.

4.4.3 Data normalisation

We generated time-continuous data with plate reader assays to follow the activation of fluorescent molecules. As shown in the previous section, this resulted in sigmoid curves representing the change in bacterial behaviour and fluorescence profiles (Figures 4.12, 4.13 and 4.16). For each analysed genetic construct, we gathered data from at least 3 individual plate reader experiments, where bacteria were followed in triplicates over a range from 8 to 11 inducer concentrations simultaneously. First, all samples were checked for normal growth profile, information provided by optical density measurements taken every 6 minutes. Any persistent growth defects over replication of the experiment would be sign of poor biological fitness due to the genetic circuit. If all growth profiles appeared as standard curves (cf. Figure 3.11), fluorescence data was then normalised by subtracting the blank medium fluorescence from the observed signal (background noise) and normalised by uninduced samples to obtain fold inductions. By applying this protocol over a set of biological replicates, we averaged experimental data and used non-linear regression with our Hill model to generate model fitting curves. These curves, by their proximity to experimental data, defined the goodness of the model. Directly inferred from fitting curves, we extracted residuals (closeness to experimental data) and plotted these over time to visually and numerically assess a model goodness.

4.4.4 Quorum sender devices modelling

When the model cannot be fit to experimental data, there are two potential solutions: revising the theoretical model or the genetic device. The most common approach is to modify genetic sequences, since theoretical models are meant to translate a "perfect" scenario of biological behaviour, thus the objective to reach. Within our library of genetic devices, our first constructs provided bad fit of the model to experimental data. This was shown in the development of the first version of our sender genetic device in Figure 4.22 S2L. This circuit encoded LuxI and mCherry under the control of the pBAD promoter on a low copy-number plasmid. As we demonstrated the optimisation of biological circuits with the use of alternative fluorophores and regulatory mechanisms, modelling results significantly improved by reducing distance from the theoretical model to experimental results to a minimum (low sum of residuals). As seen in Figure 4.22, the difference between red and green fluorescence (S2/S3) provided an attractive target to improve signal quality. For red fluorescence studies with mCherry, the signal:noise ratio was much higher than with green fluorescence proteins. Therefore, this created constant variations in signal levels and made the modelling process more challenging. Green fluorescent sfGFP exhibited very good kinetics (up to 500x fold induction compared to 8x for mCherry) and thus provided a better fit to the model. Expressing circuits from higher copy-number plasmids was also shown to optimise the genetic models behaviour: providing higher expression levels allowed reduction in the signal:noise ratio. Therefore, based on the knowledge gained from engineering quorum sender units, we optimised sensor strains with sfGFP as a biological reporter.

4.4.5 Sensor strains modelling

For the modelling of AHL sensor strains, two types of genetic circuits were developed: basic AHL detectors and AHL amplifiers. Both were described by the same model and are summarised in Figure 4.23. Four AHL receiver constructs were compared based on two different regulatory mechanisms of LuxR: native (P_{L-lux} controlled) or optimised conditions (strong promoter/degradation tag). These are respectively displayed in Figure 4.23 left-/right-hand side panels. For all of these constructs, the model provided a very good fit to the experimental data and allowed us to assess which versions performed better than others. However, the same genetic devices set on low copy number backbones provided much noisier signals. We thus decided to stick with higher copy number plasmids as long as there was no cytotoxicity problems. In the next section,

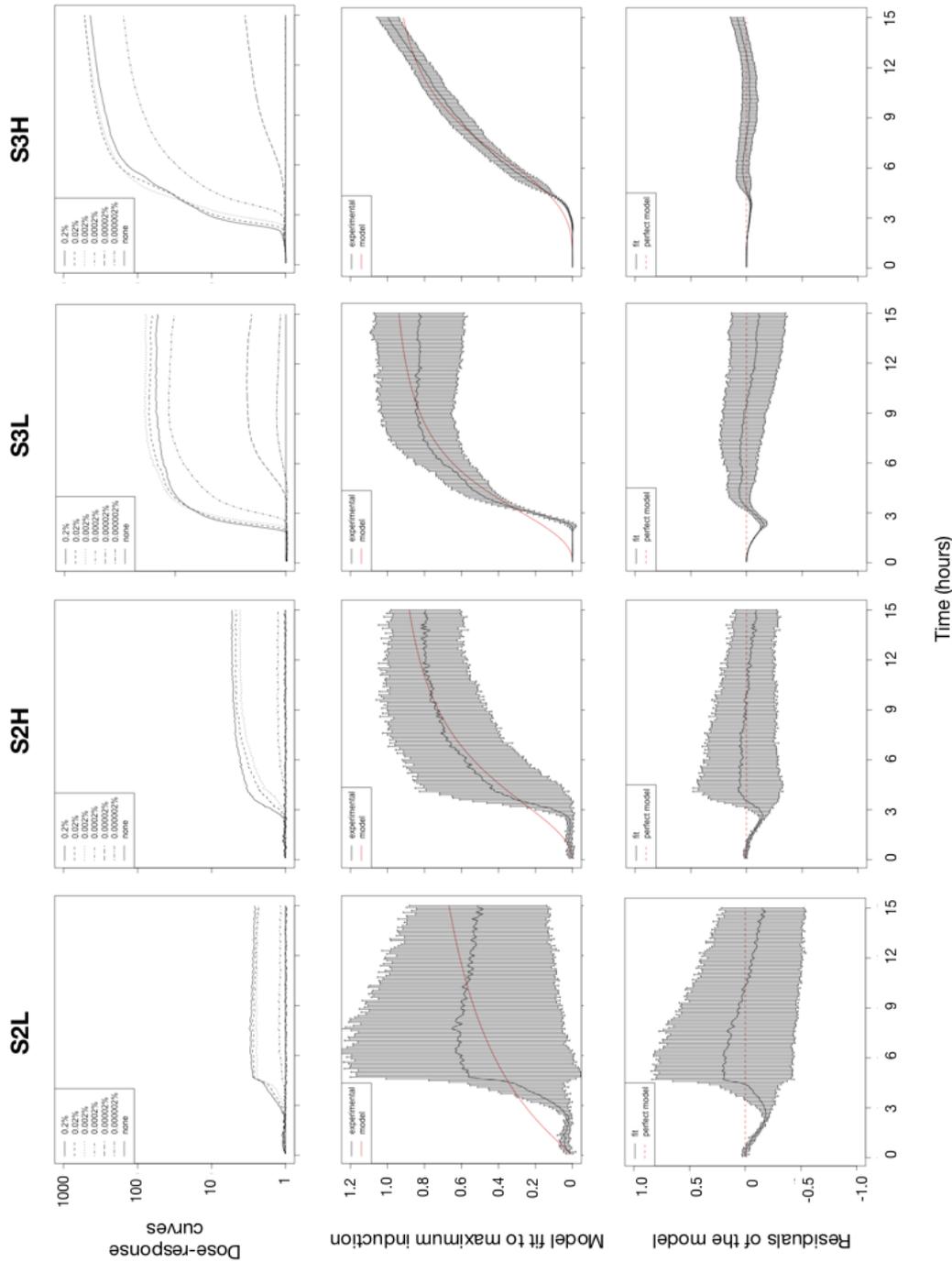


FIGURE 4.22: Sender genetic circuits modelling results. S2 and S3 stand for sender devices carrying mCherry or sfGFP respectively. The letter L describes the circuit on a low plasmid copy number and H on a high copy number backbone. Top row displays averaged fluorescence data (y-axis) over time (x-axis) for the four different circuits S2L, S2H, S3L and S3H. For each of them, the middle row describes a specific (maximum) induction profile (black line) and its fit to the mathematical model (red line). Bottom row shows residuals of the model (distance between model and data) over time. Perfect fit is shown as a dotted horizontal red line. Error bars show standard deviation across samples.

TABLE 4.2: Mathematical model fitting to individual sender and receiver strains fluorescence induction profiles. Strain represents each individual construct that was tested: S2L corresponds to the mCherry reporter protein on a low copy-number plasmid. All other constructs show relative fluorescence units (RFU) over time for sfGFP reporters. S3L is the sender device on a low copy-number plasmid while S3H has the same circuits on a high-copy number plasmid backbone. R75 and R114 are basic receiver devices with weak or strong constitutive expression of *LuxR*. R104 and R120 respectively represent R75 and R114 with *LuxI* inserted on the reporter mRNA, making it an AHL amplifier. Fluorescence fit (n) shows how well a Hill model fits average induction by its inducer. Transfer (n) is the cooperativity index of a Hill function that was used to characterise genetic circuit transfer functions. Significance values were obtained from nonlinear regression: * = 10^{-01} , ** < 10^{-02} , *** < 10^{-03} .

Strain	Reporter	Copy-number	Fluo. fit (n)*	Transfer (n)*	K_D
S2L	P_{BAD} -LuxI-mCherry	Low	2.383***	14.419***	7.034×10^{-4}
S2H	P_{BAD} -LuxI-mCherry	High	3.377***	15.296***	7.034×10^{-4}
S3L	P_{BAD} -LuxI-sfGFP	Low	2.688***	7.819***	6.161×10^{-5}
S3H	P_{BAD} -LuxI-sfGFP	High	3.498***	10.225***	7.369×10^{-5}
R75	P_L -luxR/ P_R -sfGFP	High	3.258***	12.765***	5.228×10^{-7}
R104	P_L -luxR/ P_R -LuxI-sfGFP	High	3.695***	12.067***	4.691×10^{-7}
R114	$J23104$ -luxR/ P_R -sfGFP	High	4.088***	16.780**	6.436×10^{-9}
R120	$J23104$ -luxR/ P_R -LuxI-sfGFP	High	3.820***	13.189***	1.093×10^{-7}

we provide an overview of the aforementioned genetic circuits, given modelling results that helped to optimise genetic designs.

4.4.6 Transfer function modelling

Given dose-response curves, it is possible to derive system transfer functions that characterise their transition from inactive to active states (or the opposite). For the modelling of biological complexes, we characterised transfer functions for the production of AHL in sender genetic circuits, and for detection of quorum molecule in sensor devices. Table 4.2 summarises modelling characteristics of the different genetic devices presented thus far. The transfer function was based upon discrete induction points of continuous experimental data (after 10h of culture), and was used to deduct respective dissociation constants (K_D). The same Hill function as the one used in previous subsections was employed to obtain transfer functions. For sender devices, the change from red to green fluorescence protein improved the arabinose detection threshold 10 fold. As for sensor circuits, strong expression/regulation of LuxR was shown to be the best combination to obtain lower AHL detection thresholds. In Figures 4.24 and 4.25, we show the transfer function curves that were obtained for a set of sender/receiver genetic circuits respectively. A better fit was generally observed for higher copy number

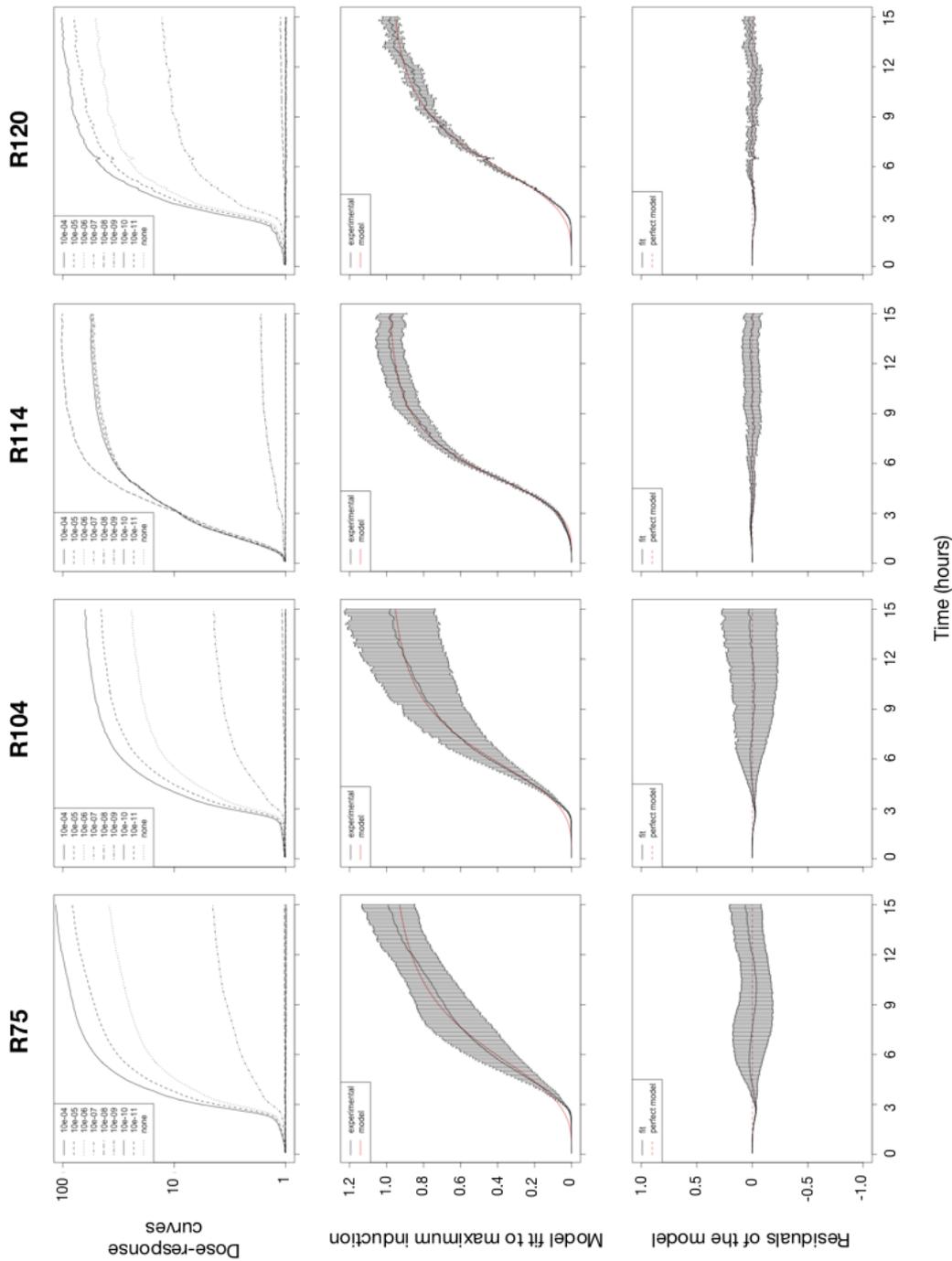


FIGURE 4.23: Receiver genetic circuits modelling results. R75 and R104 stand for sensor devices under the control of the *lux* promoter. R114 and R120 represent the same circuits with LuxR under the control of the strong J23104 promoter. R75 and R114 are AHL sensor only strains, while R104 and R120 amplify detected signals with LuxL. Top row displays averaged fluorescence data (y-axis) over time (x-axis) for the four different circuits. For each of the genetic circuits, the middle row describes a specific (maximum) induction profile (black line) and its fit to the mathematical model (red line). Bottom row shows residuals of the model (distance between model and data) over time. Perfect fit is shown as a dotted horizontal red line. Error bars show standard deviation across samples.

plasmids and modelling fluorescence induction profiles allowed us to revise genetic designs to gain control over expected and experimental behaviour.

4.4.7 Optimised bioengineered circuits

As shown in previous sections, we optimised the design of genetic circuits in order to show a better fit to our theoretical model. For the different AHL sender/sensor strains, we show the average fold induction they displayed *in vivo* under induced conditions in Figure 4.26. In the sender devices, we demonstrated that sfGFP was a better reporter protein than mCherry (S2L/S2H vs. S3L/S3H). We also constantly observed a better signal in high copy number settings for the same fluorophore (S2L/S2H and S3L/S3H). As shown in Figure 4.22, there was a significant variation in induction profiles from experiment to experiment, and changing fluorophore appeared to be the most straightforward solution to tackle this problem. The issue of getting poor signal/noise ratio with red fluorescence partly rises from the fact that the plate reader equipment used in this thesis did not have great capabilities for handling red fluorescence emission/detection: this was performed over a spectrum that did not fully overlap with the one of the mCherry, thereby producing noisier output data. For sensor circuits, amplification of AHL response always slightly decreased the observed signal, likely due to the metabolic load applied by coexpressing LuxI along sfGFP. Lower copy number versions of these circuits are not shown on this graph but averaged maximum fold induction profiles of 10x, comparable to the mCherry sender model on low copy plasmid. Overall, we showed for the optimised AHL sender/detector/amplifier circuits an average cellular response characterised by a ≥ 100 fold induction of fluorescent signals.

All-in-all, we constructed a series of genetic devices, characterised them *in vivo* and simulated their fitness with the theoretical model. A cyclic process of optimisation was carried out to improve the fit between model and experimental data, and our two best circuit candidates for AHL production/detection were respectively summarised in Figure 4.27. The most efficient AHL production device was based on the arabinose inducible promoter with coexpression of LuxI and sfGFP, both with degradation tags. This circuit displayed a dissociation constant of $K_D = 49\text{nM}$ arabinose. The best receiver plasmid was a simple detector (no amplification of AHL) with strong expression of ssrA-tagged LuxR controlling sfGFP production upon binding of AHL. This device allowed us to reach a detection threshold $K_D = 6.1\text{nM}$ AHL, about one or two orders of magnitude below most receiver versions. In the final section of this chapter, we

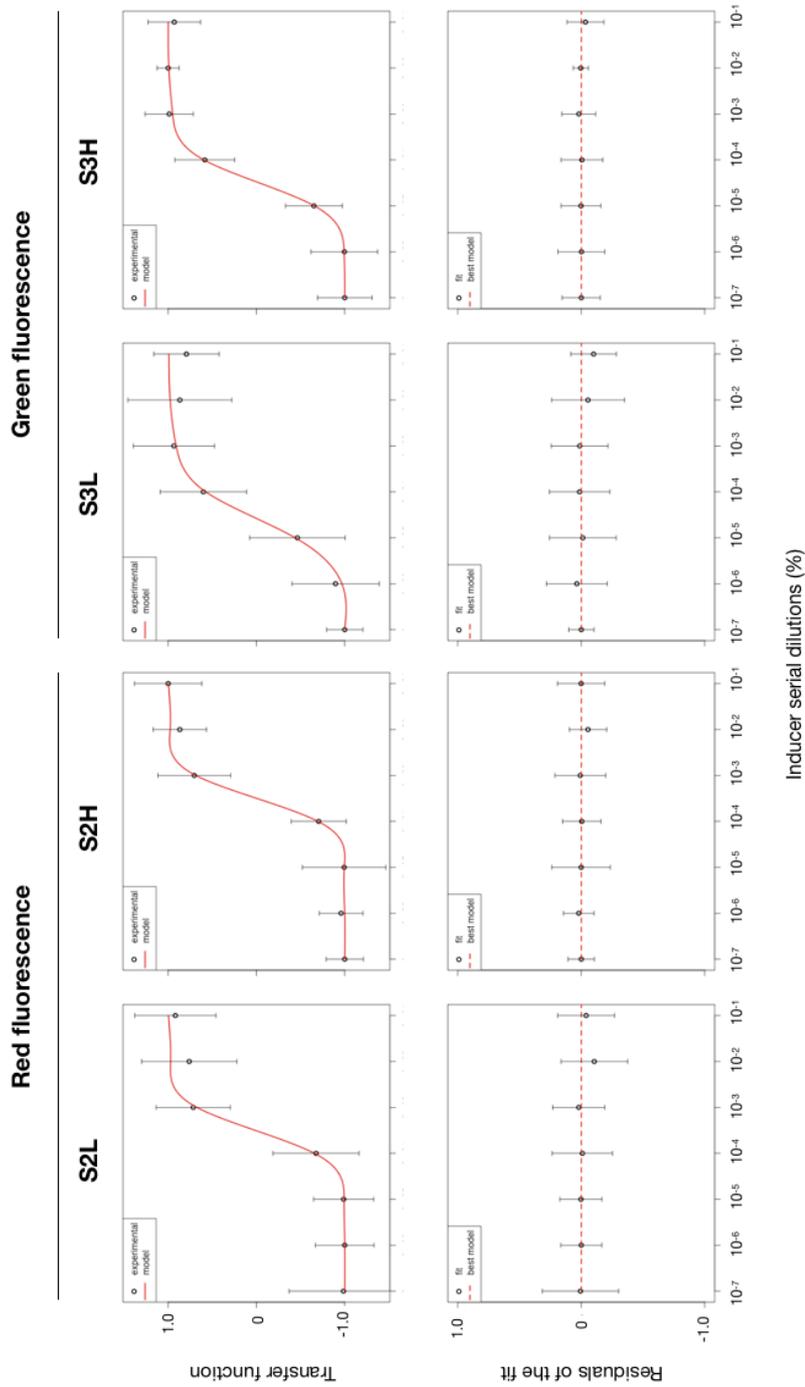


FIGURE 4.24: Transfer functions and regression residuals for different sender genetic circuits. Discrete time points were collected from dose-response curves at $t=12h$, averaged and fitted to a Hill model. Error bars show standard deviation across samples.

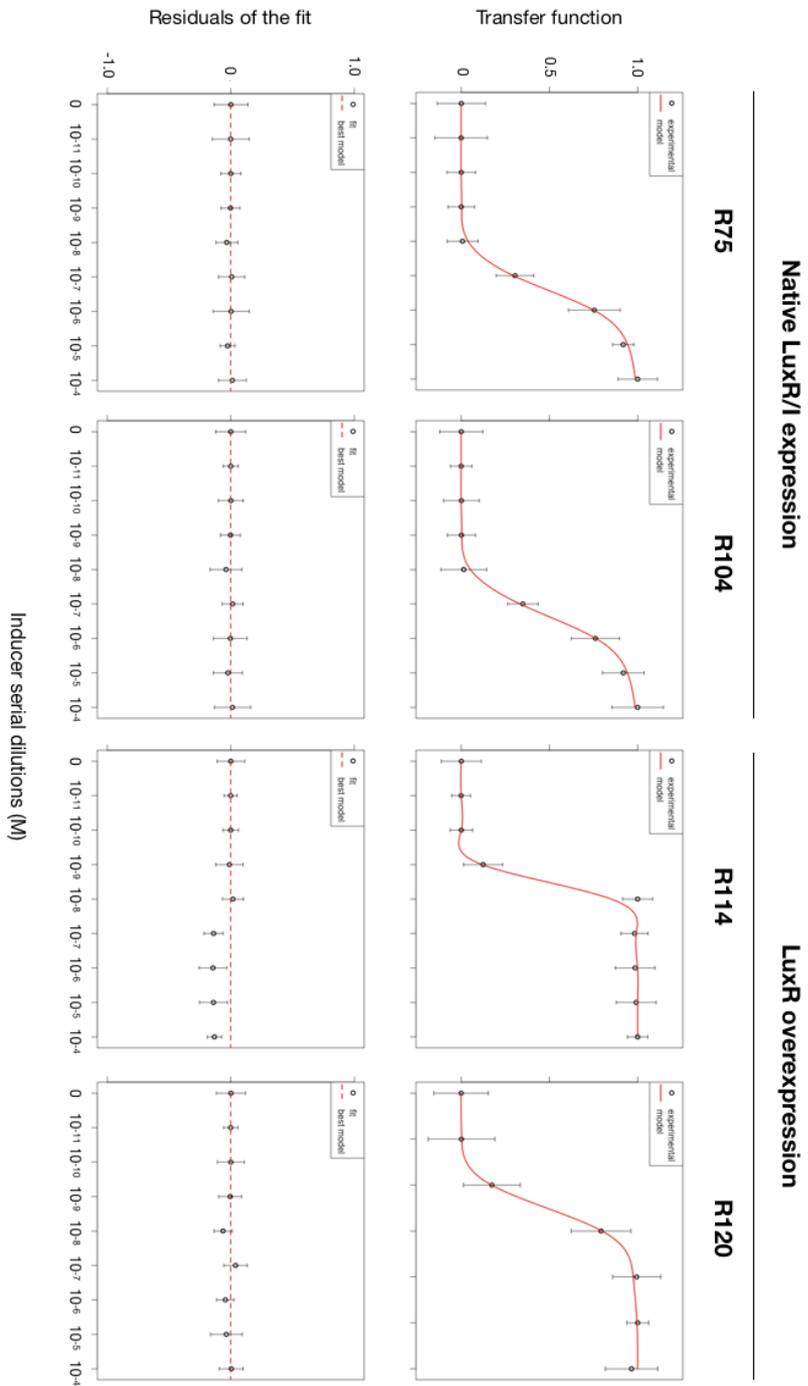


FIGURE 4.25: Transfer functions and regression residuals for different receiver genetic circuits. Discrete time points were collected from dose-response curves at $t=12\text{h}$, averaged and fitted to a Hill model. Error bars show standard deviation across samples.

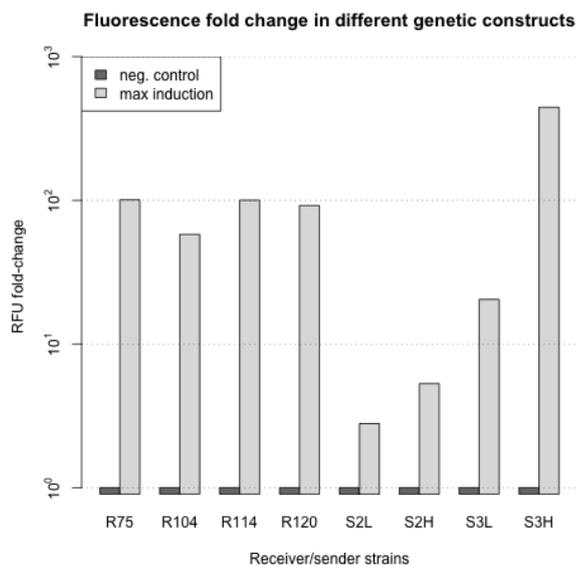


FIGURE 4.26: Fluorescence fold change in different genetic constructs: AHL sensor (R75/R104/R114/R120) and sender (S2L/S2H/S3L/S3H) strains. Data is normalised by uninduced samples and corrected for growth medium background. The y-axis shows relative fluorescent units (RFU) fold changes in fluorescence.

provide an overview of the setup of genetic circuits and their *in vivo/in silico* characterisation methods.

4.5 Summary

Genetic engineering is a complex process that involves altering biological units to observe a physiological change in living organisms. Here, we describe the cyclic optimisation of genetic devices to obtain robust and well-characterised circuits. We constructed biological devices based on quorum communication: sender circuits were set to produce a certain compound (AHL) remotely detected or amplified by sensor units. In this study, we built a documented library of genetic constructs and demonstrated best model features by comparing experimental resources with modelling data. For the development of our barcoding platform, these characterisation studies led to a thorough documentation of biological circuits, and were the first parts of the Bac2code biorepository. In the next chapter, we detail the fabrication of high-throughput microfluidics biochips for characterisation of genetic libraries. We further describe in

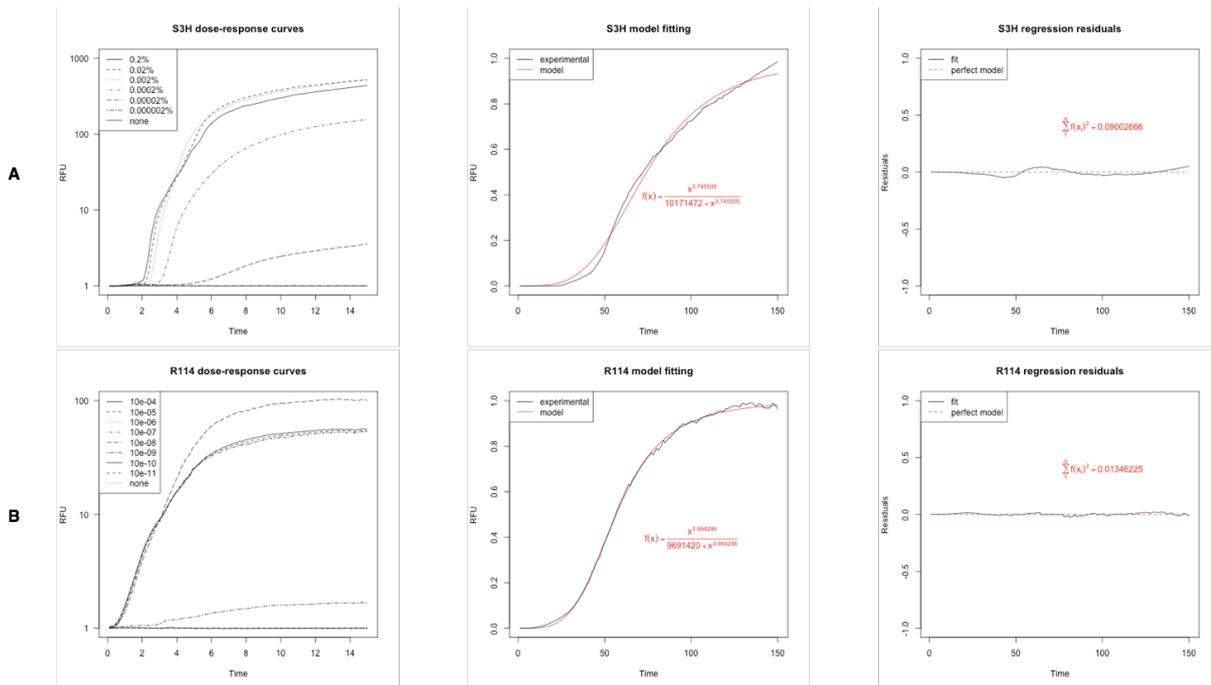


FIGURE 4.27: Optimised genetic circuits for production (top) and detection (bottom) of AHL. Top panel (A) describes the best sender circuit (production of AHL under P_{BAD} regulation) and bottom panel (B) the best receiver device (LuxR under the control of $J23104$). First column displays average experimental data, the second shows the fitting curve to maximum induction profile and the third column represents the model residuals.

Chapter 6 how the setup of a genetic library - together with a microfluidics platform - may be integrated into a barcoding or screening platform.

Chapter 5

The making of microfluidics devices for bacterial screening

In this chapter, the work presented for the channel-type biochips was based on a silicon wafer that was designed by Dr. Sunny Park.

5.1 Introduction to microfluidics studies

Modern technologies allow for the casting of specific microfeatures onto a range of templates. Arguably, one of the most famous fabrication methods is 3D-printing[7, 36, 36, 119]. This kind of technique has been proved successful at creating microenvironments suitable for bacterial growth, but the resolution of current equipment rarely goes under the $100\mu m$ order of magnitude. Since *E. coli* cells average a diameter of $1\mu m$ and a length of $2 - 5\mu m$, the study of specific bacterial communities needs a higher resolution, which is achievable by microfluidics. In fact, the production of microfluidics is analogous to "reverse 3D-printing": instead of polymerising material, microfluidics templates are often made by etching chemical layers over a silicon surface. Whilst 3D-printing involves polymerising material onto a template, the process of photolithography used to create microfluidics templates usually makes patterns in the opposite orientation of the ones generated via 3D-printing. This provides several advantages based on accuracy and reproducibility of higher resolution features. Also, silicon wafers used for microfluidics design printing are routinely used in microelectronics, which makes them well-studied and reliable materials. In this study, we designed and fabricated PDMS microfluidics chips to exhibit series of micron-/submicron-size features suitable for bacterial growth. Here, we introduce specific notions that drove the design process of polymer biochips.

5.1.1 The use of microfluidics chips with bacteria

The use of microfluidics chips for the study of bacteria allows the programming of dynamic functions, just as bacteria can be engineered with genetic devices. A system of interconnected channels dictates fluid direction within a microcircuit in which bacteria multiply[223, 65]. Chambers exhibiting micron-scale features usually trap cells and render time-lapse observations possible in a set of confined compartments. Microfluidics chips are connected to a wiring system ensuring continuous input of fresh nutrients and output of waste products; in general, this manifold is linked to devices that exert an external control over fluids flow (syringes and pumps). This external control allows for the fine-tuning of environmental conditions in order to adjust physical forces and differential input effects over bacteria. Altogether, this well-controlled environment makes the *in vitro* tracking of cells possible and provides a single-cell resolution solution to the study of particular bacterial behaviour[72, 110]. In particular, biologically engineered cells can be tracked via an automated imaging system that renders reconstitution of discrete cellular events over a continuous scale feasible, and provides a better understanding of cellular physiology.

5.1.2 Small world effect

Due to tiny dimensions, microfluidics use very little medium and generate an equally low volume of waste. For instance, compared to large scale chemostats that require litres of sterile medium, microfluidics only use a few millilitres which is sufficient to test one sample over several hours. Since the dimensions of chips are so small, it is easy to engineer devices with combinatorial versions of a design and, therefore, to be able to capture tens of different conditions at once[211]. In particular for biological studies, the ability to reproduce experiments over a number of iterations is essential to draw meaningful conclusions. Although microfluidics do not directly solve the problem of replication of experiments, they tackle it by providing the possibility to set series of design variations and repetitions, all observable at the same time in the same biological context. This adds to the statistical significance of studies, where representative data can be averaged from a large pool of samples. Thus, small dimensions in the microfluidics world - even if they affect the laws of physics - should be considered as an asset to ease technical and biological replication. However, even if microfluidics experiments are cheap and repeatable, the process of making specific biochips may sometimes involve considerable labour.

5.1.3 PDMS substrates

For the setup of microfluidics chips, we chose to use PDMS as a substrate over agarose in order to obtain higher resolution, more stable features[214]. PDMS has very advantageous properties that make it a polymer of choice for replication of small template features[142, 151]. It is a highly viscous compound with good flexibility - that makes it easy to work with - and exhibits unique physico-chemical properties (stable and biologically inert). At room temperature, PDMS is a gas-permeable transparent material that provides a good support for microscopy studies involving fluorescence measurements and direct observation of cells. While it prevents water-based solutions from leaking due to its high hydrophobicity, PDMS is also a porous medium that gases can easily permeate. In the context of fluorescence bacterial studies in particular, it is beneficial to provide enough oxygen for efficient protein folding[190]. After fabrication, PDMS substrates can be plasma activated and surface silicon groups bonded to oxygen atoms[101]. By applying the same treatment to borosilicate glass coverslips, this provides the ability to form a strong chemical bridge between both substrates and results in creating sterile microdevices ready for experimentation.

5.1.4 Scaling up bacterial screening

For the screening of bacterial libraries, microfluidics can adapt high-throughput designs to allow combinatorial *in vivo* studies. In the context of fluorescence microscopy, it is better to grow single layers of cells to obtain higher resolution over large fields of view. In practice, this was accomplished by hindering *E. coli* within a $1\mu\text{m}$ height surface area. By reducing the possibility of vertical stacking of cells, bacteria could be studied in single cell or small community contexts and followed over time-course experiments.

We have shown in Chapter 4 the possibility of wiring bacteria through certain chemical inducers. We postulated that these may be characterised in high-throughput microfluidics settings and fabricated polymer biochips to detect signal transduction via intercellular communication[177]. As shown in Figure 5.1, small bacterial communities may be wired and respond to chemical stimuli[223]. For the physical setup of bacteria in microfluidics, we engineered observation chambers that connect two potentially interacting colonies over variable distances. For each design, we replicated a range of conditions in which cells may show differential behaviour due to a change in

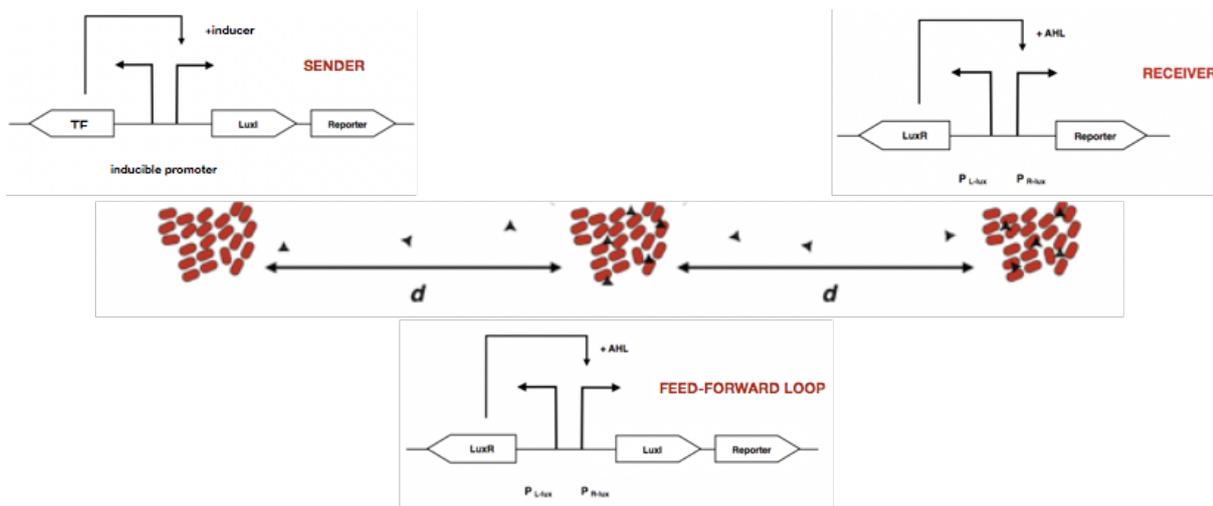


FIGURE 5.1: Schematics of communicating microcolonies of bacteria. Each colony carries a specific genetic circuit aimed at producing or sensing a quorum molecule (AHL, small triangles between red cells from different colonies) over a certain distance (d).

inducer concentration. Technically, over distances ranging from $1\mu m$ to $100\mu m$, we fabricated PDMS chips to study communication between microorganisms, allowing very small size colonies (1 – 200 cells) to be observed in continuous time.

5.2 Design of biochips

In the bacterial world, cell behaviour and timing of replication are often dictated by the availability of fresh nutrients. Standard bacterial cultures are usually grown in flasks and always reach a phase of saturation when left for a long enough time. However, chemostats can be used to sustain continuous bacterial growth with a constant input/output flow. Microfluidics are a micro-alternative to chemostats: in fact, microfluidics chips developed for use in biology are chemostats. They simply follow the same principles, but in a much more confined space[199, 65, 17, 199]. Due to the microscopic size of microfluidics, this small space affects the laws of physics that rule fluid dynamics behaviour. As shown in Figure 5.2, fluids in microfluidics behave more like a river in which particles in water would not mix (with a low Reynolds number)[114]. The Reynolds number is defined as:

$$Re = \frac{\sigma UL}{\mu}$$

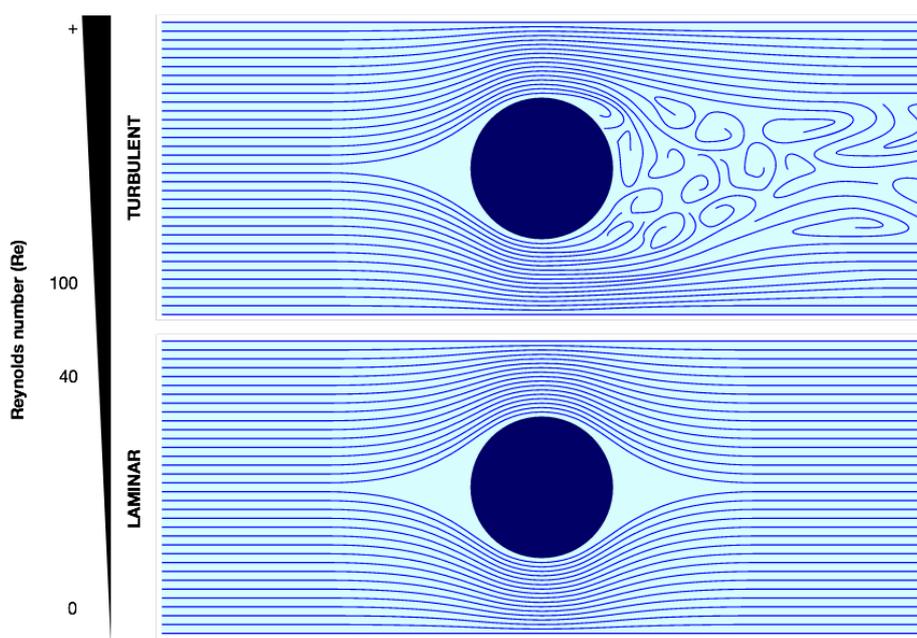


FIGURE 5.2: A Reynolds number provides the ratio of inertial forces over viscosity in fluids. Fluid dynamics change and become turbulent for high Reynolds numbers while low numbers are associated with an undisturbed laminar flow.

with μ the molecular viscosity, σ the fluid density, U its velocity and L the length scale. It corresponds to the ratio of inertial vs. viscous forces, and as the scale L changes, characteristic Reynolds numbers can jump by orders of magnitude. For instance, the Reynolds number for water in standard microfluidics size channels is 10^{-2} , which is three orders of magnitude below the flow water would have in a river. This is due to a higher impact of fluid viscosity over inertia, which results in making mixing in microfluidics a tedious process[235]. Understanding the physics of microfluidics is therefore important to ensure the successful creation of chips[196]. Although the rules of fluid dynamics vary at the microscale, overall, they form an asset to predict the path of particles in space and time.

5.2.1 Photolithography considerations

We fabricated microfluidics devices from silicon wafer templates produced via soft-photolithography[151]. We casted negatively-printed circuits onto hard-PDMS and transferred the pattern onto soft-PDMS, single-use biochips (cf. protocols in Appendix D). The circuits reproduced from silicon wafers in PDMS were first imprinted by etching different resins on template materials. In photolithography, a resin is evenly

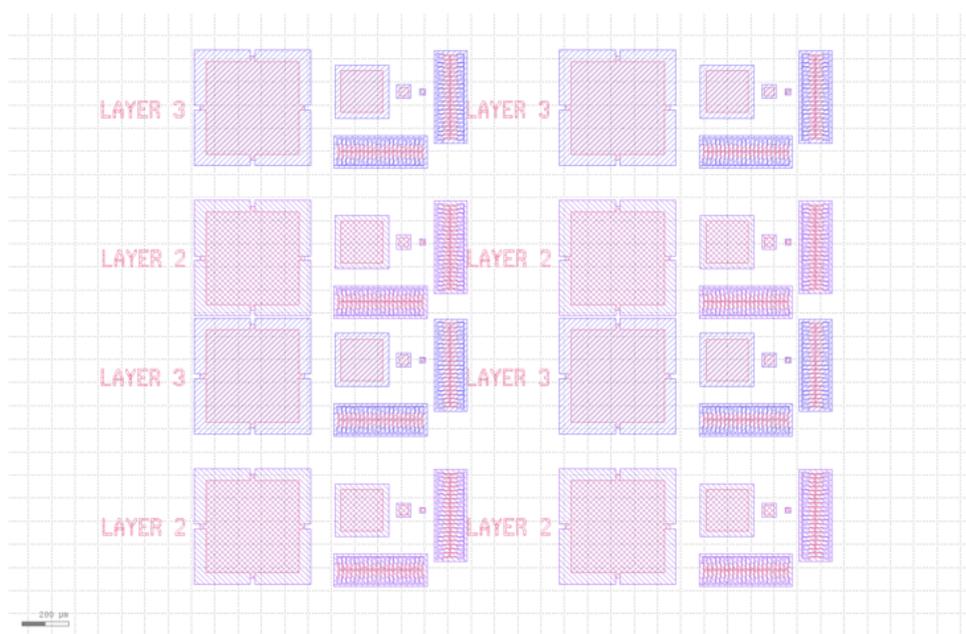


FIGURE 5.3: Redundant mask alignment pattern for the construction of multiple depths on silicon wafers. Each colour represents a specific depth layer.

coated onto a substrate and different masks are aligned to etch microstructures intended from the design. For each depth level of the design, an additional mask is necessary to build specific features. The outlook of a specific wafer is always subject to the photoprinting company manufacturing it. In our study, we delegated to two Newcastle-based companies the production of 6 inch diameter silicon wafers with microfluidics imprints. In order to align the different masks and to produce the intended structures, each wafer contained redundant features aimed at focusing a mask aligner for photopolymerisation or etching. Such features are shown in Figure 5.3, where each layer can be aligned via overlapping and unique features between depth levels. After correct alignment, each depth level is created successively starting with the thinnest and ending with the thicker layers. Altogether, this procedure resulted in the production of three individual wafers including 33 biochips of interest to this study.

Microfluidics structures were imprinted in negative onto silicon wafers. In Figure 5.4, we review the main differences in the design of microfluidics chips for bacterial intercellular communication. Since photolithography etches specific resins down to different depths, there were two possibilities to offer a communication surface area between distal bacterial colonies. This was achieved by either making a sub-micron size channel ($0.7\mu\text{m}$ width/depth) or a thin surface area (200nm depth) between communicating bacteria. In the following, we refer to these models of chips respectively

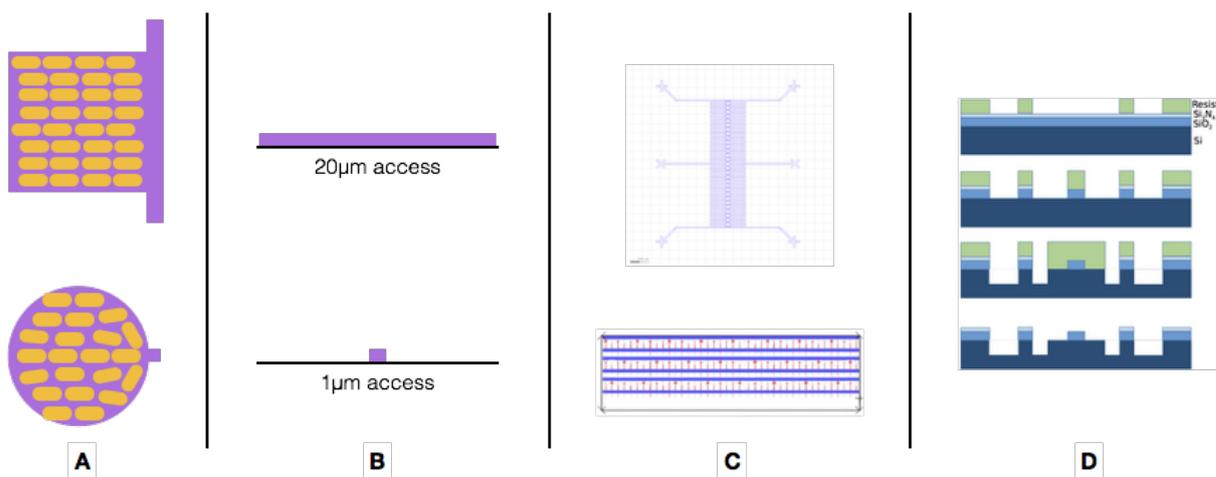


FIGURE 5.4: Considerations for the design and fabrication of microfluidics chips. (A) Top view of bacterial traps as circular or rectangular shapes in respectively channel- and surface-type chips, (B) corresponding aperture from traps to flow channels, (C) CAD design of microfeatures and (D) photolithography process overview.

by channel and surface chips. Both models were produced on different silicon wafers and compared based on cellular traps ease-of-access and their reliability for bacterial culture. As shown in Figure 5.4 A and B, observation chambers were subject to two distinct types of design that had a differential impact over input/output of cells from bacterial traps. General biochip CAD designs are displayed in Figure 5.4C, while the top chip masks/layers used for photopolymerisation/chemical etching of the silicon wafer are streamlined in panel D. To create this particular chip, silicon (Si), silicon oxide (SiO_2) and silicon nitride (SiN_4) resins were coated onto an SU-8 layer and etched in multiple steps via three different masks. Since chemical etching was orthogonal between the three resin substrates, we could ensure relative dimensions intended in the multilayer device would be respected. We compared the design methods of channel and surface biochips for the setup of high-throughput bacterial screening in microfluidics, and we optimised PDMS chips fabrication protocols (cf. Appendix D) for extremely small features reproduction. In the next sections, we describe the fabrication and main features of microfluidics devices that were created in this study.

5.2.2 Microfluidics chip compartments

Fabricating microfluidics chips is a multi-step process that allows the production of cheap and reproducible microdevices. Micrometre scale features are printed on a photoresist, which is in turn used to mould structures and shapes on a hard substance

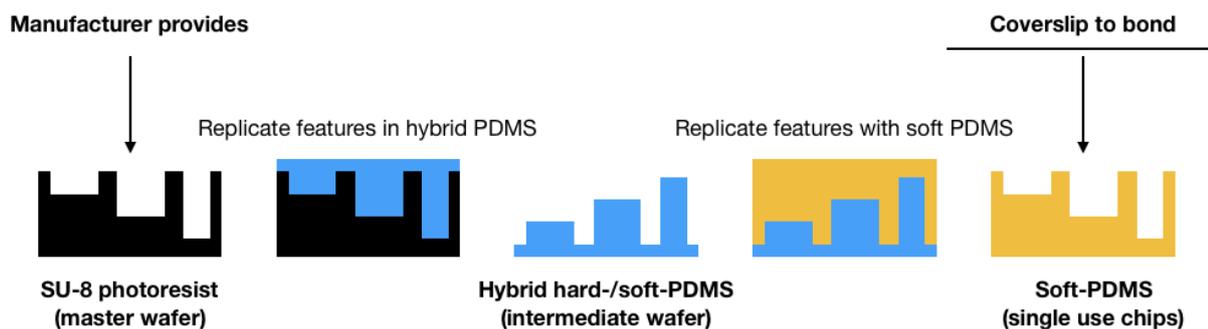


FIGURE 5.5: A master wafer is first manufactured by photolithography and contains specific microfluidics patterns to be cast onto PDMS. In order to obtain the same pattern orientation, we first generated an intermediate mould made of hybrid hard-/soft-PDMS. This intermediate was then used for the casting of final single-use biochips.

(usually agarose or polydimethylsiloxane (PDMS))[157, 92, 249]. The hard substance covers the entirety of a given microfluidics design of channels and chambers, and seals structures as a mirrored imprint. Depending on the negative or positive orientation of features present on the first polymer, it is sometimes necessary to reproduce the design on a second hard substance mould to obtain usable devices. As shown in Figure 5.5, in this study, we built sub-micron scale features that needed to be first imprinted in negative due to size and precision limitations of their manufacturing process. Therefore, a two-step process was necessary: first, a reusable intermediary mould was produced which then was used as template for the making of single-use individual chips.

5.2.2.1 Physical considerations associated with PDMS structures

Very small structures are rendered best with PDMS[92]. While agarose chips are easier to make, they degrade faster and are not particularly suitable for small design features. PDMS is a hard polymer that can support high height:ceiling length ratios and provides an improved rigidity necessary for small structures to hold[65]. Generally, if the design for observation of bacterial cells is intended to comply with 1D-/2D-growth (in a line or a plan, respectively), it is recommended to keep chambers between 1 and 1.5 times the diameter of bacteria. For work in *E. coli*, rod-shaped cells are usually immobilised by a ceiling height of $1\mu m$. Smaller ceiling heights may not be able to trap cells while larger heights can allow for stacking of cells (3D-growth), and render single-cell measurements more challenging.

5.2.2.2 Observation chambers

In a microfluidics device, due to space restrictions and flow control, bacteria move quickly. Since we aim to characterise single-cell or small population scale events, it is necessary to provide windows in biochips, where cells can be fixed by a lower ceiling height[65]. These observation chambers, or traps, immobilise bacteria and allow time-lapse observation of their growth. Besides averaging a correct height for cellular immobilisation, traps should also provide a suitable interface for bacterial growth[220]. This implies that chambers should display a wide enough aperture to capture cells flowing in high-speed channels and enough fresh medium, but also that this aperture should be relatively small to limit bacteria from exiting the trap. When a chamber is full of cells, the design should account for a way to discard cells and to keep the system growing as a continuous culture. This is generally done via the same aperture as the one cells use to enter the trap, which implies that nearby high-speed channels should provide a waste removal possibility. While this approach suits best 1D-/2D-growth, 3D-growth often requires bigger apertures (causing higher fluid velocity) that rise the need for alternative chamber exits. One-dimensional growth consists of a growing line of cells. However, two-dimensional growth aims to cultivate cells on a planar surface, and this has to account for a specific geometry that complies with bacterial growth characteristics[231]. As shown in Figure 5.6, *E. coli* cells show a self-organisation pattern when reaching higher cellular densities in a 2D-growth space. This natural phenomenon usually simplifies time-lapse tracking of bacteria and eases post-experimental image processing. Altogether, a thoroughly planned and rational design of chambers should allow the observation and tracking of single-cell resolution events on arrays of microscopic chemostats.

5.2.2.3 Channels

In order to build a chemostat, any circuit needs to be closed and under constant pressure to ensure stable continuous culture. This is achieved by connecting growth chambers to higher-speed flow channels that: (i) provide an input/output interface to traps and (ii) wire microfluidics chips to external flow control equipment (syringes, pumps, waste)[65]. In general for *E. coli* studies, a good working depth is about $10\mu\text{m}$. If cells are to be studied in a two-dimension context, this exhibits a 10 : 1 ratio between channel height and observation chambers, which fits the requirements of PDMS products[65]. The minimal achievable microfluidics circuit involves a single input/output system, where two individual wires need to provide an input flow (first wire connected to an input syringe), which is in turn propagated through the chip channels

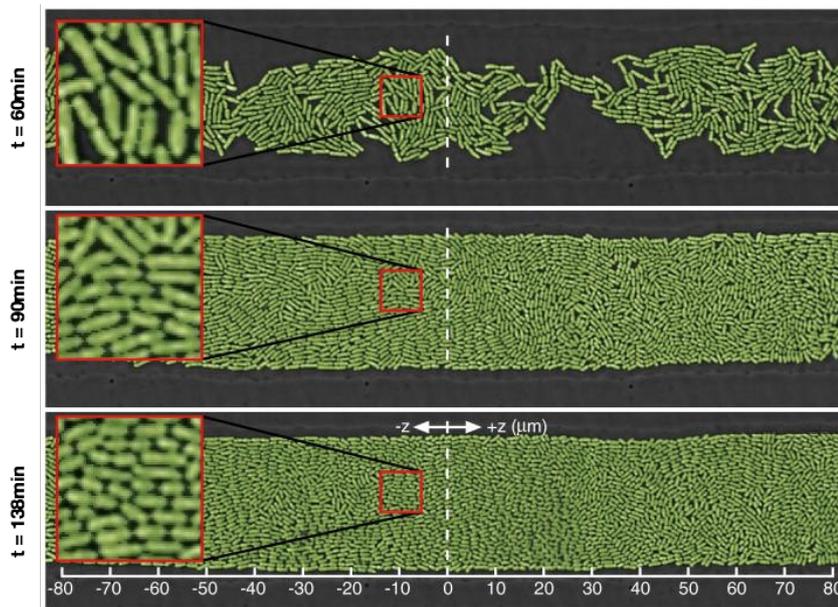


FIGURE 5.6: *E. coli* growth was observed from low to high cell densities at three time points (after 60min, 90min and 138min of growth) and showed robust cellular self-organisation. Adapted from [231]

and chambers, and exits the circuit via an outlet (second wire collecting used samples). More complex wiring may be designed to allow the mixing of multiple input/output systems, but these should be carefully planned and verified by fluid dynamics simulation to obtain parameters that would allow harmonious input/output flow rates. In general, microfluidics channels are connected to external equipment by a physical hole punched into the microfluidics chips before being bonded to a glass coverslip. Using a hypodermic needle of a slightly larger diameter than the one used to punch the chip is a cheap alternative to seal biochips to input/output ports [154]. Microfluidics channels are usually visible to the naked eye but the chamber system may be indistinguishable. Therefore, ports that connect the chip to external needle/tubing system are normally designed as larger features (1 – 3mm), with a recognisable pattern. Punching holes in PDMS is thus streamlined by the presence of these specific structures. Once a chip is bonded to a coverslip after plasma activation (necessary to bond two surfaces together), connecting inlet/outlet tubing via needles to its ports starts the chemostat, and then samples can be set for imaging.

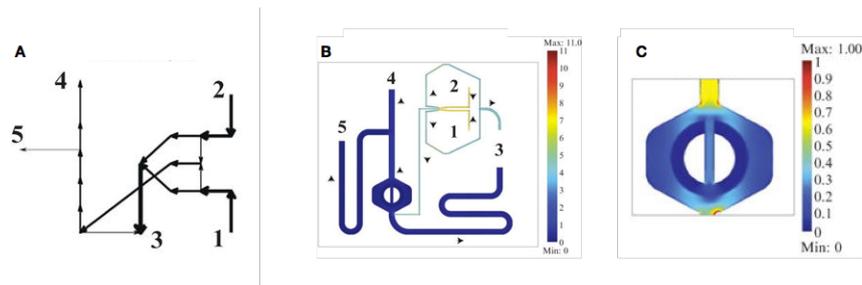


FIGURE 5.7: For a given chip design, (A) nodal analysis describes flow directions and pressure drops, (B) finite element methods provide finer fluid dynamics methodologies including (C) variable flow regimes in cross-sections. Adapted from [65]

5.2.3 Fluid dynamics

For any microfluidics design, we recommend running at least one modelling tool to predict the fluid dynamics of the circuit. This allows prediction of the behaviour of fluids and the spotting of any design errors. Two approaches are used for modelling this process: a simple node analysis or a more complex finite state elements method[65, 60]. Although the latter provides better visuals for flow simulation (Figure 5.7), it requires a software license and specific modules to run. Alternatively, OpenFOAM is a free software that can also solve fluid dynamic problems[243].

5.2.3.1 Node analysis

In fluid analysis, one method is to abstract segments of continuous channels as pipes, and junctions between channels as pressure drops. By design and according to the law of mass conservation, microfluidics circuits have an input/output system. For whatever volume of fluid there is to enter the chip, the same volume of fluid needs to be extracted from it by an outlet system[65, 18]. In node analysis, nodes (or junctions between channels) are characterised as a system where the node can only handle a certain amount of fluid at a given time.

The aim of this analysis is to check that design features are rational. Should they not be, results from this analysis would indicate faulty junctions that fail to provide viable input/output rates. Overall, node analysis is a straightforward method to verify microfluidics designs feasibility, and can be achieved relatively easily via a set of free/commercial existing tools (R, Matlab or specialised apps).

5.2.3.2 Finite element modelling

Methods of modelling exist to assess dynamic properties of microfluidics systems[261, 60]. However, these are a combination of complex algorithms and are mostly registered with specialist companies. The most common one offering these methods with tailored algorithms for microfluidics is Comsol. Even with the use of Comsol, for a better compatibility, microfluidics chip designs should always be exported as computer-aided design files (.cad, or CAD files). Different encoding formats exist for the description of design files but the CAD system is one of the few universally recognised. Specialised software such as AutoCAD or L-Edit can be used to design and to export models for such purposes.

In finite element modelling, a mesh is first created throughout the microfluidics circuit. Given determined input and output parameters, the software runs fluids through the chip and their dynamics are simulated in order to predict any undesirable behaviour in the channelling system. As an additional feature, diffusion processes can also be modelled and used to predict the behaviour of certain species of interest. In bacterial signalling, it is essential to verify that metabolites used as signals go where they should.

5.2.4 The making of biochips

Two types of polymer biochips were made in this study. Due to design restrictions given photolithography techniques, microfluidics circuits were negatively printed on SU-8 photoresists. It means that photoresists first needed to be mirrored into a positive and hard-PDMS pattern (master template), that could then be used for production of single-use soft-PDMS chips in series (cf. Figure 5.5). The photoresist manufacturing process was delegated to specialised companies: either INEX in Newcastle (UK) for the surface-type biochips, or a company based in Belgium for the channel-type biochips (name undisclosed). The subsequent steps of making chips were carried out by myself, and protocols on how to proceed can be found in Appendix D. In the next sections, we characterise the two main types of biochips and discuss their use with bacteria.

5.3 Channel-type chips

5.3.1 Design

The first type of microfluidics biochips we tested was developed by Dr Sunny Park. From a silicon wafer carrying specific design features, we replicated the circuit into PDMS substrates and used these in a cellular context. This design contained features that are represented and simplified in Figure 5.8. Two bacterial strains may be studied in proximity from one another by feeding individual flow channels with respective bacteria. Cells were trapped in circular observation chambers and communicated via a $0.7\mu\text{m}$ diameter channel that allowed diffusion of cellular compounds between colonies such as inducers. The distance that separated microcolonies was invariant, but the number of colonies placed either side of the diffusion channel could fluctuate (cf. Figure 5.8).

5.3.1.1 Cellular flow

Under steady-state conditions, as per design, cells should grow in a single plan and achieve intercellular communication via diffusion channels. As we demonstrated in Chapter 4, the activation of genetic circuits can trigger physiological changes, observable at a single-cell level via fluorescence microscopy. In order to analyse differential communication behaviour in cells, we studied microcolonies of variable size undergoing the same physiological changes simultaneously. As shown in Figure 5.8, colony sizes may vary from a single cell ($1\mu\text{m}$ width channel, similar to a mother machine device[233]) up to tens of bacteria ($40\mu\text{m}$ diameter circular trap). In these settings, variable activation of a second distal bacterial colony could be compared in high-throughput settings.

5.3.1.2 Unique identifiers

The identifiers used for channel-type biochips were designed by Dr. Sunny Park. Combinatorial studies provide a strong advantage to analyse quantities of biological replicates in parallel. However, the data obtained from such experiments needs to be thoroughly documented in order to provide meaningful information. In microfluidics, structures usually cannot be recognised by the naked eye, and they need specific labelling distinguishable by microscopy for correct identification. For the construction of channel-type biochips, two individual identifiers were appended to the pair of communicating bacterial colonies within each feature (cf. Figure 5.10C). On the one hand,

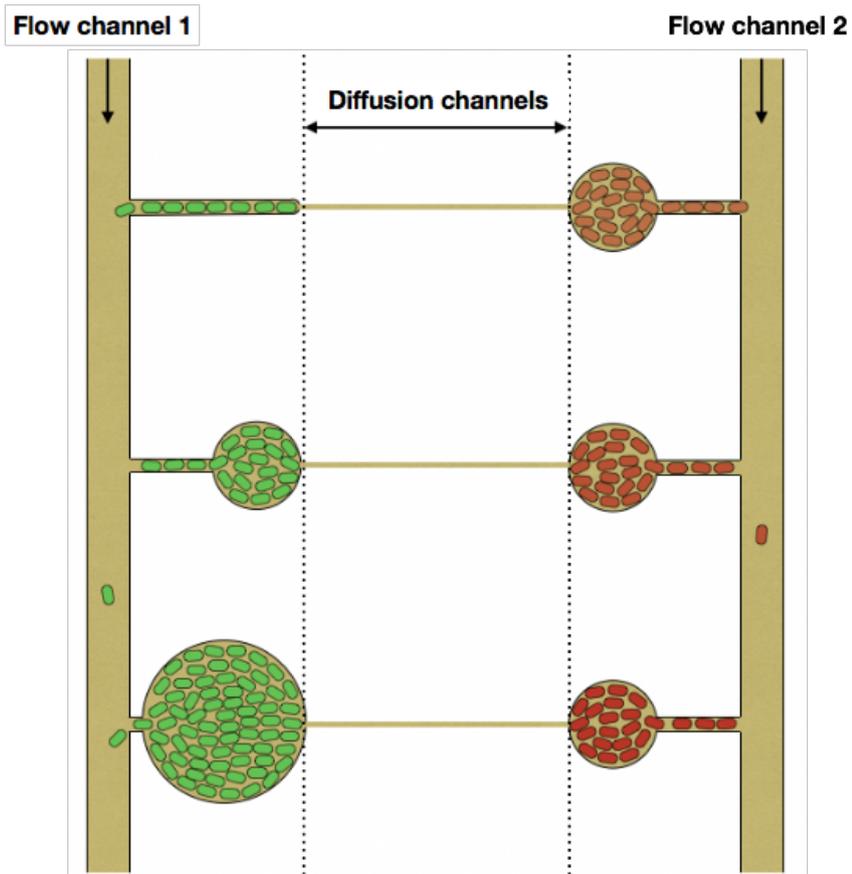


FIGURE 5.8: Simplified schematics of channel-type microfluidics biochips. Bacterial strains can be observed simultaneously in variable size microreactors connected by a thin diffusion channel. Inducers can be added to flow channels directly or propagated in diffusion channels.

each combination displayed a human-readable code to help the experimenter to locate and understand microfluidics features while setting up the experiment. On the other hand, this identifier was complemented by a machine-readable code. This code consisted of a pattern with unique description features (similar to a barcode) that image-processing pipelines may recognise for automated processing of microscopy images. Thus, the combination of interconnected bacterial strains together with specific identifiers allowed an easier mapping of high-throughput data to the theoretical design features.

5.3.2 Fabrication of PDMS devices

Channel-type microfluidics chips were produced from silicon wafers in a two-step process detailed in Appendix D. General microfluidics protocols should be adapted to different designs. For the setup of micron- and submicron-size features, one must ensure that all surfaces are well covered and conditioned for an easy separation between template mould and replicated PDMS circuit. For instance, silanisation of templates is generally performed overnight in a closed container but needed to be carried out in high-pressure vacuum chambers for our microfluidics devices, in order to provide a uniform coating of silanising agent on all surfaces. This step was very important as small features may tear whilst separating the template from newly creating PDMS chips, rendering both substrates unusable. Moreover, the use of harder PDMS substances for circuit replication was also crucial but trickier to cast into tiny dimensions. Again, under vacuum conditions, we used a high-speed spin coater to overcome this issue and to reproduce microfeatures with high accuracy. In the following, we provide an overview of the channel-type PDMS microfluidics chips examination.

5.3.2.1 SEM screening

As it was described above, we optimised PDMS chips fabrication protocols in order to obtain circuits that accurately reproduced the intended microfeatures. To verify the integrity of intermediate PDMS chips (hard-/soft-PDMS hybrid) features, we used scanning electron microscopy (SEM) to observe samples with a great level of detail (cf. Figure 5.9)[150, 179]. An overview of channel-type biochips is provided in Figure 5.10. PDMS chips were first cut with scalpels into small samples fitting SEM observation chambers, then coated with a thin layer of gold (5nm) and finally mounted onto the microscope. SEM equipment relies on blasting a high-power electron beam onto different specimens. Object surfaces and shapes are then computationally reconstructed

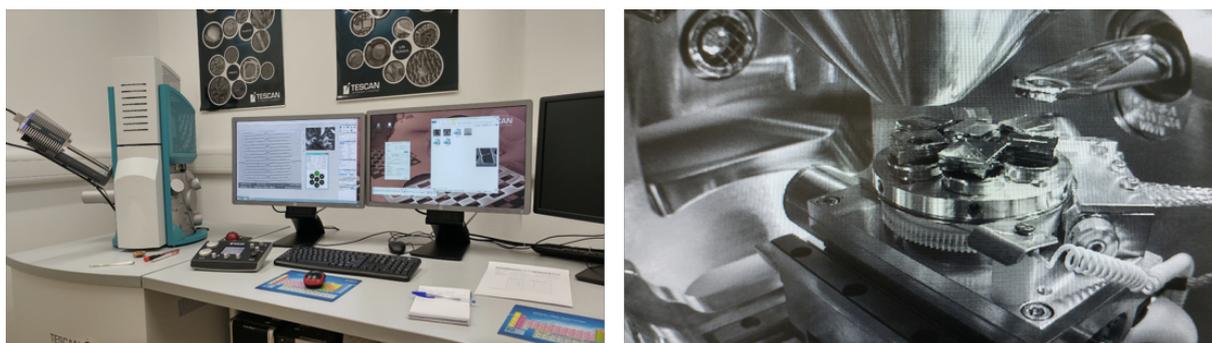


FIGURE 5.9: Overview of a scanning electron microscope. On the left, we show a picture of the microscope setup, with an incubation chamber for specimen visualisation, an electron beam apparatus and a computerised platform for dynamic treatment of data. On the right, we display the view from the interior of the specimen observation chamber. A central platform features up to 7 specimens that should be gold-coated before imaging.

based on the differential impact of electrons over focal zones. Coating samples with a conductive metal layer (5nm gold in our case) allowed to limit undesirable effects such as sample degradation whilst imaging. Although it is usually not necessary to use a clean room for the construction of microfluidics devices, we report that doing so greatly increases cleanliness of small features, as observed by SEM. Throughout this chapter, we aimed at keeping the microfluidics chips fabrication process as clean as possible in order to provide a better chance to get functional designs in *in vivo* conditions.

Figure 5.10 displays an overview of the features reproduced in PDMS from the channel-type silicon wafer. In panel A, we show a large view of five potential cellular combinations, where a cellular compartment varying in a $1\mu m$, $10\mu m$, $20\mu m$, $30\mu m$, $40\mu m$ size range communicates via a thinner diffusion channel with three individual complementary *mother machine* devices[233]. In these *mother machine* features, a single mother cell is blocked at the end of an observation channel, and its daughter cells are extracted by pushing each other through the aperture of the trap. For each of these combinations, a negative control is included on the other side of the flow channel of sensor bacteria. In these settings, higher flow in deeper channels avoids any signal to be detected by cells placed on the opposite side of a specific experimental condition. In Figure 5.10B, we provide a zoomed-in picture of diffusion channels: while cell chambers show a size of approximately $1\mu m$, diffusion channels are designed to be thinner in order to block the passage of cells and eventual biological contamination. Identifiers that were discussed in the previous section are visible in panel C. This specific example reads out as the combination number 35, where 4 cells (3 test samples and 1 negative

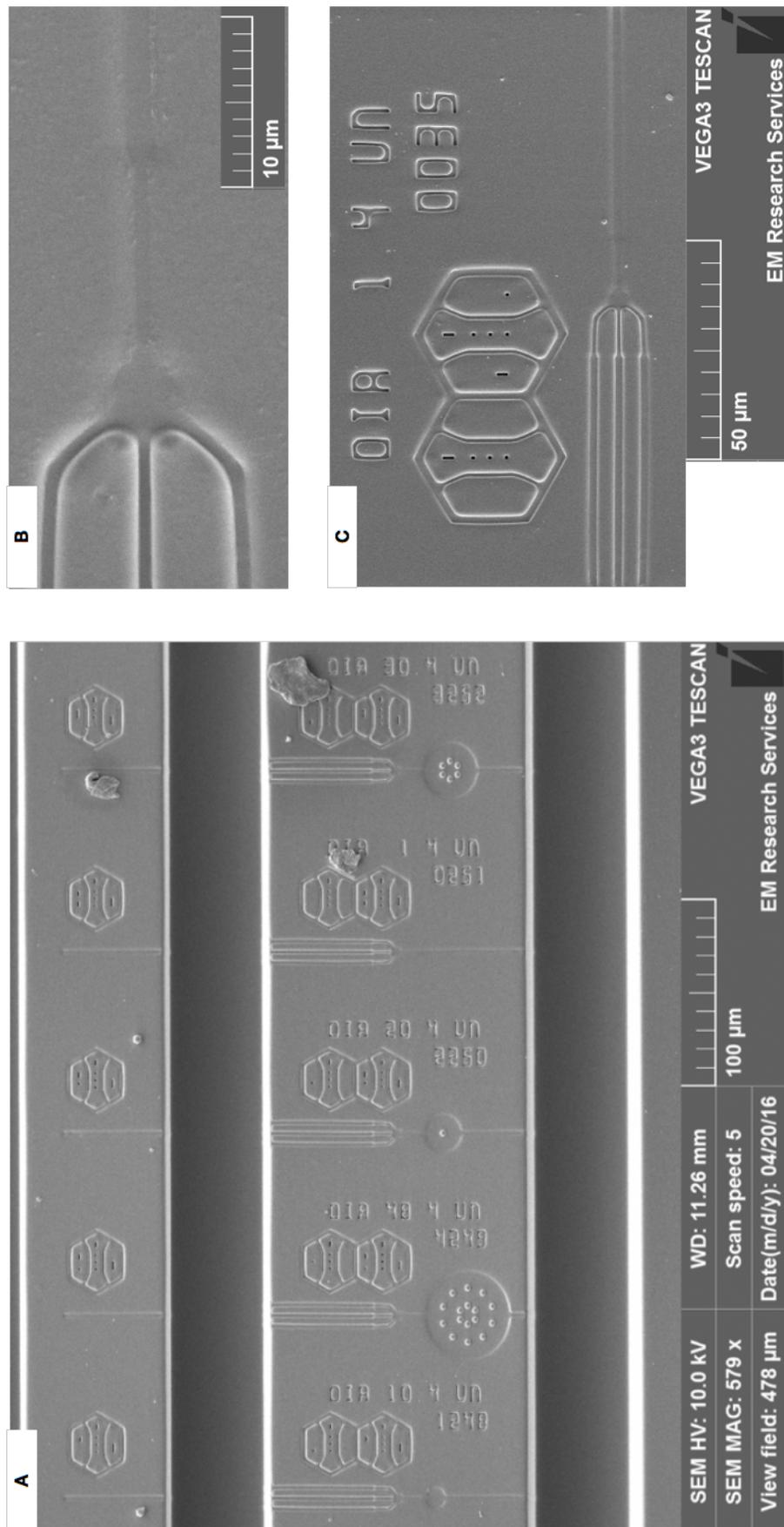


FIGURE 5.10: Channel-type microfluidics biochips screened by scanning-electron microscopy. (A) displays an overview of microfeatures that allow communication between three *mother machine* type channels and opposite variable size microreactors. (B) shows a zoomed in view of 3 vs. 1 connected *mother machine* devices and (C) displays a single sender/receiver cell combination, potentially usable for bacterial screening.

control) are set in proximity of a $1\mu\text{m}$ trap. Alphanumeric indices refer to wafer-specific abbreviations providing information about design location and chip version. Overall, with the use of SEM technology, we ensured that PDMS chips accurately replicated wafer design features and that a correct behaviour should be observed following the input of cells in coverslip-bonded microfluidics chips.

5.3.2.2 Adequate flow verification

Before injecting a flow of cells for direct and time-lapse observation in microfluidics chips, it is recommended to verify the functionality of the biochip flow system. This was done by using a food colourant visible to the naked eye as a red dye. Besides providing a direct visual verification over the directionality of flow channels, it also helped with visualising smaller features on a fluorescent microscope setup. Since the dye used to input flow in the microfluidics chip was coloured, red-channel fluorescence provided strong evidence for locating immersed regions of cellular traps. In Figure 5.11, four individual combinations of cellular chambers are shown over the fluorescence channel. At the bottom of the different images, small pillars may be observed in bigger cellular compartments: these were set on purpose in order to strengthen the observation chambers. For all channel-type microfluidics chips that we verified through this characterisation method, we always observed a correct distribution of medium over microfeatures and obtained correct flow directionality. Here, we demonstrated the concordance between theoretical design and *in vitro* characteristics of PDMS microfluidics chips. In the following section, we expand our model to *in vivo* conditions with the input of cells to channel-type biochips.

5.3.3 *In vivo* studies

For the study of bacterial communication in these chips, we used two genetic circuits that were engineered in Chapter 4 and thoroughly characterised via a range of techniques in discrete and continuous time series experiments. These corresponded to the optimised sender and receiver devices presented in Figure 4.27.

5.3.3.1 Cell size control

Our study relied on trapping cells in specific microfluidics chambers. However, with the design features of channel-type microfluidics chips, observation channels/chambers were inaccessible to *E. coli* under normal size conditions. The standard protocol to load *E. coli* into microfluidics chips involves growing bacteria from a fresh colony, refreshing

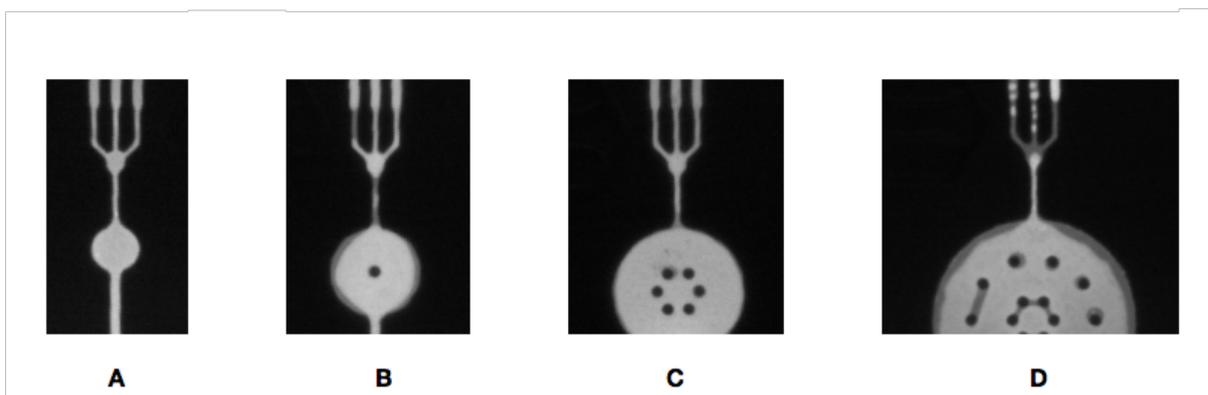


FIGURE 5.11: Fluorescence images of channel-type microfluidics features immersed in a red dye. Features visible over a black background correspond to cellular traps and diffusion channels. (A), (B), (C) and (D) respectively represent bottom traps widths of $10\mu m$, $20\mu m$, $30\mu m$ and $40\mu m$. Top cellular channels have a width of $1\mu m$ and diffusion channels $0.7\mu m$ on average.

saturated cultures and providing exponential phase cells to the input system. Given the facilities provided by the Centre for Bacterial Cell Biology, we may have been able to load cells via an optical tweezer[182], but this method would have implied to move the microfluidics apparatus after cellular loading to a fluorescence microscope for live imaging. Since *E. coli* cells undergo physiological changes as they grow in different phases, we rather tested inputting cells extracted from higher cellular density conditions - thus smaller cells - into microfluidics chips[126, 136]. Although this step did not help the efficient loading of cells into biochips, we overcame this issue by exploring poorer growth media alternatives *prior* to live cell imaging. We tested a range of minimal media, derivative of M9 salts, and managed to reduce *E. coli* size by an average of 3-4x fold compared to cultures that were grown in richer media (LB-based). This poor growth medium consisted of 1x M9 supplemented with 0.02% casamino acids and 0.5% glycerol. By growing cells to exponential phase in this poorer medium, we significantly improved cellular loading in microfluidics chips, from ≤ 1 cell per trap to several cells per observation feature (cf. Figure 5.12A).

Microfeatures aimed at maintaining cells in observation chambers were initially designed to accommodate standard *E. coli* cells grown in rich medium. Therefore, we could not keep mini-cells obtained from starvation in poor medium in bacterial traps. These were too tiny and could move in small compartments supposed to hinder cells. We thus needed to swap medium after loading of smaller cells for a richer, standard

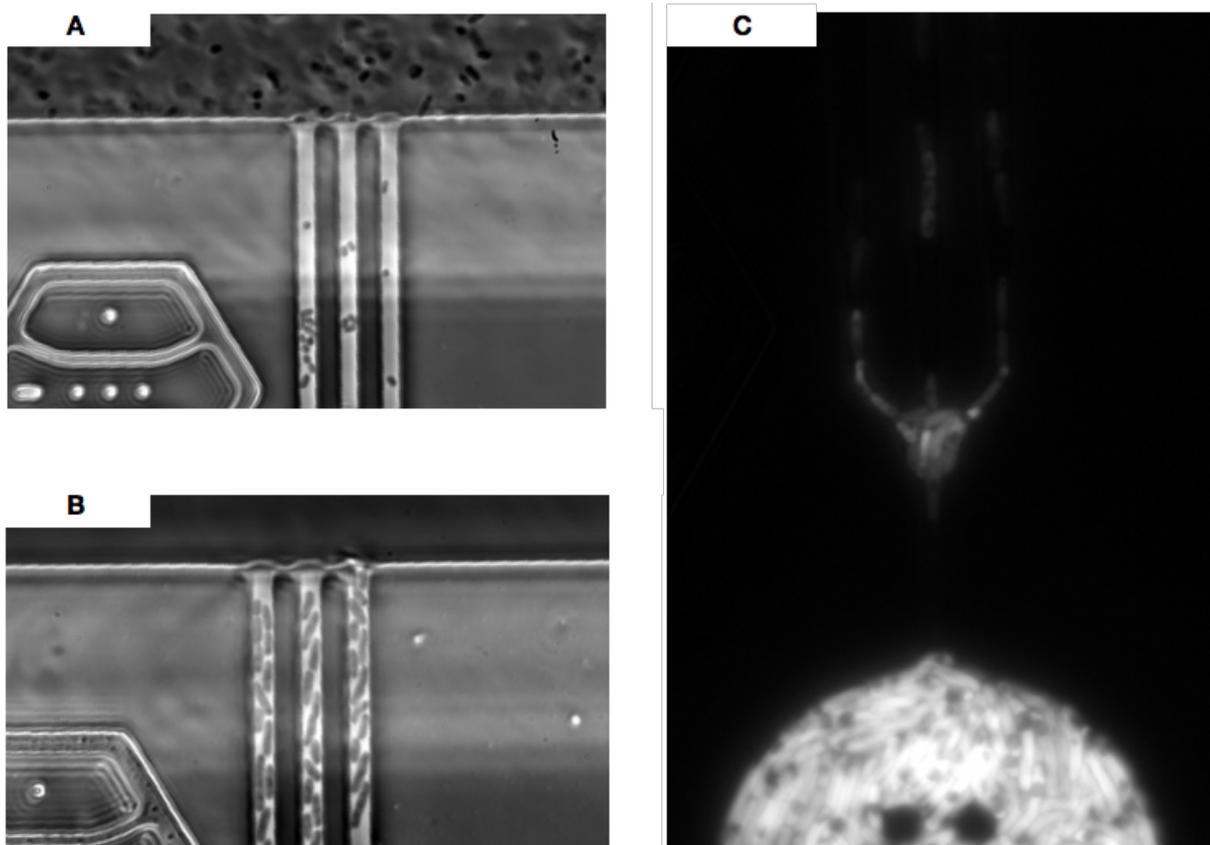


FIGURE 5.12: *E. coli* cells in channel diffusion microfluidics chips. (A) and (B) show the same location within the biochip at two different time points: after cell starvation (small cells, A) and after reverted physiological change in LB (bigger cells, B). (C) displays the fluorescence levels observed in two bacterial communities respectively sending (bottom) and sensing (top) AHL after 2h of sender culture induction.

LB growth medium supplemented with eventual inducer. Cells started recovering normal size within half an hour after providing richer medium. In general, we allowed a 2-3h recovery time for cells to grow back to normal size, and afterwards performed a final medium change to induce cultures. The physiological change from small cells (just loaded) to recovering cells after 1h culture at 37°C is shown in Figure 5.12A and B. After the obtention of normal size *E. coli* that would fill bacterial traps, cultures were induced and followed by fluorescence over time. Figure 5.12C displays a fluorescence snapshot of two sender and receiver communicating colonies in channel-type microfluidics chips. Here, sender devices were located at the bottom of the image in the large cellular compartment, while sensor cells were growing above these. Over ≥ 10 experiments, we consistently started to detect AHL sensing by the activation of sfGFP signal in receiver colonies after 2h of sender cultures induction. Overall, we managed to successfully load bioengineered cells into channel-type biochips, and we observed significant signal for intercellular communication between microcolonies of bacteria separated by $50\mu\text{m}$ within 120 minutes.

5.3.3.2 Limiting design factors

We have previously demonstrated how channel-type microfluidics chips may be engineered and adapted for *in vivo* studies of *E. coli*. However, after fabricating channel-type PDMS chips and visualising their features, it became quickly obvious that trapping cells in $1\mu\text{m}$ wide aperture channels would become challenging (cf. Figure 5.12). Although the process of cellular loading could be optimised for our purpose as detailed in the previous paragraph, we ran into major challenges when reaching high cellular densities in bacterial traps. The design of $0.7\mu\text{m}$ diffusion channels appeared to be fine to block the passage of cells over the first couple of hours, but subsequent cellular divisions ended up squeezing cells within diffusion planes. Individual bacterial colonies would then join and contaminate each other, rendering the screening of bacterial behaviour impossible in extended time-course assays. Figure 5.13A and B show the same bacterial trap at 6h of interval during which cells gradually overfilled the observation chamber. At the top of the circular trap, we observe in panel A a single cell starting to enter the diffusion channel getting progressively filled over a few hours, as shown in panel B. Besides the stacking of cells within observation chambers rendering fluorescence measurements less precise (*e.g.* cells growing in 3D), we also noticed a big deterioration in *E. coli* physiology. As shown in Figure 5.13C, cells under observation eventually acquired mutations and developed significant phenotypical changes. The cell shown here and other occurrences displayed atypical lengths of $\geq 50\mu\text{m}$. The

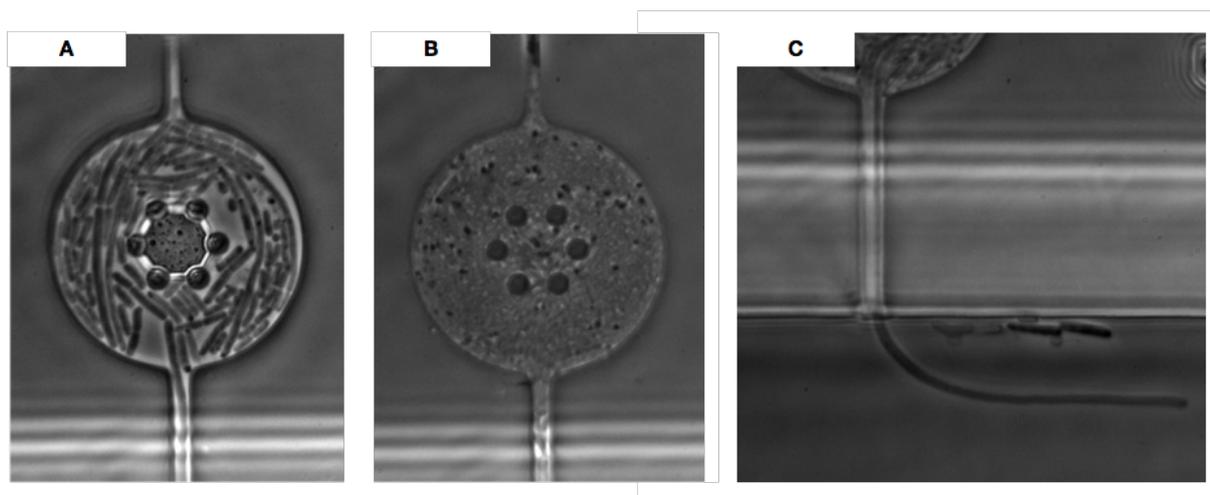


FIGURE 5.13: *E. coli* cells in channel diffusion microfluidics chips after extended culture. (A) and (B) show the same location within the biochip at two different time points: at cell induction (A) and after 6h of culture (B). (C) shows a bacterial trap where a mutant cell is hanging in deeper flow channels after an overnight culture.

development of such abnormalities could be due to the trigger of a SOS response that ends up affecting the cellular division machinery in subsequent biological pathways. It decreases cellular fitness and is proof that cellular observation chambers were not optimal for bacterial growth. This could be reasonably well explained by an aperture for input/output of fresh medium/cells too small to accommodate larger colonies. Overall, channel-type biochips were poorly adapted for bacterial growth. This stresses out the need for a correct balance between ideal design and compromises for bacterial fitness to achieve a certain function when using microfluidics chips. Therefore, based on the knowledge gained from experimenting with these biochips, we engineered other types of microfluidics chips aimed at screening for intercellular communication, and present these in the next section.

5.4 Surface-type chips

We engineered a second type of microfluidics chips and drove their design to optimise the limiting factors that were observed with the use of channel-type biochips. While we designed microfluidics patterns and fabricated all PDMS products, the manufacturing of the silicon wafer was outsourced to the INEX company, based in Newcastle-Upon-Tyne (UK). We developed these microfluidics chips to improve channel-type biochip properties observed in *in vivo* conditions. As they featured smaller channels than the

channel-type chips previously described, the casting process of PDMS needed to be optimised for a better replication of microfeatures. This optimisation process was based on the making of PDMS chips under different conditions, to then test for the adequate pattern replication via microscopy. In this section, we provide an overview of the design, validation studies and *in vivo* assays that were undertaken for the characterisation of surface-type microfluidics chips.

5.4.1 Design features

5.4.1.1 Improving biological fitness

In order to obtain a functional genetic circuit in *in vivo* conditions, two properties should be observed: (i) the genetic device should be characterised and provide robust behaviour, and (ii) cells should grow in near-optimal conditions. We have described in Chapter 4 how to optimise the design of genetic circuits and demonstrated their functionality in *E. coli*. However, the structure of observation chambers in microfluidics may impact on cellular growth conditions. The design of microbial culture chambers is subject to a number of limiting factors, inherent from the dimensions of the cells and their requirements to proliferate. To improve cells fitness and to avoid the kind of scenario described in Figure 5.13, we changed the shape of cellular traps from circular to rectangular and we expanded the aperture opening towards flow channels (Figure 5.4A-B). First, as we described earlier in this chapter (Figure 5.6), *E. coli* can adapt to an organised planar growth in straight shape channels. Therefore, we based all observation chambers design around rectangular shapes. This approach needed to provide a larger opening from flow channels to observation chambers, so we adjusted the aperture of bacterial traps to the size of their relative width. Although this approach would provide higher chances for cells to escape traps due to stronger lateral flow, it better accommodated higher cellular densities that were found to be a problem in channel-type biochips.

5.4.1.2 Creating thinner diffusion features

In the characterisation of channel-type microfluidics chips, we demonstrated that cells started entering diffusion channels shortly after the beginning of time course assays. In that design, the dimensions of diffusion features were adapted from the tiniest features achievable via soft-photolithography on a same alignment mask. Standard sizes achievable by common soft-photolithography techniques resolve around $1\mu m$. Therefore, the design of submicron size features can only be approximated when other

features (e.g. cell traps) are placed on the same dimension. To overcome this issue, we based the diffusion process between distal bacterial colonies in surface-type biochips on another photolithography mask (additional vertical dimension). At INEX, the smallest separation between silicon layers on wafers was of 200nm for the equipment that we used. So, we designed diffusion planes of a 200nm depth to separate distal observation chambers. For the most uniform propagation of inducer, we also adjusted the width of these surfaces to the bacterial traps.

In Figure 5.14, we provide an overview of one of the few types of surface biochips we developed in this study. Shown in the bottom frame of this figure, the feature underneath the 515 index displays two bacterial traps of $20\mu m \times 50\mu m$ separated by a $1\mu m$ diffusion plane. For surface biochips, we engineered 5 individual types of bacterial traps to study variable cell numbers in simultaneous conditions. These were labelled based on their size from 1 to 5 (1 being the smallest single-cell size trap, 5 the biggest trap suitable for up to 200 cells) and a similar scale was set for diffusion distances (from $1\mu m$ up to $100\mu m$). In the example given in Figure 5.14, the code 515 is read out as a first size 5 bacterial trap connected via a small diffusion distance (size 1, $1\mu m$) to a second large bacterial trap. On each side of these features are flow channels that bring cells, fresh nutrients and collect waste materials from observation chambers. Bacterial communication features are located at U-turns of deeper flow channels. These are split into a main channel that takes a sharper turn, while a smaller lateral subchannel brings cells in the proximity of traps. This subchannel reduces width as it approaches the observation trap, creating a higher flow and making cells pushing each other into the target $1\mu m$ depth area. As per design, the higher flow hereby created should also ensure that cells are correctly extracted from bacterial traps and receive enough nutrients to sustain robust growth.

5.4.1.3 Correcting design by fluid dynamics

In microfluidics, flow channels are the main effectors to input cells and to provide continuous fresh medium for optimal bacterial growth. Since several variations may be designed from a basic microfluidics pattern, we screened for best combination of features in order to obtain predictable and adequate flow behaviour near bacterial traps. First, we optimised the design of entire microfluidics chips by replicating the same conditions twice per chip. As shown in Figure 5.14 in the top right corner, if an input flow is provided in position S or R, then it is equally split over top and bottom parts of the chip biological replicas (towards S1/S2 or R1/R2). We simulated this behaviour by fluid dynamics with Comsol (analysis software) and showed the division of input

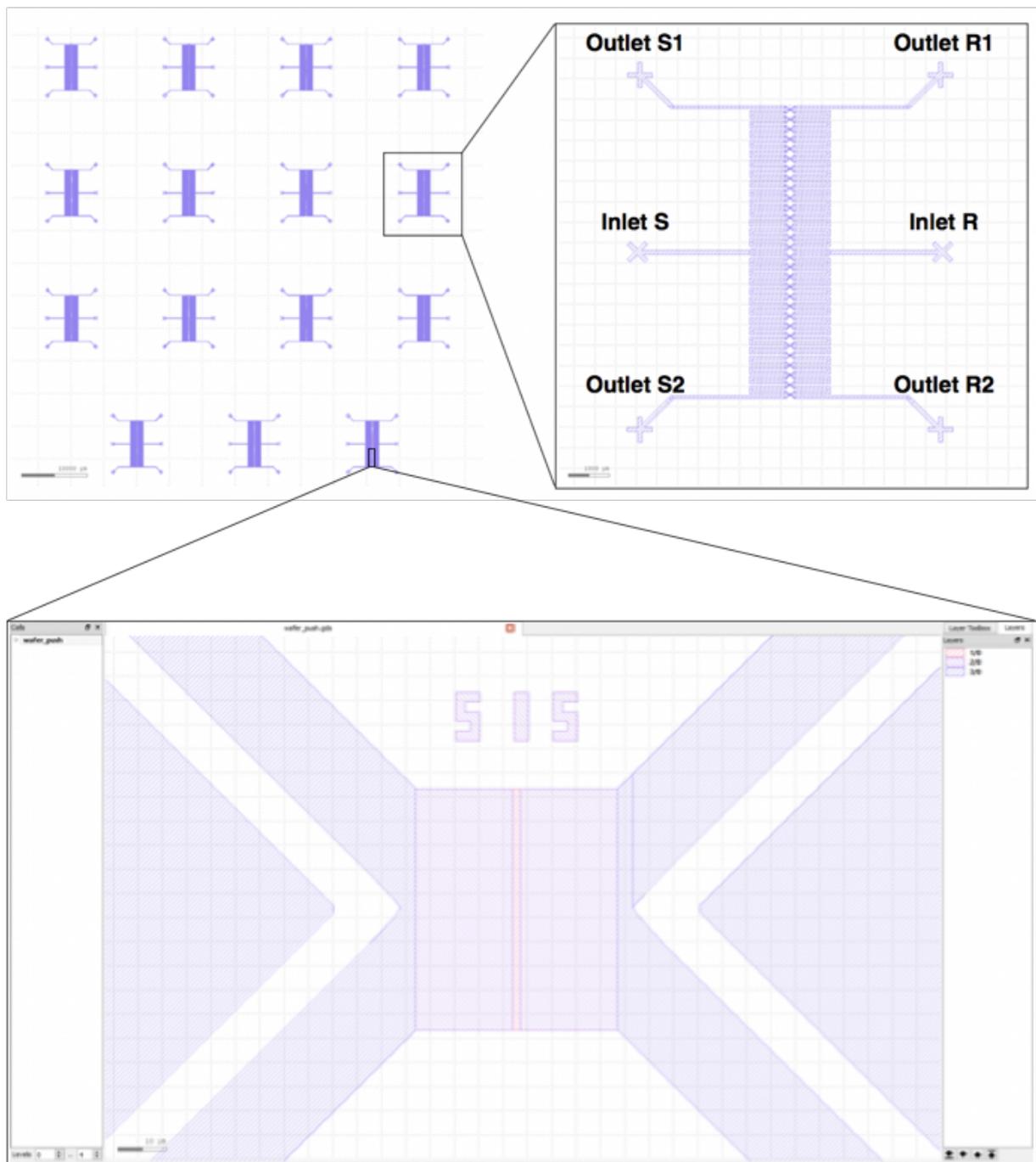


FIGURE 5.14: Overview of surface type microfluidics chips design. Top left panel represents the location of different chip designs as they should be printed on silicon wafers. The zoom-in frame in the top right corner displays the bigger features of biochips, while the bottom frame corresponds to the L-edit design an observation area[131]. In this feature, 2×200 bacteria from different colonies are separated by a $1\mu\text{m}$ diffusion plane. Along a single microfluidics chip, there is an array of 30 individual communication features as the one shown in the bottom panel, where sender and receiver colonies may be observed in different numbers at a variable distance.

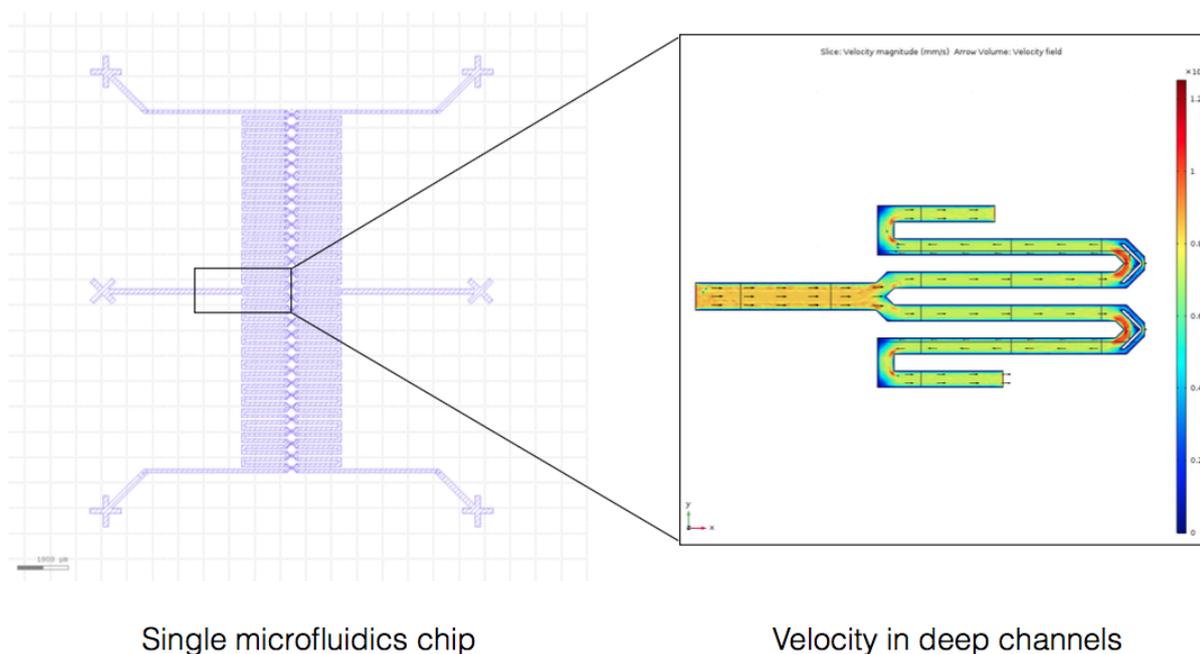


FIGURE 5.15: Cmsol fluid dynamics simulation results in surface biochips deep flow channels. Near input channel ports, a junction equally splits the main channel into top and bottom flow feeding individual biological replicas.

medium in Figure 5.15. As well as correct directional fluid flow, we designed deep channels to form a pattern similar to a spring in order to provide good mixing of the input flow and waste products. Subsequent U-turns are a simple method to mix liquids in microfluidics, and this would ensure that waste products from cells higher up in flow channels would stay separated from further bacterial traps. In these settings, we could ensure that continuous fresh medium could populate microfluidics chips and behave according to theory in larger-scale features ($10 - 60\mu m$).

In a second optimisation process, we investigated the differential impact of flow speed near bacterial traps. As it was described earlier, a small subchannel from the main input flow is diverted towards micron size features at deep channel U-turns. Here, we designed different features to restrict subchannel dimensions near observation chambers. We present the design and flow simulation of a basic microfluidics communication feature in Figure 5.16. In this instance, we ensured that deep flow channel speed would be slightly accelerated near bacterial traps, and that eventual waste products (medium or surplus of cells) could be easily evacuated from growing cells regions. Based on this microfluidics template, we generated series of microvariations in which we included PDMS features that physically pushed bacteria for cellular

loading of microfluidics chips.

With the restricted ability to load large *E. coli* cells observed in channel-type biochips, we simulated the behaviour of 5 different PDMS devices for an easier loading of bacteria. The best model that we found is presented in Figure 5.17 besides the basic system introduced in Figure 5.16. In this device, we show that the theoretical loading of a single cell can be improved by reducing the distance between cell traps and the back of flow channel PDMS features. For a $1\mu\text{m}$ depth level, we can observe a much higher flow near bacterial traps after reducing this distance to $3\mu\text{m}$, which implies that cells are more likely to be squeezed in observation chambers near these features. However, we had to consider that this higher flow and reduced distance may also have an impact over clogging of microchannels. In microfluidics, as we showed earlier for the example of channel-type biochips, one should always make compromises during the design phase to obtain an optimal function *in-vivo*. Therefore, we selected two sets of biochip designs for manufacture of silicon wafers, and we created analogous circuits to the basic and higher flow surface chips presented in Figure 5.17.

5.4.2 Characterisation of PDMS devices

Two surface chips silicon wafers were manufactured by INEX and corresponded to respective designs of basic and high flow near bacterial traps biochips. On each 6 inch wafer, we printed 15 individual chip circuits and replicated their design in hard and soft PDMS, as described in Appendix D. To inspect correct reproduction of microfeatures from wafer to PDMS, we characterised PDMS samples by scanning electron microscopy by the same method as described in Figure 5.9 for channel-type biochips.

5.4.2.1 SEM screening

In the previous section, we demonstrated the adaptation and optimisation of PDMS device fabrication for the construction of channel-type biochips. Using the same optimised protocols, we generated series of intermediate (h-/s-PDMS) and single use chips (s-PDMS), and we examined the accuracy of the microfeature reproduction process. In Figure 5.18, we display two views of PDMS chips coated with 5nm gold and mounted onto a SEM observation platform. In these assays, we managed to obtain an entire single microfluidics chip per SEM sample. This allowed us to fully characterise the accuracy of the molding process, with biological replicas located on top and bottom halves of individual biochips. In all SEM experiments with surface-type microfluidics chips, we also report that PDMS devices that were prepared in a clean room showed

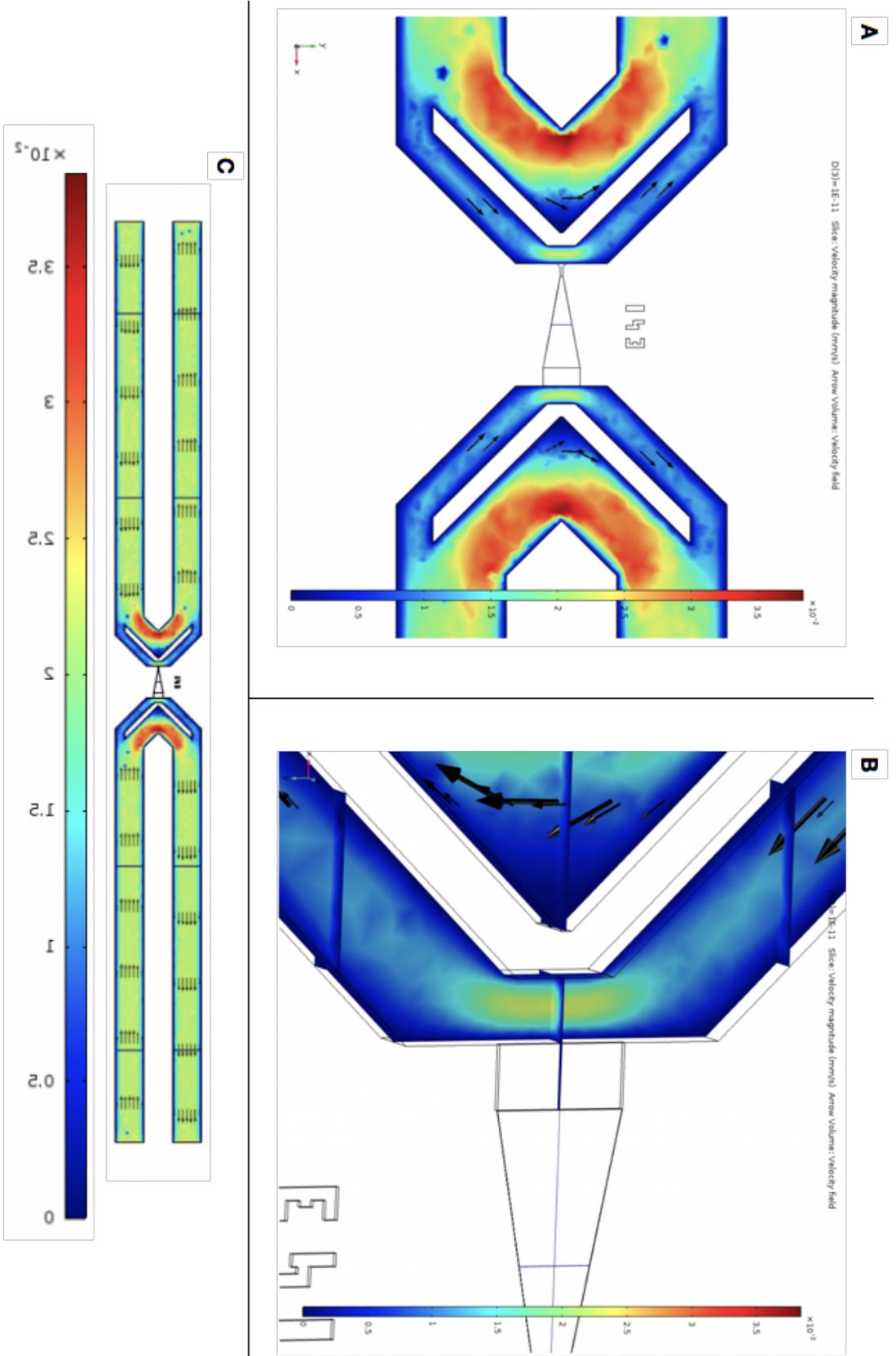


FIGURE 5.16: Consol fluid dynamics results overview for surface diffusion microfluidics chips. In all three cases, a colour gradient located next to microfluidics circuits provides legend information for flow velocity within channels for water. Flow representation is displayed for a $5\mu\text{m}$ depth (half the height of flow channels). The top left corner focuses on the simulation of a specific intercellular communication feature, where both traps around the diffusion plane are included. In the top right corner, a zoom-in view of a deep flow channel neighbouring a single bacterial trap is shown. The bottom half shows a larger overview of microfeatures including input/output deep flow channels.

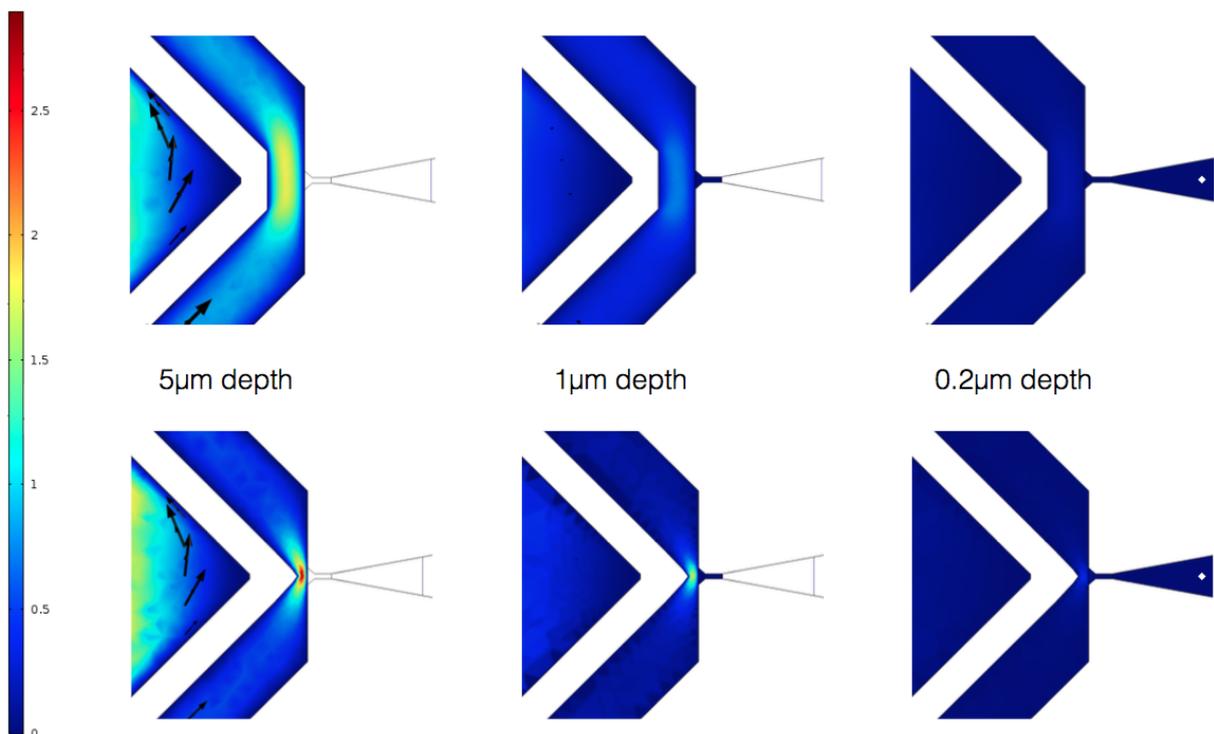


FIGURE 5.17: Models of flow acceleration towards bacterial traps in surface biochips. The three top simulation panels describe the basic flows that occur, from left to right, at $5\mu\text{m}$, $1\mu\text{m}$ and $0.2\mu\text{m}$ depths levels in basic surface chip design. Bottom frames correspond to the same simulation results for a chip that displays higher flow near bacterial traps. The left-hand side legend displays colour-coded differential flow velocities.

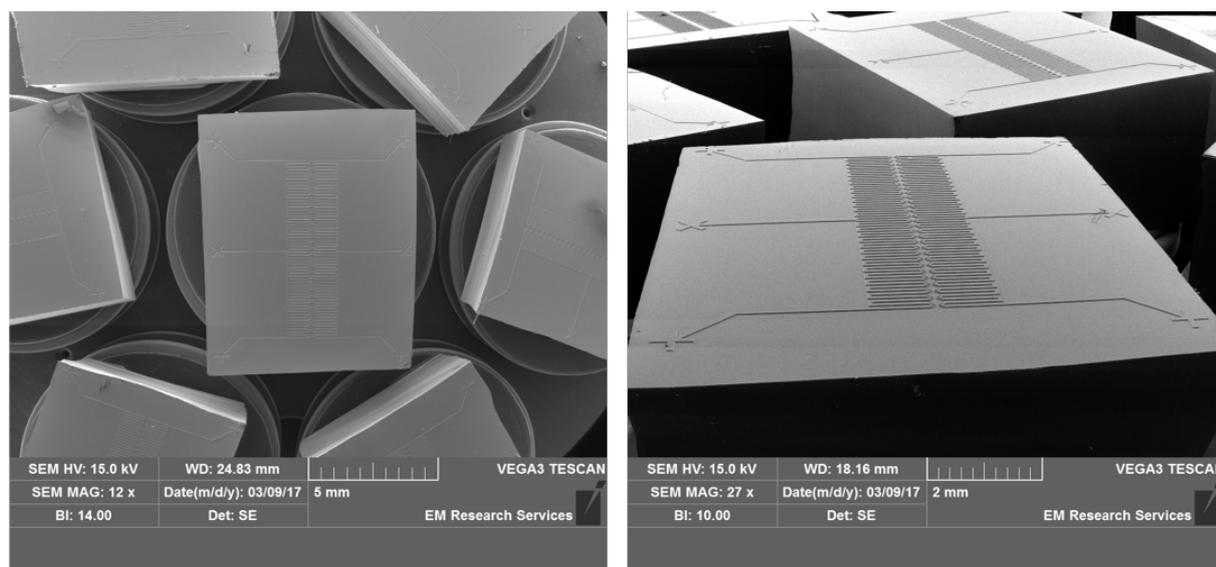


FIGURE 5.18: View of gold-coated surface biochips from the SEM internal camera. On the left, we display a overview of the 7 gold-coated specimens to observe on the rotary central platform. On the right, we show a low magnification of the entirety of a microfluidics chips at a 55° angle (chip CWqs21).

more accurate microstructures. For chips that were prepared in traditional lab conditions, we could observe the presence of multiple dust particles that may impair micron size structures (cf. Figure 5.10A on the right side). Therefore, we constructed all microfluidics chips under fume hood or vacuum conditions.

In the screening of a large set of microfluidics chips, it is inevitable to detect abnormal microstructures. In Figure 5.19, we show the reproduction of a teared flow channel rendering both intermediate and soft PDMS substrates unusable for a specific feature. The aim of optimising PDMS fabrication protocols was to minimise this error in order to generate functional circuits. In Figure 5.20, we provide an overview of a single surface-type biochip and display its associated microfeatures for half of the design replicas. This chip shows the smallest distance two bacterial colonies may be separated by in an *in vivo* context ($1\mu\text{m}$). In a general case, all hybrid PDMS chips (positive mould from silicon wafer) as the one presented in Figure 5.20 reproduced the intended design submicron size features with great accuracy. By this step of the PDMS devices fabrication pipeline, we could thereby ensure that intermediate moulds replicated all microfluidics features with good fidelity.

Following the production of hybrid PDMS chips, we generated single-use, soft PDMS chips from the intermediate templates, and examined microfeatures correct replication from mould to sample. In Figure 5.21A and B, we show the same part of a

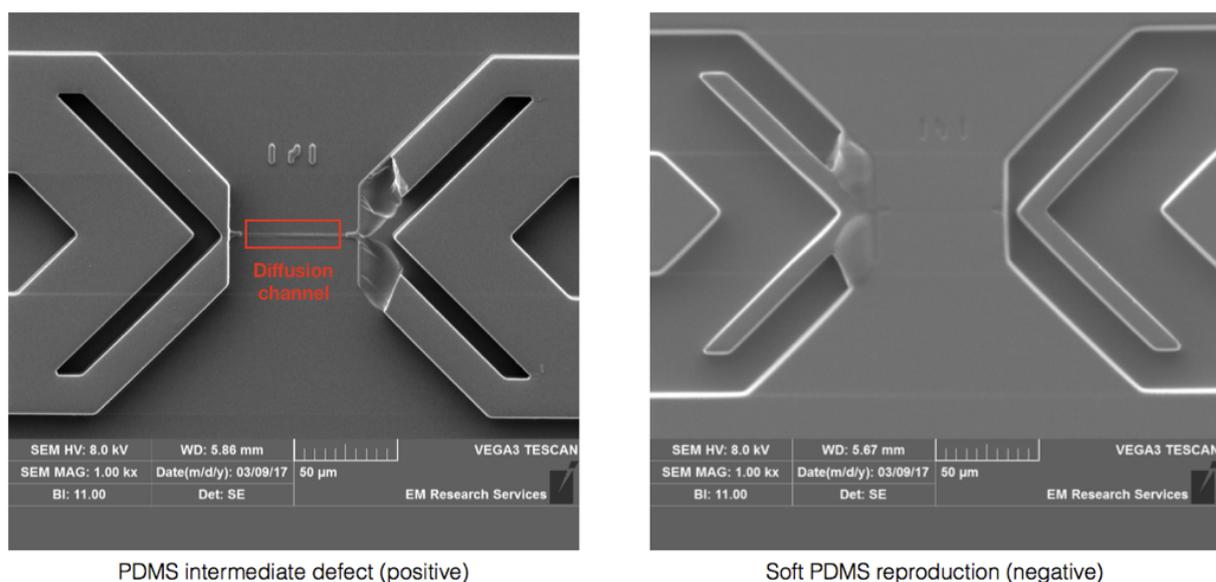


FIGURE 5.19: Potential defects in the reproduction of PDMS microstructures from silicon wafers. Left panel displays the intermediate mould that tore a microfeature, likely during separation from template. The red rectangle is placed around the diffusion channel that, despite flow channel collapse, was accurately reproduced. Right panel shows reproduction of the same defect onto soft PDMS chips.

microfluidics chip resulting from the separation of hybrid template from s-PDMS products. Features that were observed so far for intermediate PDMS devices were printed positively on hard PDMS, and panels A and B display the evidence for a consistent good replication of microfluidics features from hard to soft PDMS. In Figure 5.21C, we focussed on the observation of the smallest biological feature of all surface-type biochips. Progressive magnifications show the localisation of two observation chambers that consist of individual mother-cell devices, while a $200nm$ plane separates these by $1\mu m$. In a biological context, this could allow the study of intercellular communication between two single cells directly facing each other. Altogether, we demonstrated by SEM imaging that nearly all PDMS microfluidics chips could replicate exact submicron size features, proof that surface-type biochips may be used in an *in vivo* context.

5.4.2.2 Liquid flow check

As it was described for channel-type microfluidics chips, we performed the same simple colorimetric test to verify the functional flow of surface biochips. We injected a

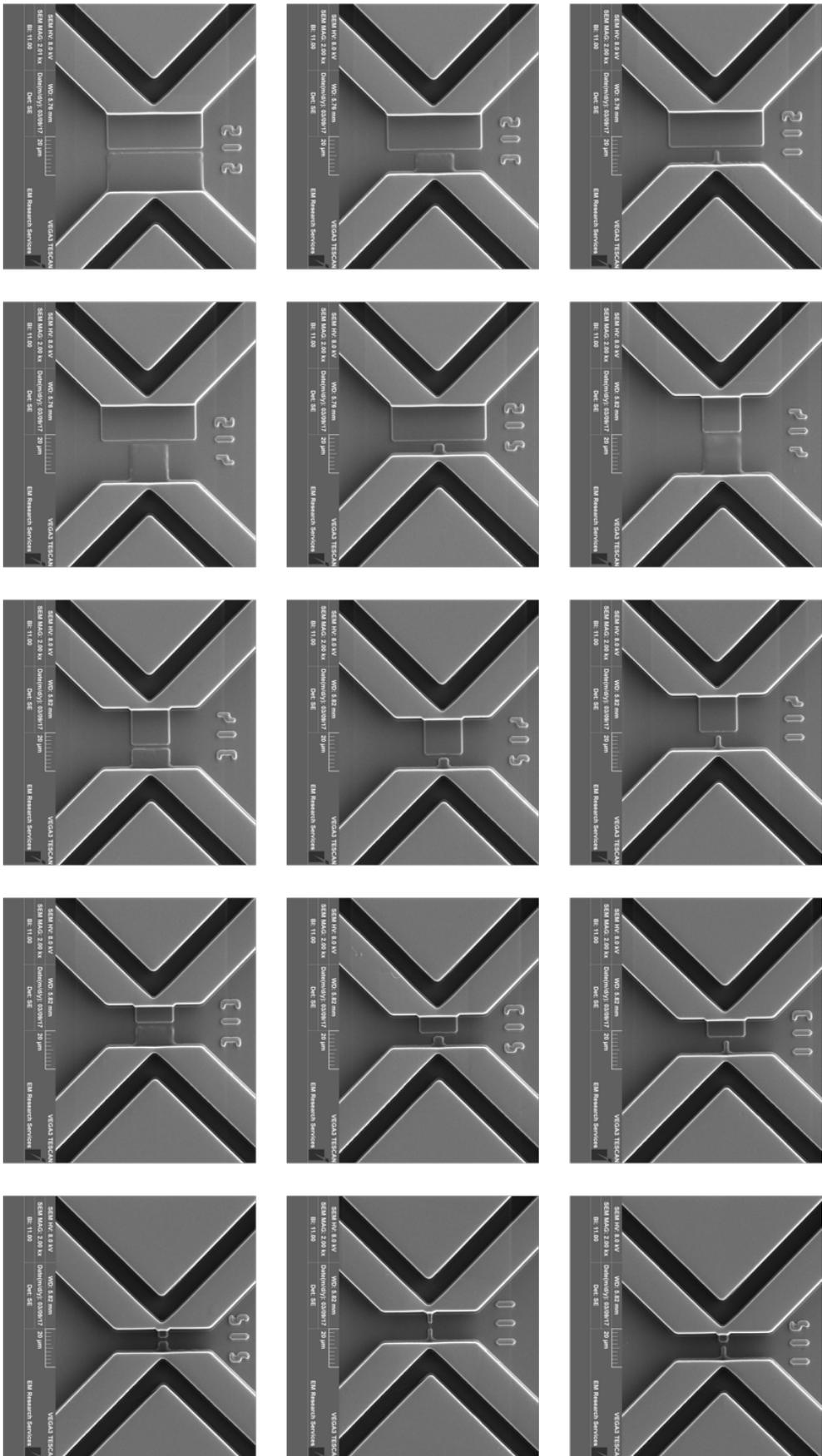


FIGURE 5.20: Microfluidics chip CWqs21 SEM overview. This chip displays the smallest diffusion distance between two separate bacterial colonies ($1\mu\text{m}$).

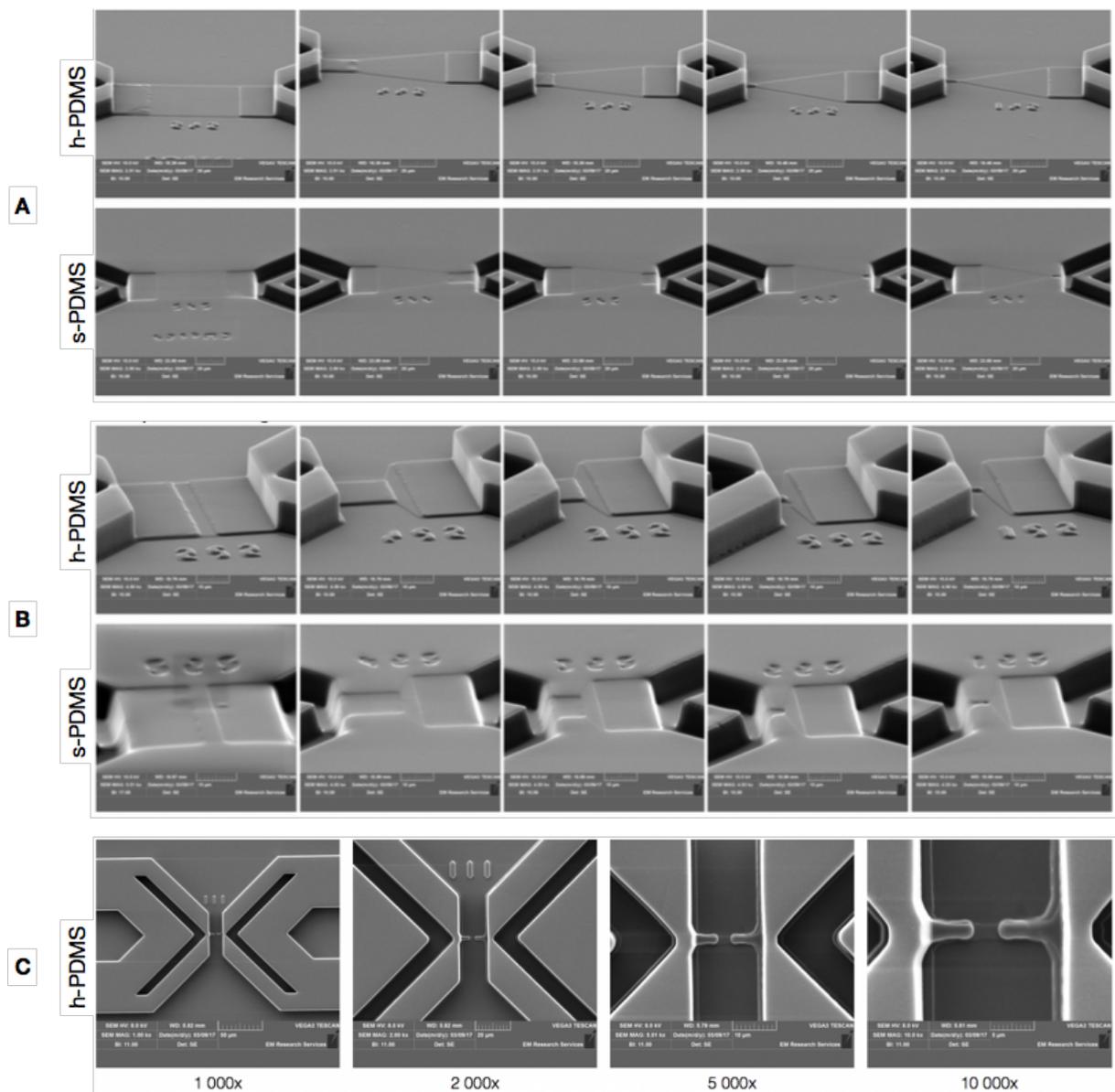


FIGURE 5.21: Overview of surface biochip microfeatures by SEM. In (A) and (B), each column compares features reproduction on hard PDMS (intermediate mould, h-PDMS) and soft PDMS (single use chips, s-PDMS). The largest diffusion distance separating potential bacterial colonies ($100\mu\text{m}$) is presented in (A) while a $5\mu\text{m}$ diffusion plane is shown in panel (B). In (C), we display different magnifications of the smallest microfeature present in any chip: 2 single cells communicating via a $1\mu\text{m}$ diffusion channel.

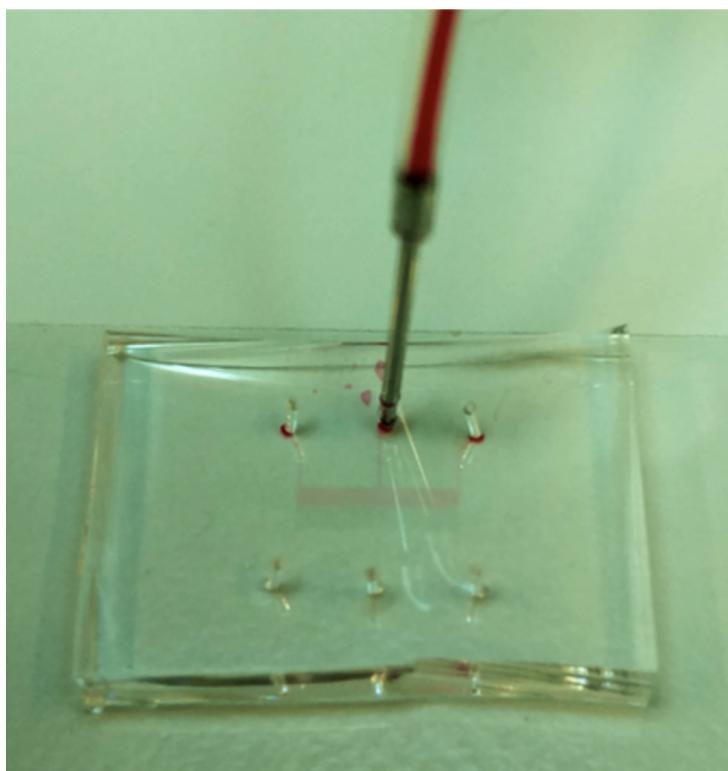


FIGURE 5.22: Flow dynamics test of a surface-type biochip with a coloured compound. Input flow from a syringe plugged into the top middle port splits symmetrically into left-/right-hand part of the microfluidics design (biological replicas). The flow is contained within the top half of the chip and separated from bottom replicas by a narrow diffusion surface.

diluted red food dye into input ports of surface chips and observed the expected behaviour of the liquid dynamics. As it is displayed in Figure 5.22, theoretical simulations were supported by the colouring of only half of the biochip. The other half, that corresponded to a potential separate bacterial strain, was well isolated by a small diffusion surface that did not allow the flow from the opposite side to leak. Together with this final verification step and PDMS chips imaging by SEM, we proved that we could replicate submicron size microfluidics features accurately and obtained identical flow behaviour as simulated *in silico*. In the next section, we reveal *in vivo* characteristics of surface-type microfluidics chips.

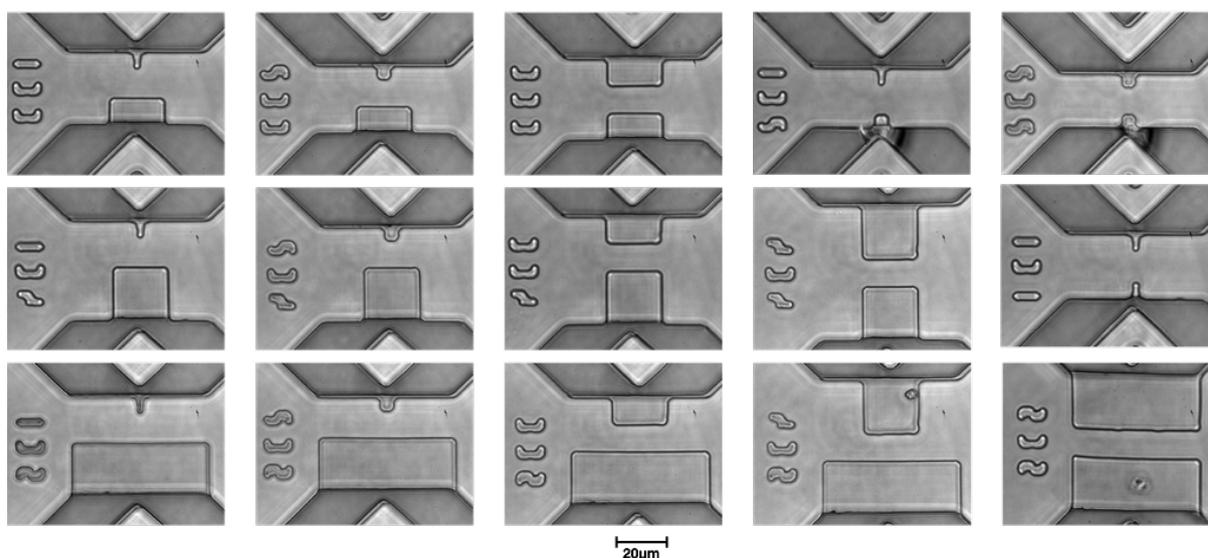


FIGURE 5.23: Microfeatures of surface biochips immersed in LB medium under an inverted microscope. Soft PDMS was bonded to a coverslip, plugged at input/output ports and observed in microscopy before applying a flow of cells.

5.4.3 *In vivo* cellular fitness

5.4.3.1 Cellular loading

After a similar plasma bonding protocol as the one used for channel-type chips, we fixed microfluidics chips to thin borosilicate glass coverslips and mounted the resulting microfluidics apparatus onto fluorescence microscopes. According to the standard microfluidics cellular loading technique, we grew cultures overnight, refreshed them in the morning and injected cells extracted from exponential phase into surface-type biochips. Before loading cells, we immersed PDMS devices into culture medium in order to reduce the formation of eventual air bubbles within the circuit. Figure 5.23 provides an overview of the bacterial traps present in a single microfluidics chips. In standard Nikon-Ti inverted fluorescence microscopes, we could not always obtain a precise focus of diffusion surfaces by vertical alignment. However, strong unequal flow from either part of the microfluidics input ports always showed medium leakage into the opposite side patterns. This meant that diffusion planes may not have been visible under fluorescence microscopes, but were still present in the final, single use chips. After correct immersion of all microstructures, which may take up to 5 minutes without an external pump, we simply swapped culture medium reservoirs for specific

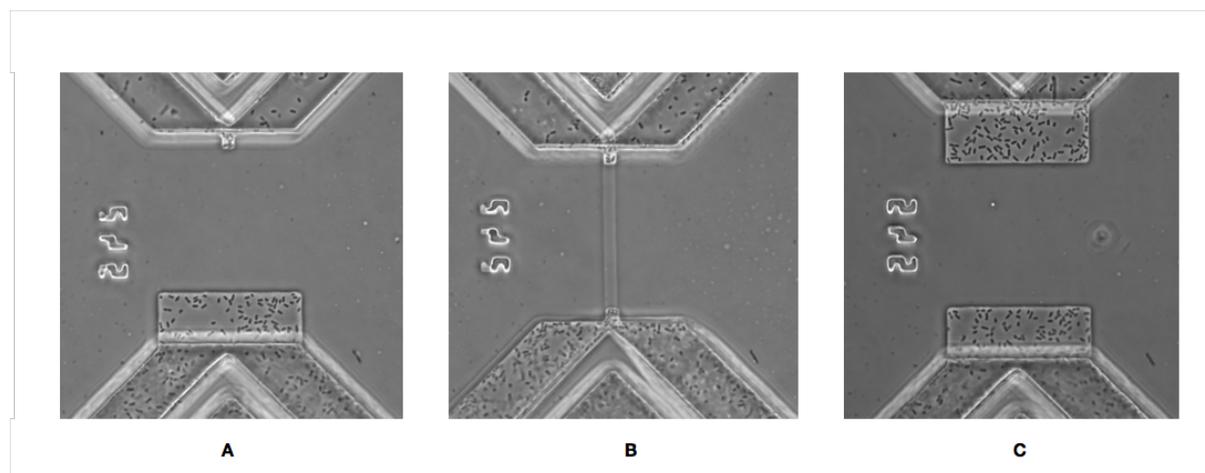


FIGURE 5.24: Surface biochips bilateral cellular loading. Two *E. coli* cellular reservoirs were plugged to a surface diffusion microfluidics chip and let to flow in circuit channels without external pressure. (A), (B) and (C) represent top views of different intercellular communication features, where colonies are separated by a distance of $50\mu m$. (A) displays the combination of a small and a large colony tested together, (B) two small and (C) two large bacterial colonies. Diffusion surface is visible in (B) as a channel connecting the two growth chambers.

E. coli strains. With this method, contrary to the channel-type biochips, we consistently obtained robust cellular loading in terms of number of cells naturally trapped in observation chambers. Here, we mean that cells would directly enter observation chambers without any external stimuli or change in the syringe input/output system. This showed some contrast with the loading of channel-type biochips, where cells first needed to be starved to reduce their size before loading, after what the input medium would have to be changed to allow for the recovery of standard-size *E. coli*. Noticeably, surface biochips that displayed a model of higher flow near bacterial traps were very efficient at hindering cells within observation chambers. As shown in Figure 5.24, bacterial traps of different sizes were easily filled by cells. This setup allowed us to start time-lapse experiments in a favourable context, where a good starter amount of cells could be symmetrically provided to all observation chambers. Overall, these results supported the correctness of our theoretical model and proved the adequacy of surface diffusion microfluidics chips for bilateral cellular loading. We deemed cellular loading sufficient to pursue time-lapse experiments when cells would be hindered within the $1\mu m$ observation chambers and would multiply over 3 successive divisions, whilst remaining stuck in the bacterial trap.

5.4.3.2 Bacterial growth stability over time

From the setup performed by a robust cellular loading process, we grew cells hindered in observation chambers over 12h periods in continuous culture at 37°C . Over the first 2h of the assays, we observed small observation chambers gradually emptying themselves. This was likely due to the high pressure of the nearby flow channel that displaced cells under observation. However, in microcolonies growing in chambers of $10\mu\text{m} \times 20\mu\text{m}$ or larger, we could follow the growth of bacterial species over time. As shown in Figure 5.25, three different situations generally emerged after following continuous growth of bacteria in surface-type biochips: (i) reasonable bacterial growth filling bacterial traps as expected from the design (Figure 5.25A), (ii) abnormalities in design features developed around higher cell density regions (Figure 5.25B) and (iii) clogging effect caused by PDMS impurities and cellular aggregates (Figure 5.25C). In a suitable setup for bacterial growth, discrete final time-points revealed the correct activation of the bacterial genetic circuits tested in this study. However, the most recurrent situation was the emergence of PDMS feature failure, starting by diffusion channels colonisation by bacteria. Although we demonstrated the correct dimensions of surface-type microfluidics earlier in this chapter, wild-type laboratory *E. coli* appears to be able to significantly shrink its size to enter 200nm deep surfaces with the help of PDMS flexibility. Since PDMS is an elastic polymer, we cannot exclude that it may have bent under cellular pressure, eventually allowing cells to create an additional and undesirable growth layer. However, it is difficult to quantify this effect at a sub-micron scale. In all instances of diffusion surfaces overtaken by bacterial invasion, which would take place after 2-3h of culture, we observed the detachment of PDMS from the glass-coverslip over a variable size area. This abnormality could have differential impact over the microfluidics features and - in isolated cases - may create a break between respective flow channels, thus cross-contaminating samples. Nevertheless, in most cases, only the specific feature was affected and rendered unusable for experimental measurements. Also, as displayed in Figure 5.25C, PDMS residues sometimes ended near bacterial traps and triggered clogging of high-speed biochips (with only $3\mu\text{m}$ of flow channel width). Therefore, due to these different detrimental effects, it was difficult to follow continuous growth over long periods of time.

Altogether, the model of surface biochips we developed in this study was successful in short-time assays, but inadequate for longer time course experiments due to design mechanical constraints. As we describe in the following section, these constraints were partly inherent from the photolithography resolution that was used for the construction of microfluidics devices.

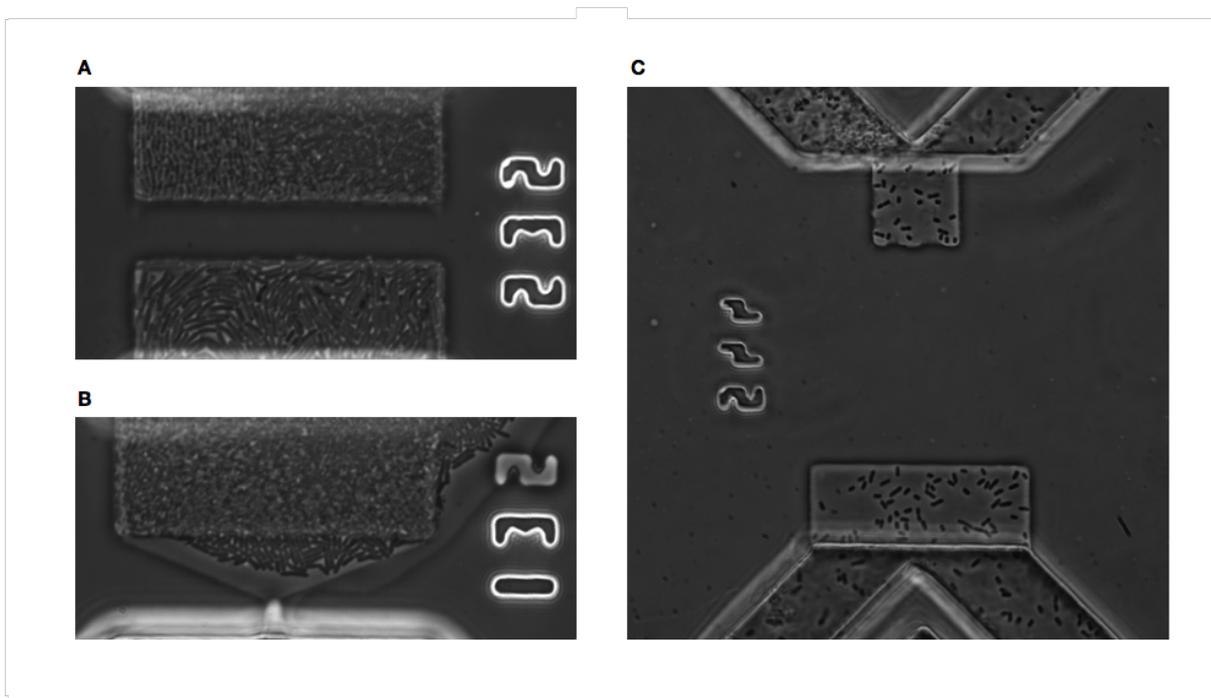


FIGURE 5.25: Potential situations following continuous bacterial growth in surface-type biochips. (A) describes an adequate bacterial growth, (B) colonisation of the diffusion surface by cells and (C) a clogging effect (top left) due to residual organic matter stuck in a flow channel.

5.5 Intercellular communication in microfluidics

Microfluidics provide a platform for the setup of biological high-throughput assays. In this chapter, we presented the set of microfluidics chips that were developed in this thesis. Designing microfluidics chips is relatively straightforward but should always be bespoke to the requirements of the specific application. Engineering the right microfluidics circuit is a cyclic process where biochips are in turn designed, simulated, created and modified to improve their fit to application-specific needs. In this study, we engineered microfluidics biochips to study combinatorial events taking place in microscopic bacterial windows. Here, we discuss the limiting factors that were observed for the study of microscale intercellular communication via genetic devices engineered in Chapter 4, and further circuits that may be developed as expansions of the current microfluidics design.

5.5.1 Microscale reaction-diffusion systems

5.5.1.1 PDMS device fabrication

We built microfluidics devices to study individual communicating colonies of bacteria at the microscale. For these means, we designed observation chambers to trap bioengineered cells at distal locations, connected by diffusion areas. For the study of *E. coli*, microscopic dimensions of these circuits first required the optimisation of the PDMS device fabrication process (see Appendix D). If microfluidics chips are tailored to the observation of single cells and display hollow features, a clean and robust replication of the microfluidics wafer patterns is essential to obtain usable devices. Although there is usually no need to perform PDMS chips fabrication in a clean room, this appeared to be critical step to limit the amount of microscopic dust that could block chip compartments after bonding to glass coverslip. Due to the size of features on SU8-photoresists, we also had to make a thick hard-PDMS layer to enable the intermediary wafer to accurately replicate features into soft-PDMS chips. In consequence, detaching PDMS from master wafer became more challenging after a simple silanisation step[65]. We thus needed to optimise the incubation of the master wafer by sealing it in a vacuum chamber under high pressure (27psi) in the presence of the silanisation agent. We screened appropriate replication of PDMS features onto intermediate and final single-use chips by SEM, and provided a thorough explanation of the PDMS chips molding process (Appendix D). With these protocols, we demonstrated very accurate reproduction of submicron-size microfluidics features, and could ensure that a specific design was correctly imprinted before use in biological assays.

5.5.1.2 *E. coli* growth in observation chambers

Microfluidics allow the characterisation of genetic circuits in bacteria via two main functions. First, cells are immobilised in specific compartments where they can be followed over time. Then, a dynamic control can be applied over fluid flow to provide appropriate growth conditions for the studied cells[65]. In practice, this is achieved by creating multiple channel depths over a uniform PDMS surface, and by plugging external devices such as input syringes and waste outline into a microfluidics chip. Altogether, the wiring of such device results in a well-controlled environment for bacterial proliferation and the study of specific physiological behaviour. However, there are multiple factors that can impact cellular behaviour in microfluidics settings, and we provide in the following an overview of the key elements that we observed during continuous growth of *E. coli* genetic circuits.

We presented above two types of biochips: channel and surface diffusion devices. As their name suggest, they differed in the process of diffusion for intercellular communication. First, we tested channel-type biochips that had been developed by Dr Sunny Park and connected individual bacterial strains by $0.7\mu\text{m}$ size channels. Later, we manufactured and reproduced a variant design of channel biochips with thinner diffusion surfaces (200nm depth) instead of channels between microcolonies of bacteria. These two designs of microfluidics chips, as described in Figure 5.4 and produced via the protocols listed in Appendix D, were tested *in vivo* and showed a much easier cellular loading for the second type of surface biochips. One of the main design differences between these chips was the way cells flowed near bacterial traps. While channel-type chips featured straight flow channels, we optimised the amount of cells naturally trapped in observation chambers by altering these and actively pushing bacteria into $1\mu\text{m}$ depth compartments. Nevertheless, the latter caused clogging issues to arise in a couple of instances[50]. In nature, cells tend to adhere to surfaces and to produce different chemicals that help the setup and development of microcolonies. For instance, this is what happens during the creation of biofilms[175]. In microfluidics, a single cell adhering to a PDMS wall can replicate every 25min and manage to attract and attach itself to further flowing cells. Over a few hours of experiment, this effect can be multiplied and completely block some parts of the flow channels. In Chapter 5, we presented two types of surface diffusion biochips where cells were loaded by either a $3\mu\text{m}$ or a $10\mu\text{m}$ wide flow channel. In practice, we observed that a narrower distance ($3\mu\text{m}$) performed better at loading cells in bacterial traps, whilst this made chips also more likely to develop clogging of cells combined with PDMS residue (cf. Figure 5.25). On the contrary, $10\mu\text{m}$ features never displayed cellular clogging but only loaded half

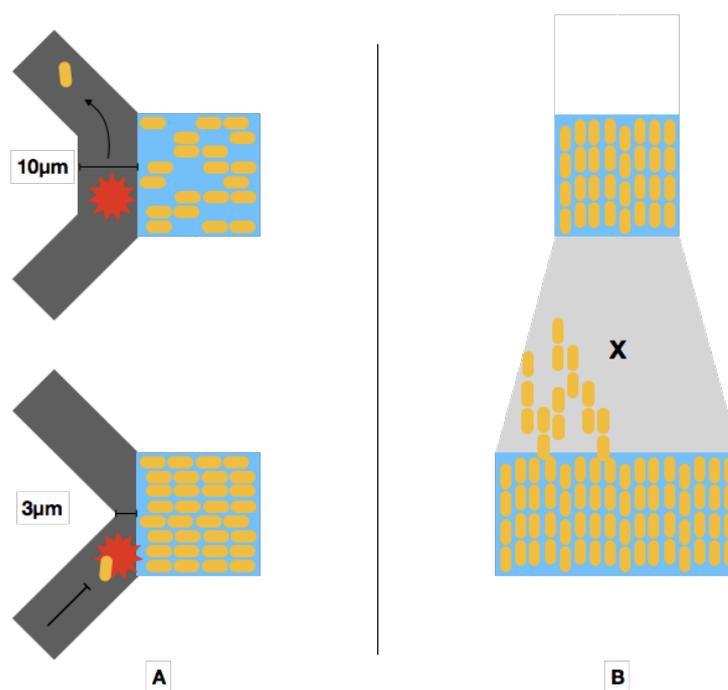


FIGURE 5.26: Surface diffusion microfluidics chips limitations with *E. coli* cultures. In all panels, deep flow channels are shown in dark grey, observation chambers in blue and diffusion area in light grey. Individual yellow rods represent bacteria in different settings. (A) shows the possibility of clogging in two type of biochips. Top and bottom panels shows different distances of flow channels reduction nearby bacterial traps and eventual clogging effect via PDMS residue (red star). (B) displays colonisation of a surface diffusion area ($X = 200nm$) by individual cells over growth.

the amount of cells attainable with the alternative design. This process is outlined in Figure 5.26A with clogging of the $3\mu m$ wide channel. Altogether, the design of each microfluidics chip should be considered for the optimisation of multiple factors, and allowing for model variations greatly helps troubleshooting specific issues during *in vivo* studies.

In surface diffusion microfluidics design variations, we included the modification of observation chamber aperture size. In channel-type chips, we overcame poor cellular loading issues by growing *E. coli* cells in minimal medium, and swapping back to a richer medium after biochip loading. However, as a consequence of the design, this model of bacterial traps did not display a good fitness for growing microcolonies of bacteria. In these settings, observation chambers did not allow for easy removal of the excess cells[65, 50]. Besides a stacking effect observed in a theoretically two-dimensional environment, *E. coli* cells also regularly developed mutations that significantly affected their morphology and underlying physiological processes. Therefore,

as suggested by the literature and as we observed in surface-type biochips[65], we report that providing a larger aperture to observation chambers is a strong determinant in bacterial growth fitness over time. In general, multiple design considerations should be accounted for in the optimisation of microfluidics circuits, and these should be carefully reviewed for potential implications with the culture of bacterial cells.

5.5.1.3 Diffusion system limitations

We have previously demonstrated how the setup of surface-type microfluidics chips helped the cellular loading and culture of *E. coli* cells. In the biological context of this study, we investigated diffusion of quorum signals between specific bioengineered strains presented in Chapter 4. Although computational and wet lab studies have characterised the wiring of bacteria by quorum compounds at macroscopic distances, there is little known about microscopic interactions and how autoinducers act at a smaller scale[223, 47, 152, 206]. In our microfluidics chips, we designed the thinnest achievable features by soft-photolithography to represent planar diffusion surfaces between two individual strains. While cells were anchored for growth in a $1\mu\text{m}$ depth surface area, excess bacteria were extracted from observation chambers via flow channels. Since channel-type biochips revealed that $0.7\mu\text{m}$ were insufficient to retain *E. coli* from entering diffusion surfaces, we designed these planes as 200nm to minimise bacterial entry into theoretically cell-free compartments. Whilst this approach appeared to accommodate low to medium cellular densities, the elasticity of PDMS and the increase in pressure exerted by higher cell density allowed colonisation of 200nm diffusion areas after a few hours of experiment (cf. Figure 5.26B). At the current time, the development of micro-/nanoengineering techniques accelerates the study of biological samples in large scale screening assays. However, soft-photolithography did not provide a suitable platform for the study of bacterial communication endeavoured in this study. For the development of further microscale intercellular communication devices, we would likely consider photolithography alternatives such as E-beam lithography that can achieve nanometer-scale features[5, 62]. This manufacturing process increases the resolution obtained compared to standard photolithography but nevertheless multiplies the costs of wafer production.

5.5.2 Wiring and controlling bacterial communication

Quorum sensing systems are a landmark for the study of bacterial intercellular communication[155, 225]. They perfectly mimic the process of reaction-diffusion observed

in systems we studied in Chapters 4, and we generated synthetic constructs that allowed the specialisation of certain strains to either produce, sense or amplify environmental quorum molecules. So far, we have described the microfluidics results obtained for a simple sender-receiver system, and demonstrated the production and sensing of quorum signal in *E. coli*. However, the propagation of chemical waves within or between bacterial colonies naturally follows rules that can be distinguishable at a higher scale[227]. In the following subsections, we describe potential microfluidics applications that may use more complex genetic circuits that were presented in Chapter 4.

5.5.2.1 Quorum signal repression

In nature, quorum sensing is a mean of synchronisation across bacterial species. Quorum-controlled behaviour is usually observed at high cell densities and depends on local concentrations of an autoinducer. A positive feedback loop normally amplifies the production of quorum compounds, and cells undergo different subsequent physiological changes. In this study, we used a promoter fusion between the activable P_{R-lux} promoter and the cI repressible binding site known as $P_{RcI-lux}$ [48]. A cI binding site was placed between lux operators, preventing luxR to dimerise and activate the transcription of P_{R-lux} in the presence of AHL and the cI protein at the same time. In this context, environmental cI protein is considered as an inhibitor of the quorum circuit. This means that, regardless of the presence of a quorum signal, one could reset the state of a sensor strain. For the study of computation with bacterial cells, this sort of feature allows one to obtain spatiotemporal control over growing cells, and to observe dynamic physiological behaviour in time-lapse experiments[56, 15, 220, 164, 193, 137]. Together with a repressor system, one may use the quorum strains engineered in Chapter 4 and attempt to adapt the ON/OFF quorum response of sensor devices to oscillatory circuits.

5.5.2.2 Pulse generator devices

Pulse generators are molecular devices that express a signal molecule to trigger cyclic or unidirectional regulation of complementary compounds. In both regulatory processes, the iterative activation and repression of biological units can transduce an input signal and be identified via the formation of a specific pattern over macroscopic distances[16, 46, 4]. In Figure 5.27A, we simulated the behaviour of a quorum circuit split between individual bacterial strains that, in turn, produced, sensed, amplified

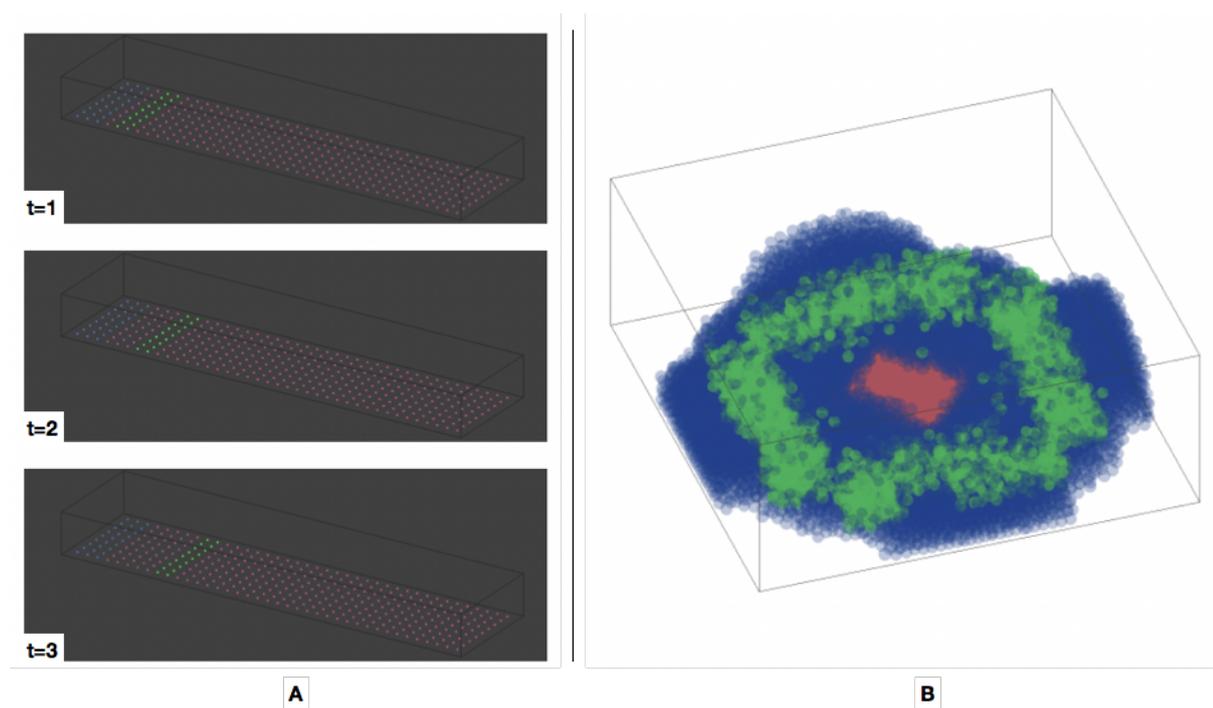


FIGURE 5.27: Simbiotics simulation results of a pulse generator device. In both models, cells are represented as individual coloured circular entities growing in the grey cage outline. (A) displays three different time points after inducer production (blue) in microfluidics settings (e.g. two-dimensional growth). Pink cells are inactive and green cells amplify quorum signal. (B) presents cells producing a quorum molecule (red) or sensing this signal (blue and green) in a three-dimensional space allowing stacking effect. Blue cells are repressed and inactive while green cells are activated for quorum production.

and repressed signalling molecule detection. This snapshot represents the physiological state of different colonies after bacterial growth in static conditions, and is concordant with other *in vivo* models that have been explored for quorum devices in the literature[16, 15, 203]. To simulate growth from separate bacterial species, we used a three-dimensional simulation platform developed by Jonathan Naylor called Simbiotics (<http://simbiotics.org>)[169]. Using the same approach, we also simulated the behaviour of these circuits in microfluidics-like conditions, where single layers of cells should be placed and grown in relative proximity from one another (cf. Figure 5.27B). While the real complexity for the study of intercellular communication in microfluidics relies on maintaining bacterial cells controlled in *in vivo* conditions, this type of *in silico* work greatly helps the setup and design of potential circuits.

5.6 Summary

In this chapter, we described the theoretical design and fabrication process of microfluidics chips for high-throughput characterisation of intercellular communication. Firstly, we optimised PDMS devices making for micron or submicron size features replication. Then, we developed different microfluidics circuits to improve cellular growth in these settings. Based on the genetic devices engineered in Chapter 4, we observed signal transduction (*e.g.* production/sensing of quorum molecule) between microcolonies of bacteria distal of $20\mu\text{m}$ to $100\mu\text{m}$, but it was impossible to gather significant experimental data from small cellular compartments due to limitations of PDMS devices themselves (*cf.* Figures 5.13 and 5.25). In the fabrication of silicon wafers, soft-photolithography allows the setup of small structures but has a resolution limited to a few hundred nanometres. Therefore, physical constraints are imposed on the design of specific circuits and while we demonstrated correct features replications in PDMS, *in vivo* conditions may differentially affect microfeatures through pressure drops and high cellular densities. Here, we tested the smallest achievable features that could be printed by standard soft photolithography, and we showed that these were inefficient at keeping well-controlled growth conditions intact over long-term continuous culture. The development of high-throughput approaches in microfluidics was shown to help the screening and understanding process of cellular behaviour. Although the tiniest features were unstable, it would remain interesting to detect bacterial communication events at the single-cell level. For instance, this could provide some valuable insights into how bacteria may differently communicate when they are alone or surrounded by other cells. In the next chapter, we integrate the results and discuss the data that were presented in Chapters 3, 4 and 5.

Chapter 6

Integrating bioengineering approaches in synthetic biology

In this thesis, we have reviewed and explored a few of the basic elements of bioengineering in synthetic biology. In this respect, we presented the setup of a DNA encoding method to identify and track bacteria via DNA barcodes. Our work focussed on building the abiological sequences and in their integration in living organisms. Chapter 3 provides details about barcode stability *in vivo*. We engineered and characterised a library of genetic circuits that could be used as model for barcoded strain documentation profiles (Chapter 4). In Chapter 5, we show the potential for microfluidics to scale up and optimise genetic circuits single-cell characterisation. At the submission time of this thesis, we are in the process of developing a spin-off platform about DNA barcoding with the help of Newcastle University. Here, we discuss some key aspects of the barcoding platform, and how they link the different parts of this study.

6.1 Bac2code online barcoding platform

Synthetic biology relies on using specific biological devices to perform some given action in an *in vitro* or *in vivo* context. For both types of work, the construction of tailored genetic circuits and their associated wet lab documentation is essential[29]. In order to facilitate interdisciplinary approaches complementing wet lab resources with *in silico* studies, we created and presented in Chapter 3 a barcoding platform called Bac2code. In this study, we designed DNA barcodes for the development of a strain repository that would allow the browsing of bacterial strain profiles and their associated *in vitro/vivo* and *in silico* documentation. Dr. Jurek Kozyra, who encoded the hexadecimal to DNA conversion algorithm, is developing this online repository. The following explains the prototype of such online barcoding platform.

6.1.1 Possible actions on the Bac2code website

First, possible actions on the online barcoding platform are defined by the type of user. The Bac2code system targets three main type of users with specific needs, as summarised in Figure 6.1. At the write-up time of this thesis, the main type of users of the system would be categorised as "Expert users", since the barcoding platform started as an empty repository with no strain to browse. These users are mostly interested in populating new barcoded strain profiles with specific documentation. This can be performed in a similar manner as the creation of wiki pages[109], or by simple document upload as outlined by Figure 6.2. Usually, users have a working and documented strain that needs barcoding (*e.g.* commit changes to a strain repository), and the Bac2code website provides a unique DNA barcode to track this strain. After following the barcoding procedure outlined in Appendix C, users can insert a new barcode or update a strain already barcoded with the newly generated DNA identifier. The experimental procedure to barcode a new strain or to update a barcode is identical. Although users may barcode all their strains, we envision that only "milestone strains" (*e.g.* final versions of working and documented strains) will be recorded in the Bac2code system. Finally, users publish their strain documentation on the Bac2code website and choose to make this strain profile public, private or protected by access restricted to a limited set of users.

A main feature of the Bac2code website, once populated with a community of users and a large number of barcoded strains, is to provide an easy way for retrieval of barcoded strain profiles. As shown in Figure 6.3, users are offered the possibility to retrieve specific barcoded strain content. This can be done either via direct DNA sequence input or by the upload of a sequencing read file targeted by the universal primer. The Bac2code website then finds a barcode in our database that matches the user input, and directs this user towards the specific strain profile.

Finally, our barcoding platform also offers the possibility to pre-order a batch of DNA barcodes for further strain development. The case explained in Figure 6.4 shows that users could order several DNA barcodes at once, possibly for strains that are yet to be constructed and documented. The Bac2code system thereby allows for the creation of blank barcode pages that can be edited to prepare the development of genetic circuits in research projects. As detailed above, this may be done privately or in collaboration with other users (mainly academic colleagues) that could view and/or edit specific strain profiles. To this extent, we aim to develop the Bac2code system as a version control and tracking tool for barcoded strains.

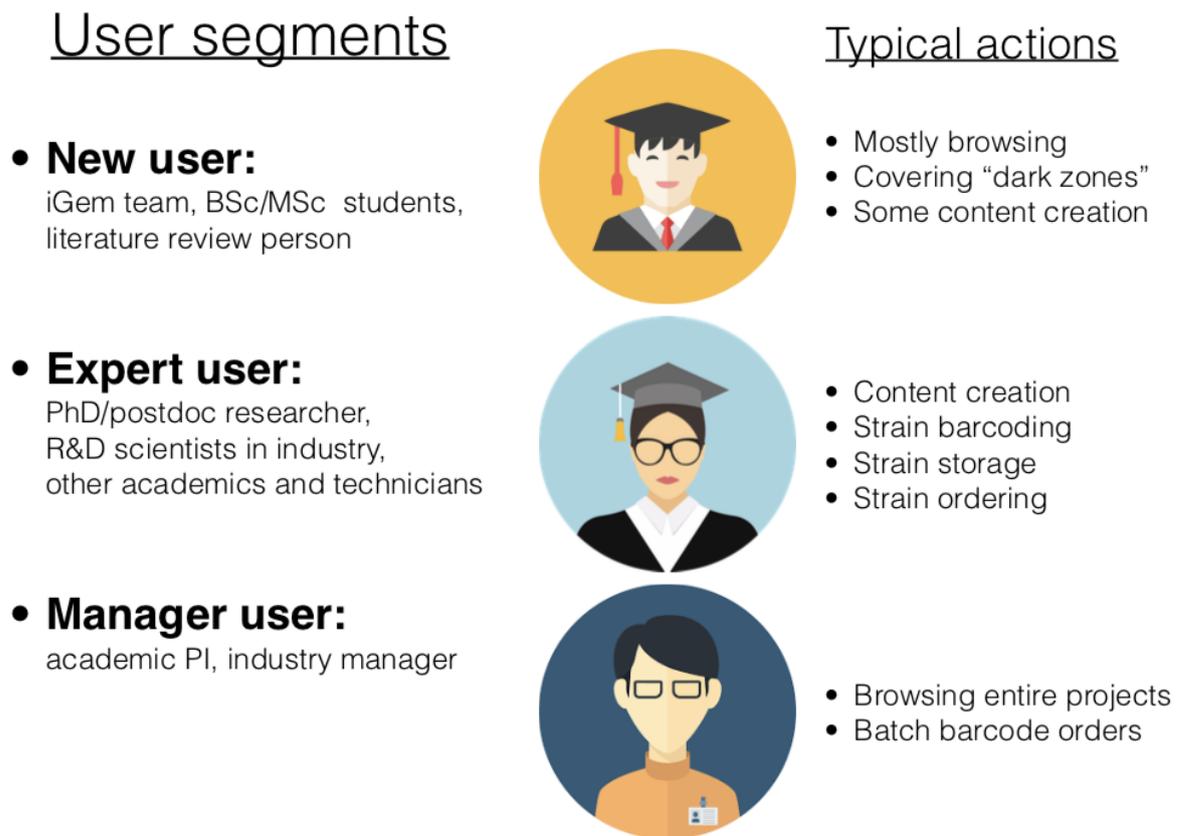


FIGURE 6.1: The Bac2code platform targets three types of users (left panel) that often perform a similar type of action on the website. New users generally browse only, and explore "dark zones", or specific details, that may be associated with specific strains. Expert and supervisory users have others needs that need to be accounted for in the possible actions they can perform.

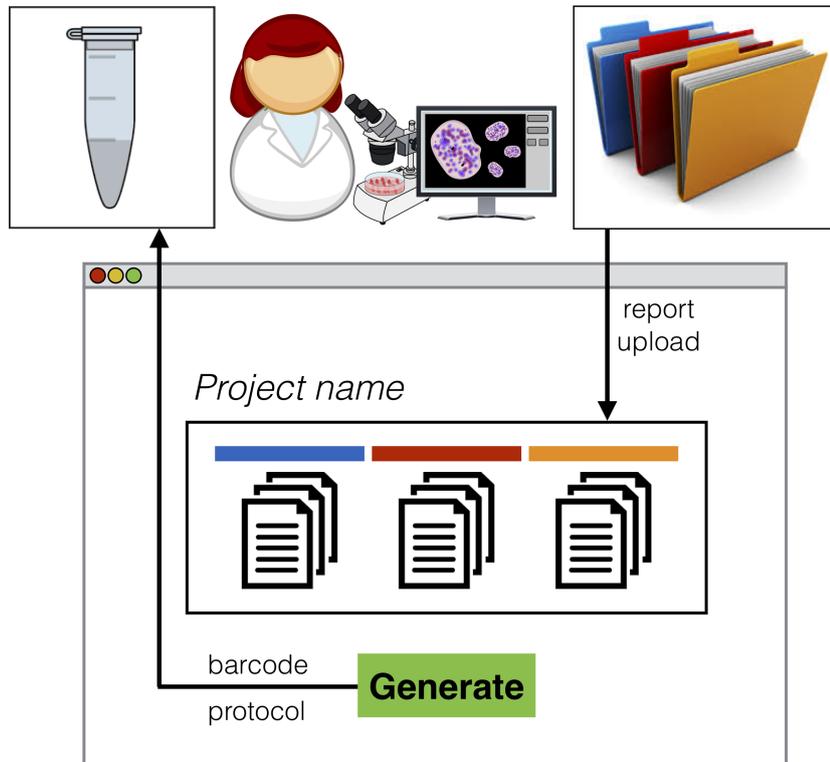


FIGURE 6.2: Barcoding with Bac2code: users can upload a set of documentation to fill in a blank barcoded strain profile that is linked to a specific DNA barcode. This barcode is automatically generated by the online platform and provided to users, who can then barcode strains using protocols listed in Appendix C.

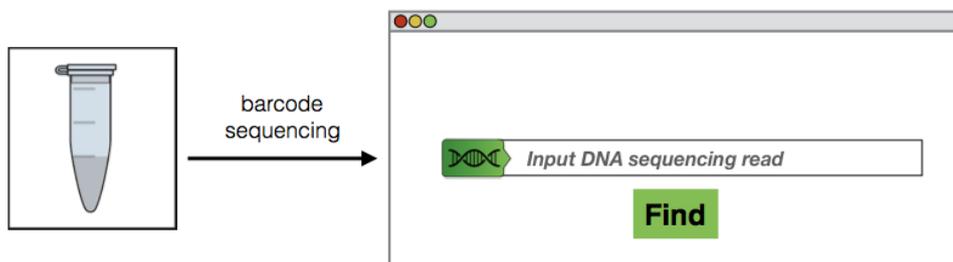


FIGURE 6.3: Retrieval of barcoded strain profile on Bac2code. Users can retrieve any barcoded strain profile given a sequencing read used to identify a specific barcode sequence via a universal primer.

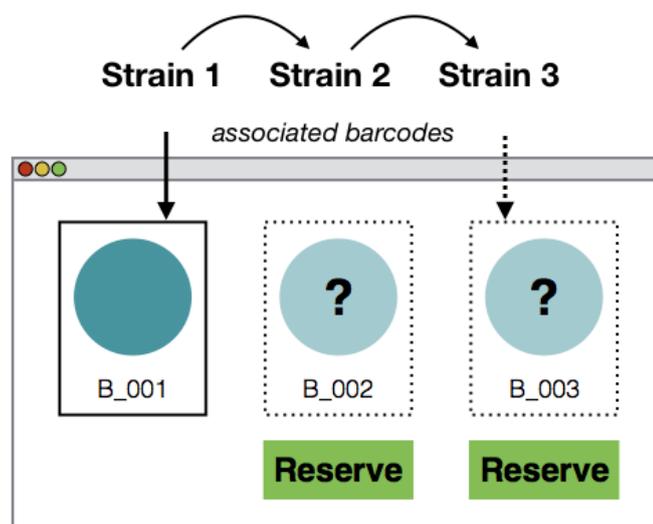


FIGURE 6.4: Possibility to order and reserve DNA barcodes prior to the construction of bacterial strains. Here, strains 1, 2 and 3 are respectively meant to be associated with barcodes B_001, B_002 and B_003.

6.1.2 Integration of *in silico* resources

In synthetic biology, large consortia of scientists have adopted the use of standards for bacterial strain engineering and documentation [149, 73, 123, 29]. By creating the Bac2code platform, we do not endeavour to simply create an additional standard. Instead, we aim at grouping large amounts of *in silico* data based on individual platforms onto a single hub, and to integrate computational studies that help genetic device characterisation. In particular for the setup of genetic libraries, such a platform would facilitate the handling of computational models, gather available information from external resources and create strain variants automatically, which currently is a tedious process for most large scale studies. With *in vitro/in vivo* data users should provide for each barcoded strain, the integration of the Bac2code platform with other online resources is key to building a system that may truly bring together experimental and computational biological studies.

6.1.3 Integration of barcoded strain specifications

We described above what we intend the Bac2code platform to be, and this includes a collection of experimental data to consult when browsing strain specifications. The data we first gathered to populate the Bac2code server was extracted from biological studies documented in Chapter 4. These were focussed on the construction and

characterisation of a library of genetic circuits. For the characterisation of these devices *in vivo*, we followed bioengineered bacteria over time, built mathematical models based on the Hill function and simulated fluorescence behaviour for individual constructs[46]. We tested slightly more complex models with additional factors affecting the Hill function such as molecule production and degradation rates, but did not observe any significant improvement in fitness between experimental data and computational model. Compared to the genetic devices initially developed, optimised constructs showed an improved fitness with Hill predictions, suggesting that simple models were good at explaining reaction-diffusion systems (cf. Figure 4.27). For the assessment of best synthetic constructs, the matching between the theoretical model and experimental measurements was a main factor of comparison. Altogether, the experimental data gathered from these studies represented a great starting point to be uploaded onto Bac2code strain profiles, where complete datasets may then be further explored to develop circuit derivatives. By building such repository of strains, we can exemplify the progression of a research project, here by the optimisation of genetic construct behaviour in response to quorum sensing signals.

The Bac2code platform is being developed to ease the transfer of information and to improve communication between researchers. We propose a system where all experimental and computational data related to a bacterial strain may be gathered in a single place. Since the amount of information provided when exchanging strains between laboratories is often very little, this would facilitate and favour the exchange of genetic parts and data in the research community. Here, we demonstrated the setup of cloning methods for the barcoding and update of *E. coli* and *B. subtilis* DNA barcodes, but our method is being developed to accommodate barcoding of a larger number of species. Although we presented barcodes encoded on the bacterial chromosome, the Bac2code online platform may be adapted to barcode plasmid systems. In this instance, barcodes could be specifically assigned to background strain or genetic construct (barcode located on plasmid directly), and a barcoded strain profile may be associated with a combination of barcodes. However, such work is beyond the scope of this thesis, and these next challenges are being undertaken by Jonathan Tellechea.

6.2 Barcoding in microfluidics

In Chapters 4 and 5, we have discussed the different methods and results obtained for the characterisation of genetic circuits, and the application of microfluidics to reaction-diffusion systems that were built. As detailed in the previous section, data obtained from these experiments can be used to document barcoded strain profiles on the Bac2code server. Here, we propose how barcoding may be taken to the next step in terms of industry, *e.g.* how to perform it in high-throughput settings. This strategy is based on the use of microfluidics for large scale processing of individual barcoded species.

In this thesis, we have demonstrated the perfect stability of barcode sequences over hundreds of bacterial generations in different experimental conditions. For these means, we used high-throughput methods that consumed relatively large amounts of culture medium, and needed external methods to extract DNA and to sequence barcode fragments. Towards the automation of bacterial strain handling, we are working on the development of a microfluidics platform that could operate two types of functions: (i) the barcoding of new bacterial strain and (ii) DNA barcode retrieval. Currently, there are three steps to barcoding cells: the assembly of recombinant DNA, bacterial transformation and the selection cassette loopout. In practice, biological techniques used to perform these actions are mainly based on temperature cycling and growth of bacterial strains. A number of applications have been developed to automate molecular cloning in microfluidics, and the type of operations that are required to barcode cells could also be adapted to the microscale[89, 88, 211, 255]. In fact, this may be achieved via a microfluidics platform that would allow region-specific temperature control for the assembly of recombinant DNA/culture of bacterial cells, and programmable methods for *on chip* transformation. Likewise, a system for barcode DNA retrieval may also be developed and tailored to grow cells, extract gDNA and perform PCR on barcode target regions. Parallelisation of processes can be easily achieved in microfluidics and, therefore, such platforms would provide autonomous equipment that would be suitable for cellular library barcoding and large scale identification of bacterial species.

Automated digital technologies are a hallmark for the development of biological applications. Synthetic biology uses these by making the fusion between wet laboratory and computational investigations. Nevertheless, the overall progress of multidisciplinary studies can only be achieved by an easy access to relevant *in vitro/vivo* and *in silico* data. Yet to be physically connected, here, we solve this problem with the Bac2code system

and endeavour to use DNA barcodes on integrated platforms that could handle multiple series of bacterial species at one time. Besides helping the screen of genetic libraries, the Bac2code platform is also being built to allow the tracking of individual strain lineage and thus, to display the capabilities of a biological version control software. As was shown for the field of computer science[247, 63], shared repositories are the solution to a smooth orchestration of multiple project strands. With Bac2code, we aim to facilitate strain referencing with an online system, and we aspire to assist collaborative work and the progress of state-of-the-art bacterial studies.

Chapter 7

Concluding remarks and further prospects

In this doctoral thesis, we developed a physical link between bacteria and their associated documentation. We propose a unified platform for the authentication and the tracking of bacterial strains. Using standard bioengineering methods, we constructed a genetic library and used it as a template to demonstrate our system's capabilities. Finally, we showed the optimisation of microfluidics for the study of bacterial cells, and discuss further considerations of this research area, directed towards the development of automated high-throughput barcoding approaches.

Synthetic biology is about optimising the design and function of novel synthetic circuits. In Chapters 2, 3 and 4, we reviewed the methods that were developed to engineer synthetic circuits in this study. In particular, we provided in Chapter 3 a thorough walkthrough of the design process of DNA barcodes, their *in vitro* assembly and *in vivo* integration in two model bacterial species. We then explained the retrieval process of DNA barcode information, and provided evidence for the *in vivo* stability of these artificial sequences. Altogether, we showed the setup of a universal cloning method that uses analogous mechanisms for different bacterial strains, how to engineer bioorthogonal barcodes, and most importantly, how to retrieve barcode information, which provides access to a unified computational and experimental resource platform. Based on this, we postulated in Chapter 6 about further developments in order to adapt microdevices to the automation of parallel barcoding (insertion of barcode) and serial strain identification (retrieval of barcode). At the submission time of this thesis, we are preparing a patent application for our barcoding platform, and consequently aim to expand to a startup company, with the help of Newcastle University.

Adequately documenting scientific work is of utmost importance to ensure the reproducibility of results. In the development of genetic circuits, novel synthetic devices should be associated with a thorough explanation of their characterisation process, in

order to provide meaningful information. In Chapter 4, we explored the growth and fluorescence profile of a series of bioengineered *E. coli* strains. For the optimisation of specific device behaviour, we used various experimental and computational methods, and demonstrated their usefulness to obtain robust physiological changes over a range of different conditions. In the context of microfluidics, we studied the previously obtained optimised strains at the microscale, and for these means had to improve standard PDMS fabrication protocols for the accurate replication of micron and submicron size features. However, in *in vivo* conditions, we reached the technological limit imposed by microfluidics printing technologies; we observed biochip failure at submicron size domains due to the forces bacterial cells exert in order to proliferate. Yet, this provides hints for the development of further microfluidics circuits, and helps us to understand the need for specialised equipment for the progress of scientific research.

As a final note, although we demonstrated how to bridge the gap between experimental and computational biology via the Bac2code platform, our technology remains at an early stage. In the plethora of available synthetic biology tools, it remains difficult to provide a new unified solution, sometimes overriding others. Therefore, our platform may, one day, be used as a standard laboratory notebook, or it may as well end up as a bacterial strain identifier test in hospitals. At the present time, both scenarios are unlikely to happen; however, we hope that the work in this thesis will eventually help the development of further applications to solve real world problems.

Appendix A

Biological protocols

Synthetic biology aims at the assembly of modular genetic circuits in order to control specific functions. Here, we provide an overview of the experimental methods and the equipment that were used for the assembly and characterisation of the synthetic devices presented in Chapters 2 and 4.

Bacterial culture medium We investigated the growth of bacterial cultures with a range of culture media, including rich and poorer growth alternatives to obtain different growth of *E. coli* and *B. subtilis* species. In general, we mainly used the Luria Bertani rich medium (LB), and favoured Super Optimal Broth (SOB) or Super Optimal broth with Catabolite repression (SOC) richer equivalents of LB for cellular recovery after *E. coli* transformation. For the use of the Agilent *in vitro* mutagenesis kit, we performed transformation recovery in NZY medium (an auto-induction LB medium). For cellular culture in poorer, or more controlled growth media, we used M9 salts derivatives for *E. coli*, in which we varied the carbon sources (glucose vs. glycerol) and casamino acid concentrations. For *B. subtilis*, we used growth media based on the Spizizen minimal medium salts (SMM), and adapted these to minimal growth, or starvation conditions.

Autoclaving Samples that needed sterilisation were autoclaved in Sanyo MLS 3751L autoclaves.

Polymerase chain reaction DNA is a very small molecule. Although we only need small amounts of DNA for cloning, the construction of synthetic devices starts by the amplification of individual fragments aimed to be assembled. This is possible by Polymerase Chain Reaction (PCR), where a template DNA is added to a mix of deoxyribonucleotides and DNA polymerase enzyme, which amplifies fragments flanked by primer sequences through specific temperature cycling. For the correct amplification

of DNA, primers should be oppositely directed, and anneal to the template at a permissive temperature before polymerisation extension. We generally used the NEB Q5, Phusion or Taq polymerases (NEB M0491, M0530 and M0273) for PCR amplification of DNA fragments. Following manufacturer instructions relevant to each polymerase, we performed PCRs in Sensoquest Labcycler gradient and Jena Biometra trio/analytic thermocyclers.

Gel electrophoresis In order to identify different DNA, RNA and protein mixtures, a simple method is to analyse biological samples based on their relative size. Samples are migrated through a gel immersed in a conductor environment, and voltage is applied from one end to separate biological species based on their molecular weight. Individual molecular species travel at a speed inversely related to their length, which makes possible their accurate identification by comparison to an appropriate ladder. A ladder is the gel reference that contains different biological fragments of known sizes[76]. Therefore, we can verify biological sample size, and thereby decide to continue or to repeat a given cloning process.

In this thesis, we used agarose-based electrophoresis techniques for DNA migration. We varied agarose concentrations (Sigma A9539) from 0.5% to 2% according to the expected fragment size. In general, we used lower agarose concentrations to separate low molecular weight species, whilst we increased the agarose concentration to separate larger samples. We mixed the agarose to either Tris-borate EDTA (TBE, ThermoFisher AM9863) or Tris-acetate EDTA (TAE, ThermoFisher AM9869). The buffer that was used to immerse the gel in a tank (ThermoFisher EH15) had to be the same as the one used to make the agarose gel. While TAE was adequate for most gels, it has a tendency to overheat and requires gel tank cleaning more regularly than its TBE equivalent. TBE was thus preferred for longer runs (≥ 30 min), since it does not produce such heat. The physical migration of samples was led by a voltage (70-140V) applied from a negative electrode towards the opposite pole, placed on the other side of the gel containing samples. After migration and staining, for which we used Ethidium bromide (EtBr, ThermoFisher 15585011), we proceeded to UV exposure on a transilluminator (Syngene bioimaging gel doc used with GeneSnap) to reveal relative biological sample sizes. This result could further be followed by appropriate cloning steps.

Centrifugation A widely used technique in microbiology is to centrifuge samples in order to separate different molecules. For the centrifugation of samples at room temperature or at 4°C, we used Eppendorf 5430R and Hettish 320 centrifuges.

Gel extraction In order to purify biological samples, one method is to load them onto a gel electrophoresis apparatus, and to extract the fragment of interest by cutting a small piece of the gel after sample migration. This is usually done with a scalpel blade, and should be performed with appropriate Personal Protective Equipment (PPE) to be protected from both UV and staining agent. Commercial kits such as the Qiaquick (Qiagen 28706) or Monarch (NEB T1020S) gel extraction kits were generally used to purify specific DNA fragments. Both protocols are based on a membrane purification method, wherein agarose-embedded DNA is first dissolved in a binding buffer at 50°C and then loaded/bound to a column through centrifugation. An ethanol-based washing buffer is then applied to clean DNA fragments, later eluted in MilliQ water or elution buffer (10mM Tris-Cl, pH 8.5).

PCR purification As a general use after PCR amplification, Qiaquick PCR purification kits (Qiagen 28106) were used to elute primers and contaminants from the amplified DNA fragments. Similarly to the Qiaquick gel extraction kit, purification is carried out by mixing samples with a binding buffer that sticks DNA to a membrane, secondly washed with an ethanol based solvent before eluting in MilliQ water or the elution buffer.

Drop dialysis For DNA purification after ligation, we dialysed samples on mixed-cellulose membranes (Millipore VSWP, 0.025 μ m pores), floating on MilliQ water in a petri dish. Thirty minutes of incubation were used to purify DNA from contaminants, and samples were then aspirated from the top of the membrane to be used for bacterial transformation.

DNA ligation DNA self-ligation can be promoted by the use of DNA ligase. A commercial kit (NEB M0202S) offers the T4 DNA ligase that allows formation of a phosphodiester bond at complementary DNA ends, ligating individual fragments to one another based on a specific cloning strategy. In general, the mixture of DNA fragments to be assembled contains various molecular species ratios, from equimolar up to 7:1 excess in one species. These ratios depend on the relative size of individual fragments to be ligated. We usually preferred higher insert to vector ratios for the assembly of low molecular weight species. Otherwise, we would generally perform a 3:1 excess in insert DNA to promote the correct assembly with a given DNA backbone. For cohesive-end ligations, reaction mixtures were left for incubation for one hour at room temperature (20-25°C), while blunt-end fragments were rather incubated overnight in

a thermocycler, repeating 30 seconds oscillating periods between temperatures of 10°C and 30°C for four hours, before stabilising at 16°C overnight.

DNA digestion via endonucleases Restriction enzymes are endonucleases that cut DNA at specific sequence they can recognise. For molecular cloning, individual enzymes were purchased from New England Biolabs (NEB), and digestion was carried out for either one hour or overnight at the enzymes optimal temperature in appropriate buffer conditions (Cutsmart buffer or NEBuffer derivatives).

Gibson-type DNA assembly By the design of overlapping arms between individual DNA fragments to be assembled, samples can be fused via a mixture of enzymes that exhibit 3'-5' exonuclease, DNA polymerase and DNA ligase activities. We used a commercial tool provided by NEB in order to perform Gibson assembly (NEB HiFi), in conjunction with their online platform to help the primer design for specific amplification of individual DNA fragments (NEBuilder HiFi DNA assembly tool).

***In-vitro* mutagenesis** In order to perform small changes on existing plasmids, we performed *in vitro* mutagenesis with the amplification of template DNA from mutagenic primers. To do this, we either used the Agilent Quickchange XL II kit or an in-house method, which only involves a 35-cycles PCR amplification and *dpnI* digestion before bacterial transformation by heat shock.

Ethanol precipitation In order to concentrate DNA, a simple method is to precipitate it, to separate it from its solution and to elute it in a smaller volume. In practice, we achieved ethanol precipitation by adding three volumes of 100% ethanol to half a volume of 5M ammonium acetate, and added this solution to the samples at a 1:1 ratio before incubating one hour at -80°C. Samples were then centrifugated at 14000rpm for 30 minutes, the supernatant was removed and residual ethanol left to evaporate for 30 minutes. DNA was then rediluted in an appropriate volume of MilliQ water and incubated for 10min at 65°C.

Plasmid DNA extraction After bacterial growth with appropriate antibiotic selection, plasmid DNA can be extracted from lysed cells, and subsequently used for further cloning steps. For high-copy number plasmids, 1.5ml of *E. coli* overnight culture was centrifuged for 30 seconds at 14000rpm, whilst 5 to 10ml of culture were used for low-copy number vectors. In order to purify plasmids from bacterial contaminants, we used commercial Qiaprep kits (Qiagen 27106). This method is based on cell lysis, RNA

degradation and protein precipitation. These contaminants are spun down as a pellet, whilst plasmid DNA remains in the supernatant. Via a similar column as the ones used in other Qiagen kits, DNA is then bound and washed on a membrane, before elution in MilliQ water or elution buffer.

Genomic DNA extraction In order to extract genomic DNA, we grew *B. subtilis* strains in rich medium up to late-exponential phase, harvested 2ml of culture and lysed the cells for DNA purification. This allowed us to avoid the purification of DNA from spores, which are much more difficult to break and represent a biological contaminant. For *E. coli* species, we harvested 1.5ml of an overnight culture before cell lysis. Bacteria were spun down for 1min at high settings and resuspended in 100 μ l 50mM Tris-EDTA (TE) supplemented with 0.5mg/ml lysozyme and 0.5mg/ml RNase A. Gram negative species were incubated 15min at 37°C while Gram-positives were left for a one hour incubation period. Samples were then processed via nuclei lysis (75mM NaCl, 24mM EDTA for 5min at 80°C), protein precipitation (Promega A7951) and spun down for 10min at 14000rpm. DNA was then extracted from the supernatant by isopropanol precipitation, washed with 70% ethanol and eluted in MilliQ water for 10min at 65°C after individual centrifugation steps.

Bacterial transformation In order to propagate plasmid DNA in *E. coli*, we used electroporation and heat-shock as the main transformation methods. Electroporation consists of applying an electric field through a bacterial sample mixed with DNA, which forces the entry of assembled DNA products through the cell wall. We used a Biorad Gene Pulser II (25 μ F) with a Pulse Controller II (200 Ω) to electroporate 100 μ l samples in 1mm cuvettes. As an alternative, we also used heat-shock as a mean to transform DNA into bacterial cells. We based our method on the Hannahan high-transformation efficiency protocol, available in the Sambrook and Green manual[83].

Optical density reading In order to measure bacterial growth, optical density (OD) readings provide an approximation of the number of cells in a sample. We followed specific bacterial growth via OD measurements at a 600nm wavelength (OD_{600}) using Biochrom Libra S35 or WPA biowave CO8000 spectrophotometers. Optical density readings may also be used in order to analyse DNA molecules, and provide an accurate approximation of a species concentration and purity in a certain volume. In order to verify DNA concentration after sample purification, we used a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific 79482).

Bioluminescence and fluorescence plate scanning During the development of genetic circuits, we used chemoluminescence as a control of adequate biological behaviour. To image this, we used an EM-CCD camera at -25°C in an ImageQuant LAS4000 mini platform. For plate scanning, a 1min illumination period was sufficient to provide a qualitative answer about biological results in different conditions. By similar means but for fluorescence imaging, a GE Healthcare Typhoon trio variable mode imager was used with ImageQuant in order to analyse green fluorescence from plates.

Flow cytometry Flow cytometers provide a technology that allows the analysis of single cells at a population scale level. In this thesis, we used two models of flow cytometer: a Partec CyFlow space and a Fortessa x20. At specific time points, bacterial cultures were diluted and loaded onto the equipment in polypropylene tubes (Sarsedt 55.484.005), and individual cells were analysed via laser refraction, providing information about cell size and fluorescence profiles. An integrated analysis software was used for the gathering of experimental data (FloMax), and individual plots were then adequately collated for sample comparison.

Fluorescence microscopy We described in Chapter 4 a series of genetic devices that used fluorescence as a reporter system. For the imaging of individual cells, we used M200 and Nikon Ti (Ti-E/Ti-U) inverted microscopes. These microscopes were combined with a live imaging system, and mostly used in oil immersion for phase contrast and fluorescence measurements.

Fluorescence plate readers For the characterisation of kinetic changes in fluorescence over time, we used automated plate readers that performed simultaneously OD and fluorescence measurements. Used with 96-well plates (Greiner Bio-one, black and flat-bottom, 655090), we performed experiments in BMG labtech Fluostar optima and Clariostar plate readers.

Appendix B

Recombining barcode DNA in *B. subtilis* and *E. coli*

In Chapter 3, we explained the assembly of barcode recombinant products for the cloning of two model bacterial species. Here, we provide the experimental protocols required to clone these assembled fragments into *B. subtilis* and *E. coli* strains. We developed and optimised these methods in order to obtain a maximum number of true-positive colonies after transformation. These protocols were also proof-tested by three additional scientists (Jonathan Tellechea, Dr. Nunzia Lopiccolo and Dr. Wendy Smith).

B.1 Transformation in *B. subtilis*

In *B. subtilis*, we used an antibiotic marker (zeocin) to propagate recombinant barcodes within the genome. This was made according to the following protocol:

1. Inoculate a single colony of the strain to be made competent (ready to receive recombinant DNA) in minimal medium (MM: 10ml SMM basic salts, 125 μ l 40% (w/v) glucose, 100 μ l 2% (w/v) tryptophan, 60 μ l 1M Mg₂SO₄·7H₂O, 10 μ l 20% (w/v) casaminoacids, 5 μ l 2.2mg/ml ferric ammonium citrate) and grow overnight at 37°C
2. Dilute saturated medium 1/100 in fresh MM and grow for 3h at 37°C
3. Prewarm a solution of starvation medium (SM: 10ml SMM basic salts, 125 μ l 40% (w/v) glucose, 60 μ l 1M Mg₂SO₄·7H₂O) and dilute the refreshed culture 1:1 in SM and set for a further 2h incubation at 37°C
4. Aliquot 400 μ l cells, add about 1 μ g recombinant DNA and shake in a 2ml eppendorf for 1h at 37°C

5. Plate dilutions on nutrient agar (NA) plates supplement with the appropriate antibiotic (zeocin at $20\mu\text{g}/\text{ml}$) and incubate overnight at 37°C

Coupled to an inducible mazF toxin[258, 162, 260], successful transformant colonies (PCR screening and sequencing) were isolated and induced to express a toxin favouring excision of the selection/counter-selection cassette according to the following procedure:

1. Pick a positive colony and inoculate a LB culture supplemented with 0.4% (w/v) glucose/ $20\mu\text{g}/\text{ml}$ zeocin, grow overnight at 37°C
2. Dilute the overnight culture to $OD_{600} = 0.1$ in LB with 0.4% (w/v) glucose only and grow to $OD_{600} = 0.3$
3. Add 1% (w/v) xylose to the culture and incubate for 8h at 37°C
4. Spot $30\mu\text{l}$ of culture on a NA/1% (w/v) xylose plate, streak to obtain single colonies and grow overnight at 37°C
5. Restreak individual colonies onto NA/1% (w/v) xylose and NA/ $20\mu\text{g}/\text{ml}$ zeocin plates and screen for zeocin sensitivity (loss of the selection cassette)

B.2 Transformation in *E. coli*

E. coli MG1655 and many other substrains do not encode recombinase enzymes that are required for homologous recombination. Therefore, we first propagated a helper plasmid in barcode recipient strains (via electroporation). This plasmid contained the PBAD inducible promoter to drive the expression of three essential proteins that allow chromosomal insertion of recombinant DNA. It also contained a temperature-sensitive replicon, and bacteria should thus grow at 30°C to avoid plasmid loss. Bacteria harboring the helper plasmid were made electrocompetent and transformed with recombinant DNA according to the following protocol:

1. Pick a single colony to inoculate an overnight culture of LB medium supplemented with helper plasmid antibiotic (ampicillin $100\mu\text{g}/\text{ml}$ for the pKD46 helper plasmid) at 30°C
2. Refresh the culture 1/100 and grow to $OD_{600} = 0.1$ at 30°C
3. Add arabinose to 0.7% and grow cultures for a further 90min at 30°C . This step provides time for the recombination enzymes to be synthesised.

4. Prepare electrocompetent cells by aliquoting 1.4ml culture and spinning/resuspending in ice-cold glycerol at 4°C twice (10min at 3000rpm centrifuge settings). Spin cells again and resuspend in 50µl of residual glycerol.
5. Electroporate cells with 500ng recombinant DNA and grow cells for 3h at 37°C, which allows recovery of transformed cells
6. Plate culture dilutions on NA/34µg/ml chloramphenicol and grow for 24h at 37°C

Positive colonies were identified by restreaking on NA/ampicillin plates (sensitive with plasmid loss), PCR and sequencing.

Selection cassette removal in *E. coli* was performed using a flipase, an enzyme recognising and cleaving at pairs of 34bp Frt sites[103]. A positive barcoded colony was picked and transformed with a helper plasmid (pCP20)[32]. Similarly to the previous helper plasmid, this one's replicon was temperature-sensitive and should not replicate at temperatures higher than 30°C. Colonies harboring the helper plasmid were then induced to lose the selection cassette by following these steps:

1. Inoculate in LB supplemented with helper plasmid antibiotics a barcoded colony and grow at 30°C overnight
2. Dilute cells 1/100 and grow to $OD_{600} = 0.1$ at 30°C
3. Swap the cultures from 30°C incubation to a 42°C shaking water bath and grow cultures to $OD_{600} = 0.9$
4. Spot 30µl of culture on a NA plate, streak to obtain single colonies and grow overnight at 37°C
5. Restreak single colonies on NA, NA/ampicillin and NA/chloramphenicol plates. Positive colonies have lost the helper plasmid (ampicillin sensitivity) and the barcode selection cassette (chloramphenicol)

Appendix C

DNA barcoding kit

The following protocol provides the experimental procedures, written as for a kit, to undertake barcoding studies and to retrieve the barcode sequence of any barcoded strain.

How to barcode bacteria? Cloning *E. coli* or *B. subtilis* with barcodes works in a two-steps process: (i) the assembly of recombinant DNA and (ii) a species-specific transformation protocol. Using the reagents supplied with the Bac2code cloning kit, follow the protocol "Assembling recombinant barcode DNA" to assemble ready-to-clone barcode DNA. According to the species you are working with, follow either the "Transforming *E. coli* with recombinant barcodes" or "Transforming *B. subtilis* with recombinant barcodes" protocol to barcode bacteria.

How to retrieve Bac2code documentation from an unknown barcoded strain? Using the cloning kit universal primer and the "Retrieving DNA barcodes" protocol, you can amplify barcode sequences from genomic DNA and sequence them to gain access to the relevant online Bac2code profile.

What comes in the Bac2code cloning kit?

- Barcode sequences (B fragments)
- Recombinant DNA fragments (L1, R1 and S1 for *E. coli* vs. L2, R2 and S2 for *B. subtilis*)
- PCR verification primers (pF1 forward and pR1 reverse primers for *E. coli* vs. pF2 forward and pR2 reverse primers for *B. subtilis*)
- Universal sequencing primer (pU1)
- *E. coli* helper plasmids (H1, H2 and H3, respectively pKD46, pIJ790 and pCP20)

- 40% glucose supplement (G, Sigma G8270)
- Nuclei lysis solution (N, Promega A7941)
- Protein precipitation solution (P, Promega A7953)

C.1 Assembling recombinant barcode DNA

Aim Prior to cloning, barcodes need to be fused to homology regions and selection cassette to promote the occurrence of double-recombination events in *E. coli* and *B. subtilis*.

Species-specific information (supplied with the Bac2code cloning kit)

- in *E. coli*: use L1 (left homology), R1 (right homology) and S1 (chloramphenicol antibiotic resistance) fragments with primer pair pF1/pR1.
- in *B. subtilis*: use L2 (left homology), R2 (right homology) and S2 (zeocin antibiotic resistance) fragments with primer pair pF2/pR2.

Reagents needed

- Polymerase chain reactions (PCR) should be carried out using a high-fidelity DNA polymerase. Examples are given for the NEB Q5 DNA polymerase (M0491).
- Step 3 gel extraction should be carried out using a commercial kit (NEB T1020 or Qiagen 28704).

C.1.1 Primer-free PCR assembly

Prepare a 25 μ l PCR reaction as follows: mix 5 μ l Q5 buffer, 0.5 μ l 10mM dNTPs, 2 μ l of L- and R-fragments, 1 μ l of B- and S-fragments, 5 μ l Q5 GC enhancer, 8.25 μ l MilliQ water and 0.25 μ l Q5 polymerase. Run the reaction in a thermocycler for 12 cycles with a low annealing temperature (TA=55°C) and a 2.5kb extension time (1min40sec) using your PCR reagents recommendations.

C.1.2 Recombinant DNA amplification

Prepare a 200 μ l PCR reaction as follows and split into 50 μ l aliquots: mix 40 μ l Q5 buffer, 4 μ l 10mM dNTPs, 10 μ l of pF and pR primers, 12 μ l of PCR products from the primer-free PCR, 40 μ l Q5 GC enhancer, 82 μ l MilliQ water and 2 μ l Q5 polymerase. Run the reaction in a thermocycler for 25 cycles with a high annealing temperature (TA=64°C) and a 2.5kb extension time (1min40sec) using your PCR reagents recommendations.

C.1.3 Purification

Mix recombinant DNA obtained from the previous step to a DNA loading dye for gel extraction, and run the whole sample on a 1.2% agarose gel electrophoresis. Using a recommended gel extraction kit, elute the correct gel size band in 30 μ l MilliQ water:

- for *E. coli*, cut and purify the 1875bp band.
- for *B. subtilis*, cut and purify the 2123bp band.

This aliquot is the recombinant barcode DNA, ready-to-use for bacterial transformation.

C.2 Transforming *B. subtilis* with recombinant barcodes

Aim Recombinant barcode DNA is naturally taken up by *B. subtilis* and selection cassette excision is based on the MazF toxin.

Reagents needed

- Luria-Bertani (LB) growth medium: for a 1L preparation, mix 10g peptone, 5g yeast extract, 10g sodium chloride with 800ml MilliQ water. Adjust the volume to 1L and autoclave 35min on sensitive cycle
- sterile 40% glucose (solution G supplied with the Bac2code cloning kit)
- minimal medium (MM): 10ml SMM basic salts, 125 μ l 40% (w/v) glucose, 100 μ l 2% (w/v) tryptophan, 60 μ l 1M Mg₂SO₄·7H₂O, 10 μ l 20% (w/v) casaminoacids, 5 μ l 2.2mg/ml ferric ammonium citrate
- starvation medium (SM): 10ml SMM basic salts, 125 μ l 40% (w/v) glucose, 60 μ l 1M Mg₂SO₄·7H₂O

- nutrient agar (NA) plates: for a 1L preparation, mix 3g beef extract, 5g peptone, 15g agar with MilliQ water. Autoclave on sensitive cycle, add antibiotics at 50°C, pour into petri dishes and keep at 4°C for further use (max 2 weeks for plates supplemented with antibiotics)
- working antibiotics/supplements concentrations: 20µg/ml for zeocin, 1% xylose

C.2.1 Barcoding cells

Inoculate a single colony from the target strain in 10mL MM and shake at 37°C overnight. In the morning, refresh the culture 1/100 in 10mL MM, prepare 10mL sterile SM to prewarm and incubate at 37°C for 3h. Add the warm SM to the cells and carry on incubation for 3h. Aliquot 400µl cells in a 2ml eppendorf tube, mix with 1µg recombinant barcode DNA, and shake for 1h at 37°C (barcode integration now occurs). Spread 100µl cells onto a NA/Zeocin plate and incubate overnight at 37°C. Restreak a single colony on NA/Zeocin plates, incubate overnight at 37°C and screen for positive transformants by PCR using the pF2/pR2 primer pair at 62°C for a 2123bp band.

C.2.2 Removing the selection marker

Inoculate a single colony from Step 2 in 10mL LB/zeocin and grow the cells at 37°C overnight. In the morning, dilute cells 1/100 in 10mL fresh LB without antibiotics and grow up to OD=0.4. Add 1% xylose to the culture and grow the cells for 10h. Streak a 30µl cells spot from the edge of a NA/xylose plate and incubate overnight at 37°C. Restreak single colonies on NA (positive colony) and NA/Zeocin (selection cassette excision control); incubate overnight at 37°C.

Single colonies from the NA plate that do not grow on NA/Zeocin are barcoded strains. A PCR using the pF2/pR2 primer pair at 62°C should lead to a 1170bp band.

C.3 Transforming *E. coli* with recombinant barcodes

Aim Prior to barcode DNA transformation, the target *E. coli* strain needs to be propagated with a helper plasmid remotely expressing essential proteins for chromosomal recombination. Induction of recombinase proteins by arabinose allows the integration of recombinant barcode DNA into the bacterial genome.

Helper plasmids information (supplied with the Bac2code cloning kit) The H1 helper plasmid contains an ampicillin resistance cassette for selection. If this is incompatible with the target strain design, use the alternative H2 plasmid conferring resistance to chloramphenicol.

Reagents needed

- Luria-Bertani (LB) growth medium: for a 1L preparation, mix 10g peptone, 5g yeast extract, 10g sodium chloride with 800ml MilliQ water. Adjust the volume to 1L and autoclave 35min at 123°C.
- sterile 40% glucose (solution G supplied with the Bac2code cloning kit).
- sterile 10% glycerol: for a 1L preparation, mix 100ml glycerol with 700ml MilliQ water and mix. Adjust the volume to 1L and autoclave 35min at 123°C. Cool down and keep the bottle sterile at 4°C.
- Nutrient agar (NA) plates: for a 1L preparation, mix 3g beef extract, 5g peptone, 15g agar with MilliQ water. Autoclave 35min at 123°C, add antibiotics at 50°C, pour into petri dishes and keep at 4°C for further use (max 2 weeks for plates supplemented with antibiotics).
- Working antibiotics/supplements concentrations: 100µg/ml for ampicillin, 34µg/ml for chloramphenicol and 0.7% arabinose.

C.3.1 Propagate the helper plasmid

Inoculate a fresh colony of the target strain in 25ml LB and shake the culture overnight at 37°C. In the morning, dilute cells 1/100 in 25ml fresh LB medium supplemented with 0.4% glucose (250µl solution G) and incubate at 37°C. At $OD_{600} \sim 0.4 - 0.5$, freeze the culture on ice for 15min and turn a cold centrifuge on. Further steps should be carried out at 4°C. Spin cells at 5000rpm for 10min, discard supernatant and resuspend the bacterial pellet in the same volume of ice-cold 10% glycerol. Repeat this step twice; discard the final supernatant and aliquot 50µl concentrated cells from the residual glycerol. Add 1µl helper plasmid (H1 solution), mix and transfer to a chilled 1mm electroporation cuvette. Place the cuvette in a gene pulser (25µF, 200Ω at 1.8kV) and electroporate cells. A time constant close to 5ms is expected. Immediately add 950µl LB to the cuvette, pipette up/down six times, transfer to a 12ml round-bottom falcon tube and shake for 1h at 30°C. Spread 50µl cells onto a NA plate supplemented with ampicillin (or chloramphenicol if H2 used). Incubate the plate overnight at 30°C.

C.3.2 Barcoding cells

Inoculate a single colony from Step 1 transformants in 25ml LB/ampicillin and shake at 30°C overnight. In the morning, refresh the culture 1/100 in 25ml LB/ampicillin and grow the cells at 30°C to $OD_{600} \sim 0.2 - 0.3$. Add arabinose to a final concentration of 0.7% and grow for 1h at 37°C, this gives time for recombination proteins to be expressed. Further steps should be carried out at 4°C. Spin cells at 5000rpm for 10min, discard supernatant and resuspend the bacterial pellet in the same volume of ice-cold 10% glycerol. Repeat this step twice; discard the final supernatant and aliquot 50 μ l concentrated cells from the residual glycerol. Add 400ng recombinant barcode DNA, mix and transfer to a chilled 1mm electroporation cuvette. Place the cuvette in a gene pulser (25 μ F, 200 Ω at 1.8kV) and electroporate cells. A time constant close to 4ms is expected. Immediately add 950 μ l LB to the cuvette, pipette up/down 6 times, transfer to a 12ml round-bottom falcon tube and shake for 3h at 37°C (recombination now occurs). Spread 100 μ l cells onto a NA plate supplemented with chloramphenicol and incubate overnight at 37°C. Restreak single colonies on NA/chloramphenicol plates, incubate overnight at 37°C and screen for positive transformants from genomic DNA by PCR using the pF1/pR1 primer pair at 64°C for a 1875bp band.

C.3.3 Removing the selection marker

Site-specific recombination Starting from a positive colony obtained in Step 2, reiterate Step 1 with the H3 helper plasmid that promotes excision of the selection cassette. After electroporation, carry out all incubation steps at 30°C as H3 contains a temperature-sensitive replicon; adjust antibiotics concentrations to 100 μ g/ml ampicillin and 17 μ g/ml chloramphenicol.

Excision of the selection cassette Inoculate a single colony from previous step in 25ml LB/ampicillin/chloramphenicol and grow the cells at 30°C overnight. In the morning, dilute cells 1/100 in 25ml fresh LB without antibiotics and grow at 30°C up to $OD_{600} = 0.1$. Swap cells to a 42°C shaking water bath and carry on bacterial growth up to $OD_{600} = 0.9$. Streak out from a 30 μ l cells spot a NA plate and incubate overnight at 37°C. Restreak a single colony on NA (positive colony), NA/ampicillin (helper plasmid loss control) and NA/chloramphenicol (selection cassette excision control); incubate overnight at 37°C.

Single colonies from the NA plate (not growing on the antibiotics plates) are bar-coded strains. A PCR from genomic DNA using the pF1/pR1 primer pair at 64°C should reveal a band at 1068bp.

C.4 Retrieving DNA barcodes

Aim For any barcoded strain, a straightforward process consisting of amplifying genomic DNA containing the barcode allows its sequencing via universal primers, and to retrieve strain information from the Bac2code server.

Species-specific information (supplied with the Bac2code cloning kit)

- in *E. coli*: use the pF1/pR1 primer pair at an annealing temperature $T_A = 64^\circ C$.
- in *B. subtilis*: use the pF2/pR2 primer pair at an annealing temperature $T_A = 62^\circ C$.

Reagents needed

- Polymerase chain reactions (PCR) should be carried out using a high-fidelity DNA polymerase. Examples are given for the NEB Q5 DNA polymerase (M0491)
- PCR purifications should be carried out using a commercial kit (Qiagen 28104)
- Luria-Bertani (LB) growth medium: for a 1L preparation, mix 10g peptone, 5g yeast extract, 10g NaCl with 800ml MilliQ water. Adjust the volume to 1L and autoclave 35min at 123°C
- sterile 40% glucose (solution G supplied with the Bac2code cloning kit)
- TE buffer: 10mM Tris-HCl, 1mM EDTA in distilled water
- RNase A (10mg/ml, store at 4°C – Sigma R6513) and lysozyme (10mg/ml, store at –20°C – Sigma L3790)
- 100% isopropanol and 70% ethanol
- Protein precipitation and nuclei lysis buffers are supplied with the Bac2code cloning kit (respectively solutions P and N – Promega A7953 A7941)

C.4.1 Harvest bacterial genomic DNA

Inoculate 2.5ml of fresh LB/0.4% glucose with the strain to identify and shake at 37°C up to $OD_{600} \sim 0.5 - 0.6$. Spin spin cells down for 3min in a high-speed centrifuge and resuspend the pellet in 100 μ l TE and add 5 μ l RNase A and 5 μ l lysozyme, incubate at 37°C for 30-45min. Add 600 μ l solution N and heat for 5min at 80°C. Cool down to room temperature, add 200 μ l solution P, vortex thoroughly and place on ice for 10min. Spin down 10min at 14000rpm and add 600 μ l supernatant to a clean 1.5ml eppendorf containing 600 μ l room temperature 100% isopropanol. Invert the tube about 10-15 times, until DNA precipitation becomes visible, and spin for 10min at 14000rpm. Tip the supernatant out, add 600 μ l 70% ethanol and spin for 5min at 14000 rpm. Carefully aspirate the supernatant and let the tube opened to dry out in a clean place for 15min. Add 30 μ l MilliQ water and let DNA dissolve for 10min at 65°C. This provides the eluted genomic DNA for a specific barcoded strain.

C.4.2 Barcode DNA amplification

Prepare a 25 μ l PCR reaction as follows: mix 5 μ l Q5 buffer, 0.5 μ l 10mM dNTPs, 1.25 μ l of pF and pR primers, 1 μ l of barcoded strain genomic DNA, 5 μ l Q5 GC enhancer, 10.75 μ l MilliQ water and 0.25 μ l Q5 polymerase. Run the reaction in a thermocycler for 35 cycles at the species-specific recommended annealing temperature with a 1.2kb extension time (1min30sec).

C.4.3 Purification and barcode retrieval

After checking 4 μ l PCR products on gel electrophoresis for the presence of a 1170bp (*B. subtilis*) or 1068bp band (*E. coli*), purify DNA using a commercial PCR purification kit and perform the final elution step in 30 μ l MilliQ water after a 10min incubation period with the washing buffer. Send this purified PCR fragment to sequencing using the pU1 primer and input sequencing results onto the Bac2code server to retrieve bacterial strain information.

C.5 Troubleshooting

- **A tip to detect barcode positive clones?** In general, after barcode DNA transformation, smaller colonies on plates reveal the right chromosome insertion.

- **What if I get low recombinant barcode DNA concentration?** This procedure should lead to a final DNA concentration $\geq 30\text{ng}/\mu\text{l}$. If DNA products concentration is $\leq 20\text{ng}/\mu\text{l}$, it is recommended to rehearse Step 23 from leftover Step 1 PCR products.
- **What if I get low genomic DNA concentration?** This procedure should lead to a final DNA concentration $\geq 100\text{ng}/\mu\text{l}$. If DNA products concentration is within the range $20 - 50\text{ng}/\mu\text{l}$, it is recommended to add more template DNA for PCRs ($\sim 100\text{ng}$). If DNA concentration is $\leq 20\text{ng}/\mu\text{l}$, it is recommended to repeat the genomic DNA extraction while being extra-careful not to aspirate DNA after the ethanol wash. Optionally, you can leave the bottom $50\mu\text{l}$ from the supernatant and dry samples overnight in a clean cabinet before eluting in water.
- **What if the right band does not show up on gel?** Always verify that you use the right primer pair for the sample organism (pF1/pR1 for *E. coli* vs. pF2/pR2 for *B. subtilis*). Always verify that you use the right annealing temperature in PCR reactions. Always verify that PCRs elongation time can cover an up to 2kb product. Typically, 1min40 for the Q5 polymerase should always be suitable.
- **What if I get low *B. subtilis* cloning efficiency?** Always prewarm the starvation medium (SM) well in advance and add it without removing cultures from the 37°C incubator. Always compare your PCR results to a negative control: *B. subtilis* non-barcoded strains show a 1002bp fragment with pF2/pR2 primers.
- **What if I get low *E. coli* cloning efficiency?** If the electroporation time-constant for plasmids or barcode DNA propagation is very different from the standard (respectively 5ms and 4ms), make sure your cells are fully resuspended during the glycerol wash for competent cells preparation. Always compare your PCR results to a negative control: *E. coli* non-barcoded strains show a 903bp fragment with pF1/pR1 primers.
- **What if I do not get any sequencing read?** Make sure PCR fragments are always very clean before sending sequencing reactions. In standard purification protocols, a longer incubation period with the cleaning buffer (usually ethanol-based) improves a lot samples purity. Additionally, you may want to try using pR1 or pR2 primers to identify barcodes from the reverse DNA strand. *E. coli* strains should provide a read with pR1, while pR2 should be used in *B. subtilis*.

Appendix D

Making PDMS microfluidics chips

The following protocols provide details of the experimental procedures required to produce intermediate hard-/soft PDMS or single-use soft PDMS biochips.

D.1 Intermediate-PDMS template

This protocol was designed to produce PDMS-based microfluidics intermediates. A thin layer of hard PDMS (h-PDMS) is deposited onto a SU-8 photoresist featuring microfluidics patterns to reproduce. This surface layer is backed with a thicker layer of soft PDMS (s-PDMS) and serves as master for the molding of further s-PDMS chips.

For this protocol, lab coat, nitrile gloves and goggles should be worn at all times. Silanising the wafer, in this one and further protocols, is used to guarantee better separation between template and PDMS after cross-linking.

1. Preparing the wafer:

- Wash the wafer with (i) 100% isopropanol, (ii) 100% acetone and (iii) 100% isopropanol
- Blow dry the wafer with pressurised air
- Silanise the wafer in a vacuum chamber: (i) center and stick the wafer to a 50ml falcon tube inverted cap using blue tak or equivalent, (ii) under a fume hood, pipette $40\mu\text{l}$ 1H,1H,2H,2H perfluorooctyl tricholasilane onto a piece of whatman paper and place it in the vacuum chamber (leave pipette tip inside, perfluorooctyl tricholasilane is toxic), (iii) apply vacuum to 27psi and incubate for 30min, and (iv) slowly reintroduce air into the vacuum chamber

2. Creating the h-PDMS layer:

- Prepare the h-PDMS mixture:

- In a 50ml falcon tube, mix 27.2g Vinylmethylsiloxane-dimethylsiloxane copolymer (pour directly from container) and 144 μ l Platinum divinyltetramethyldisiloxane complex (with a 200 μ l tip)
- Add 0.8g 1,3,5,7-Tetravinyl tetramethylcyclotetrasiloxane (disposable pipette) and mix
- Add 8g Methylhydrosiloxane- dimethylsiloxane copolymer (disposable pipette) and mix
- Add 4g hexane with a disposable pipette and mix: careful to eventual vapors
- Degas the polymer mixture in a vacuum chamber at 27psi for 15min: increase/decrease pressure at first to avoid PDMS bubbles splashing in the vacuum chamber
- Preheat a hot plate to 65°C
- Set the clean wafer (photoresist) at the centre of a spin coater on the central spin chuck
- Activate vacuum and verify that the wafer is well centered
- Spin at 100rpm for 1 min and pour degassed h-PDMS at the centre of the spinning wafer, slowly move towards wafer edges as PDMS spreads from the centre
- Place the wafer onto a clean towel in the vacuum cleaner and degas all bubbles that may have been formed after spinning (5-10min at 27psi)
- Place the wafer back on the spin coater and run it for 5sec at 500rpm, then 40sec at 1000rpm
- Place the wafer in a large petri dish and bake for 1h on the 65°C hot plate

3. Creating the s-PDMS layer:

- Mix 70g Sylgard 184 base and 7g Sylgard 184 curing agent using a serological pipette in a 200ml beaker
- Degas the mixture at 27psi in a vacuum chamber for 30min
- Fold a holder for the master wafer that is coated with h-PDMS: (i) place the wafer at the center of a piece of strong tin foil, dull side up and (ii) fold the aluminium at the edges of the wafer to build walls that will contain the s-PDMS from pouring out of the h-PDMS layer
- Gently pour degassed s-PDMS at the centre of the wafer

- Degas at 27psi in a vacuum chamber to remove any residual PDMS bubbles for 10min
- Place on a hot plate at 65°C and bake overnight (16-18h)

4. Removing PDMS from the wafer:

- Remove the wafer from the hot plate and let it cool down to room temperature
- Remove any residual PDMS that may have leaked at the bottom of the wafer
- Gently detach the tin foil from the coated wafer
- With a scalpel n11, very precisely separate PDMS from the silicon wafer by sliding the blade all around its edges
- Delicately pull the polymer layer from the wafer and trim the polymer edges as wanted for further use

D.2 Cleaning microfluidics wafers

Sometimes, removing a wafer from a photoresist can leave residual polymer behind. For a better conservation of SU-8 photoresists, wafers are best kept clean and any residue should be removed. The recommended cleaning method is to successively use isopropanol and acetone washes. Pressurised air is very good at detaching tiny pieces of polymer that may be stuck in small microfluidics features. However, if none of these methods manage to clean the wafer, one last procedure remains possible: a piranha solution wash to dissolve any organic matter on the SU-8 layer. For this protocol, besides ordinary personal protective equipment (PPE, lab coat and goggles), wear a face shield and trionic gloves at all times. Pyrex glassware should be used at all times. In a class 2+ fume hood, prepare the piranha solution: in a beaker, pour 60ml sulfuric acid and add 20ml hydrogen peroxyde millilitre by millilitre. Carefully mix the mixture and be aware of hydrogen fumes. This chemical reaction produces a lot of heat, do not touch the glassware without gloves. Pour the piranha solution in a large petri dish and tilt the wafer upside down in the solution. Gently tilt the petri dish back and forth to immerse the entire wafer and incubate overnight in the fume hood on high fan settings. In the morning, wash the wafer with deionised water for 1h and blow it dry with pressurised air before storing it in a clean wafer box. Neutralise the piranha solution with 99% NaOH: in a fume hood, (i) add NaOH millilitre by millilitre and stir the mixture, being aware of splashes generated by the exergonic reaction, and

(ii) when the pH reaches a value between 4 and 10, dispose of the solution via a drain under running tap water.

D.3 Single-use soft-PDMS

This protocol was designed to produce single-use soft PDMS microfluidics chips. A thick layer of s-PDMS is deposited onto a master intermediate mould to reproduce the features of the initial SU-8 photoresist.

For this protocol, a lab coat, nitrile gloves and goggles should be worn at all times.

1. Prepare the wafer the same way as detailed for intermediate chips with a 45min silanisation incubation period.
2. Create the s-PDMS layer the same way as detailed for intermediate chips but proceed to a 2h bake at 80°C instead
3. Separate the s-PDMS layer from the intermediate mould the same way intermediate PDMS is removed from SU-8 photoresists. Separate polymers very slowly as soft PDMS is more likely to break at small microfluidics features
4. Trim the s-PDMS layer to make individual s-PDMS chips according to design (keep track of chip identifiers)
5. Bonding chips to coverslips:
 - Features side up, delicately create holes in s-PDMS chips by using a 0.75mm biopsy punch at the tubing inlet/outlet ports
 - Turn on a pressurised plasma cleaner and prepare 3 individual s-PDMS chips to bond (3 polymer chips and 3 glass coverslips)
 - Close the plasma cleaner door and adjust the internal pressure to 500-1000mTor
 - Turn the plasma on "high" setting for 45sec and turn it off before reintroducing air into the chamber
 - Within a minute, invert s-PDMS chips onto glass coverslips and apply a gentle pressure on top of each chip. Verify that the polymer is well bonded to the glass and that no air bubbles remain.

Bibliography

- [1] *A Universally Unique Identifier (UUID) URN Namespace*. <https://tools.ietf.org/pdf/rfc4122.pdf>. Accessed: 2019-01-11.
- [2] P Acred, DM Brown, DH Turner, and MJ Wilson. "Pharmacology and chemotherapy of ampicillin—a new broad-spectrum penicillin." In: *British Journal of Pharmacology and Chemotherapy* 18(2):356-369 (1962).
- [3] C Aldridge, K Poonchareon, S Saini, T Ewen, A Soloyva, CV Rao, K Imada, T Minamino, and PD Aldridge. "The interaction dynamics of a negative feedback loop regulates flagellar number in *Salmonella enterica* serovar *Typhimurium*". In: *Molecular Microbiology* 78(6):1416-30 (2010). DOI: 10.1111/j.1365-2958.2010.07415.x.
- [4] C Alfiniyah, MA Bees, and AJ Wood. "Pulse Generation in the Quorum Machinery of *Pseudomonas aeruginosa*". In: *Bulletin of Mathematical Biology* 79(6):1360-1389 (2017). DOI: 10.1007/s11538-017-0288-z.
- [5] M Altissimo. "E-beam lithography for micro-nanofabrication". In: *Biomicrofluidics* 4(2):026503 (2010). DOI: 10.1063/1.3437589.
- [6] M Amos. "Population-based microbial computing: a third wave of synthetic biology?" In: *International Journal of General Systems* 43.7 (2014), pp. 770–782. DOI: 10.1080/03081079.2014.921001.
- [7] J. An, JE. Mei Teoh, R. Suntornnond, and CK. Chua. "Design and 3D Printing of Scaffolds and Tissues." In: *Engineering* 1(2):261-268 (2015). DOI: <https://doi.org/10.15302/J-ENG-2015061>.
- [8] C Anagnostopoulos and J Spizizen. "Requirements for Transformation in *B. subtilis*." In: *Journal of Bacteriology* 81(5):741–746 (1961).
- [9] JB Andersen, C Sternberg, LK Poulsen, SP Bjorn, M Givskov, and S Molin. "New Unstable Variants of Green Fluorescent Protein for Studies of Transient Gene Expression in Bacteria." In: *Applied and Environmental Microbiology* 64(6):2240–2246 (1998).

- [10] *Anderson promoter collection*. <http://parts.igem.org/Promoters/Catalog/Anderson>. Accessed: 2018-03-16.
- [11] M Arnaud, A Chastanet, and M Debarbouille. "New Vector for Efficient Allelic Replacement in Naturally Nontransformable, Low-GC-Content, Gram-Positive Bacteria". In: *Applied and Environmental Microbiology* 70.11 (2004), pp. 6887–6891. DOI: 10.1128/AEM.70.11.6887-6891.2004.
- [12] T Baba, T Ara, M Hasegawa, Y Takai, Y Okumura, M Baba, KA Datsenko, M Tomita, BL Wanner, and H Mori. "Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection." In: *Molecular Systems Biology* 2.1 (2006), 2006.0008–n/a. DOI: 10.1038/msb4100050.
- [13] M Baker. "1,500 scientists lift the lid on reproducibility." In: *Nature News Feature* (2016).
- [14] NQ Balaban, J Merrin, R Chait, L Kowalik, and S Leibler. "Bacterial Persistence as a Phenotypic Switch." In: *Science* 305(5690):1622–1625 (2004). DOI: 10.1126/science.1099390.
- [15] S Basu, Y Gerchman, CH Collins, FH Arnold, and R Weiss. "A synthetic multicellular system for programmed pattern formation". In: *Nature* 434(7037):1130-4 (2005).
- [16] S Basu, R Mehreja, S Thiberge, MT Chen, and R Weiss. "Spatiotemporal control of gene expression with pulse-generating networks". In: *Proceedings of the National Academy of Sciences* 101.17 (2004), pp. 6355–6360. DOI: 10.1073/pnas.0307571101.
- [17] DJ Beebe, GA Mensing, and GM Walker. "Physics and Applications of Microfluidics in Biology". In: *Annual Review of Biomedical Engineering* 4.1 (2002), pp. 261–286. DOI: 10.1146/annurev.bioeng.4.112601.125916.
- [18] KC Bhargava, B Thompson, D Iqbal, and N Malmstadt. "Predicting the behavior of microfluidic circuits made from discrete elements." In: *Scientific Reports* 5:15609 (2015). DOI: 10.1038/srep15609.
- [19] D Bindels, L Haarbosch, L VanWeeren, M Postma, KE Wiese, M Mastop, S Aumonier, G Gotthard, A Royant, MA Hink, and T Gadella Jr. "mScarlet: a bright monomeric red fluorescent protein for cellular imaging." In: *Nature Methods* 14:53–56 (2017).

- [20] *Biobrick community collection of ribosome binding sites*. http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Community_Collection. Accessed: 2019-01-12.
- [21] AW Bisson-Filho, YP Hsu, GR Squyres, E Kuru, F Wu, C Jukes, Y Sun, C Dekker, S Holden, MS VanNieuwenhze, Brun YV, and EC Garner. "Treadmilling by FtsZ filaments drives peptidoglycan synthesis and bacterial cell division." In: *Science* 355(6326):739-743 (2017). DOI: 10.1126/science.aak9973.
- [22] I Biswas, A Gruss, SD Ehrlich, and E Maguin. "High-efficiency gene inactivation and replacement system for gram-positive bacteria." In: *Journal of Bacteriology* 175.11 (1993), pp. 3628–3635. DOI: 10.1128/jb.175.11.3628-3635.1993.
- [23] J Blakes, J Twycross, FJ Romero-Campero, and N Krasnogor. "The Infobiotics Workbench: an integrated in silico modelling platform for Systems and Synthetic Biology." In: *Bioinformatics* 27(23):3323–3324 (2011). DOI: 10.1093/bioinformatics/btr571.
- [24] J Boldt. "Synthetic biology: Origin, scope, and ethics." In: *Minding Nature* 3(1) (2010).
- [25] I Boros, G Pósfai, and P Venetianer. "High-copy-number derivatives of the plasmid cloning vector pBR322." In: *Gene* 30(1):257-260 (1984). DOI: [https://doi.org/10.1016/0378-1119\(84\)90130-6](https://doi.org/10.1016/0378-1119(84)90130-6).
- [26] JR Bradford, JA Siepen, and DR Westhead. "Chemical Structure of RNA." In: 3:3.7:96 (2005).
- [27] J Brophy and C Voigt. "Principles of genetic circuit design." In: *Nature Methods* 11:508–520 (2014).
- [28] DE Cameron, CJ Bashor, and JJ Collins. "A brief history of synthetic biology." In: *Nature Reviews of Microbiology* (2014), 12(5):381–90. DOI: 10.1038/nrmicro3239.
- [29] B Canton, A Labno, and D Endy. "Refinement and standardization of synthetic biological parts and devices". In: *Nature Biotechnology* 26(7):787-93 (2008). DOI: 10.1038/nbt1413.
- [30] G Cesareni, M Helmer-Citterich, and L Castagnoli. "Control of ColE1 plasmid replication by antisense RNA." In: *Trends in Genetics* 7(7):230-235 (1991). DOI: [https://doi.org/10.1016/0168-9525\(91\)90370-6](https://doi.org/10.1016/0168-9525(91)90370-6).

- [31] S Chakrabarti and R Sowdhamini. "Functional sites and evolutionary connections of acylhomoserine lactone synthases". In: *Protein Engineering* 16(4):271-8 (2003).
- [32] PP Cherepanov and W Wackernagel. "Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant". In: *Gene* 158.1 (1995), pp. 9–14. DOI: [https://doi.org/10.1016/0378-1119\(95\)00193-A](https://doi.org/10.1016/0378-1119(95)00193-A).
- [33] T Chien, A Doshi, and T Danino. "Advances in bacterial cancer therapies using synthetic biology". In: *Current Opinion in Systems Biology* 5 (2017), pp. 1–8. DOI: <https://doi.org/10.1016/j.coisb.2017.05.009>.
- [34] KH Choi, JB Gaynor, KG White, C Lopez, CM Bosio, RR Karkhoff-Schweizer, and HP Schweizer. "A Tn7-based broad-range bacterial cloning and expression system". In: *Nature Methods* 2(6):443-8 (2005).
- [35] S Clancy. "Chemical Structure of RNA." In: *Nature Education* 7(1):60 (2008).
- [36] JL. Connell, ET. Ritschdorff, M. Whiteley, and JB. Shear. "3D printing of microscopic bacterial communities." In: *Proceedings of the National Academy of Sciences* 110(46):18380–18385 (2013). DOI: [10.1073/pnas.1309729110](https://doi.org/10.1073/pnas.1309729110).
- [37] RS Cox, JA McLaughlin, R Grunberg, J Beal, A Wipat, and HM Sauro. "A Visual Language for Protein Design." In: *ACS Synthetic Biology* 6(7):1120–1123 (2017). DOI: [10.1021/acssynbio.6b00286](https://doi.org/10.1021/acssynbio.6b00286).
- [38] F Crick. "Central dogma of molecular biology." In: *Nature* 227(561) (1970).
- [39] FH Crick, L Barnett, S Brenner, and RJ Watts-Tobin. "General nature of the genetic code for proteins". In: *Nature* 192:1227-32 (1961).
- [40] R Daniel, JR Rubens, R Sarpeshkar, and TK Lu. "Synthetic analog computation in living cells". In: *Nature* 497(7451):619-23 (2013). DOI: [10.1038/nature12148](https://doi.org/10.1038/nature12148).
- [41] T Danino, O Mondragon-Palomino, L Tsimring, and J Hasty. "A synchronized quorum of genetic clocks". In: *Nature* 463(7279):326-330 (2010). DOI: [10.1038/nature08753](https://doi.org/10.1038/nature08753).
- [42] C Darwin. "The Origin of Species by Means of Natural Selection, or Preservation of Favoured Races in the Struggle for Life". In: *John Murray* (1859).
- [43] KA Datsenko and BL Wanner. "One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products". In: *Proceedings of the National Academy of Sciences* 97.12 (2000), pp. 6640–6645. DOI: [10.1073/pnas.120163297](https://doi.org/10.1073/pnas.120163297).

- [44] G Del Solar, R Giraldo, MJ Ruiz-Echevarría, M Espinosa, and R Díaz-Orejas. "Replication and Control of Circular Bacterial Plasmids." In: *Microbiology and Molecular Biology Reviews* 62(2):434-464 (1998).
- [45] U Deuschle, W Kammerer, R Gentz, and H Bujard. "Promoters of *Escherichia coli*: a hierarchy of in vivo strength indicates alternate structures". In: *The EMBO Journal* 5(11):2987-2994 (1986).
- [46] L Diambra, VR Senthivel, DB Menendez, and M Isalan. "Cooperativity To Increase Turing Pattern Space for Synthetic Biology". In: *ACS Synthetic Biology* 4.2 (2015), pp. 177–186. DOI: 10.1021/sb500233u.
- [47] GE Dilanji, JB Langebrake, P De Leenheer, and SJ Hagen. "Quorum Activation at a Distance: Spatiotemporal Patterns of Gene Regulation from Diffusion of an Autoinducer Signal". In: *Journal of the American Chemical Society* 134.12 (2012), pp. 5618–5626. DOI: 10.1021/ja211593q.
- [48] IB Dodd and JB Egan. "Action at a distance in CI repressor regulation of the bacteriophage 186 genetic switch". In: *Molecular microbiology* 45(3):697-710 (2002).
- [49] WJ Dower, JF Miller, and CW Ragsdale. "High efficiency transformation of *E. coli* by high voltage electroporation". In: *Nucleic Acids Research* 16(13):6127-6145 (1988). DOI: 10.1093/nar/16.13.6127.
- [50] E Dressaire and A Sauret. "Clogging of microfluidic systems". In: *Soft Matter* 13 (1 2017), pp. 37–48. DOI: 10.1039/C6SM01879C.
- [51] PV Dunlap. "Quorum regulation of luminescence in *Vibrio fischeri*." In: *Journal of Molecular Microbiology Biotechnologies* 1(1):5-12 (1999).
- [52] A Eberhard, AL Burlingame, C Eberhard, GL Kenyon, KH Nealson, and NJ Oppenheimer. "Structural identification of autoinducer of photobacterium *fischeriluciferase*." In: *Biochemistry* 20:2444-2449 (1981).
- [53] *Education in Microscopy and Digital Imaging*. <http://zeiss-campus.magnet.fsu.edu/articles/probes/anthozoafps.html>. Accessed: 2018-03-16.
- [54] AJ Egan, RM Cleverley, K Peters, RJ Lewis, and Vollmer W. "Regulation of bacterial cell wall growth." In: *FEBS Journal* 284(6):851-867 (2017).
- [55] J Ehrlich, QR Bartz, RM Smith, DA Joslyn, and PR Burkholder. "Chloromycetin, a New Antibiotic From a Soil Actinomycete." In: *Science* 106(2757):417 (1947).
- [56] MB Elowitz and S Leibler. "A synthetic oscillatory network of transcriptional regulators." In: *Nature* 403:335–338 (2000).

- [57] J Engebrecht, KH Neilson, and M Silverman. "Bacterial bioluminescence: isolation and genetic analysis of the functions from *Vibrio fischeri*." In: *Cell* 32:773-781 (1983).
- [58] C Engler, R Kandzia, and S Marillonnet. "A One Pot, One Step, Precision Cloning Method with High Throughput Capability". In: *PLoS ONE* 3(11):e3647 (2008). DOI: 10.1371/journal.pone.0003647.
- [59] RM Epanand, C Walker, RF Epanand, and NA Magarvey. "Molecular mechanisms of membrane targeting antibiotics". In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1858(5):980-987 (2016). DOI: <https://doi.org/10.1016/j.bbamem.2015.10.018>.
- [60] D Erickson. "Towards numerical prototyping of labs-on-chip: modeling for integrated microfluidic devices." In: *Microfluidics and Nanofluidics* 1(4):301-318 (2005).
- [61] "Essentials of Cell Biology: How Do Cells Decode Genetic Information into Functional Proteins?" In: *Nature Education* 2 (2014).
- [62] Y Fainman, D Psaltis, and C Yang. "Optofluidics: Fundamentals, Devices, and Applications. Basic Microfluidic and Soft Lithographic Techniques". In: *AccessEngineering* 2 (2010).
- [63] C Farago, P Hegedus, and R Ferenc. "The Impact of Version Control Operations on the Quality Change of the Source Code". In: *Computational Science and Its Applications* 8583 (2014).
- [64] M Fernandez, JA Hudson, R Korpela, and CG de los Reyes-Gavilan. "Impact on Human Health of Microorganisms Present in Fermented Dairy Products: An Overview." In: *Biomedical Research International* 412714 (2015). DOI: 10.1155/2015/412714.
- [65] MS Ferry, IA Razinkov, and J Hasty. *Chapter fourteen - Microfluidics for Synthetic Biology: From Design to Execution*. Vol. 497:295-372. Academic Press, 2011. DOI: <https://doi.org/10.1016/B978-0-12-385075-1.00014-7>.
- [66] RP Feynman. "There's plenty of room at the bottom. An invitation to enter a new field of physics." In: *Engineering and Science* (1960), pp. 22-36.
- [67] IJ Fijalkowska, RM Schaaper, and P Jonczyk. "DNA replication fidelity in *Escherichia coli*: a multi-DNA polymerase affair". In: *FEMS Microbiology Review* 36(6):1105-21 (2012).

- [68] PL Foster. "Stress-Induced Mutagenesis in Bacteria." In: *Critical reviews in biochemistry and molecular biology* 42(5):373-397 (2007). DOI: 10.1080/10409230701648494.
- [69] L Freedman, M Gibson, S Ethier, H Soule, R Neve, and Y Reid. "Reproducibility: changing the policies and culture of cell line authentication." In: *Nature Commentary* 12:493 (2015).
- [70] WC Fuqua and A Eberhard. "Signal generation in autoinduction systems: synthesis of acylated homoserine lactones by LuxI-type proteins." In: *Cell-Cell Signaling in Bacteria* 211-230 (1999).
- [71] WC Fuqua, S Winans, and EP Greenberg. "Quorum sensing in bacteria: The luxR-luxI family of cell density-responsive transcriptional regulators." In: *Journal of Bacteriology* 176:269-275 (1998).
- [72] P Galajda, J Keymer, P Chaikin, and R Austin. "A Wall of Funnel Concentrates Swimming Bacteria." In: *Journal of Bacteriology* 189(23):8704-8707 (2007). DOI: 10.1128/JB.01033-07.
- [73] M Galdzicki, KP Clancy, E Oberortner, M Pocock, JY Quinn, CA Rodriguez, N Roehner, ML Wilson, L Adam, JC Anderson, BA Bartley, J Beal, D Chandran, J Chen, D Densmore, D Endy, R Grünberg, J Hallinan, NJ Hillson, JD Johnson, A Kuchinsky, M Lux, G Misirli, J Peccoud, HA Plahar, E Sirin, GB Stan, A Villalobos, A Wipat, JH Gennari, CJ Myers, and HM Sauro. "The Synthetic Biology Open Language (SBOL) provides a community standard for communicating designs in synthetic biology." In: *Nature biotechnology* 32.6 (2014), pp. 545-50. DOI: 10.1038/nbt.2891.
- [74] WRGD Galloway, JT Hodgkinson, SD Bowden, M Welch, and DR Spring. "Quorum Sensing in Gram-Negative Bacteria: Small-Molecule Modulation of AHL and AI-2 Quorum Sensing Pathways". In: *Chemical Reviews* 111.1 (2011), pp. 28-67. DOI: 10.1021/cr100109t.
- [75] TC Galvao and V De Lorenzo. "Transcriptional regulators à la carte: engineering new effector specificities in bacterial regulatory proteins." In: *Current Opinion in Biotechnology* 17(1):34-42 (2006). DOI: <https://doi.org/10.1016/j.copbio.2005.12.002>.
- [76] *Gel electrophoresis*. <https://www.nature.com/scitable/definition/gel-electrophoresis-286>. Accessed: 2017-07-03.

- [77] SY Gerdes, MD Scholle, JW Campbell, G Balazsi, E Ravasz, MD Daugherty, AL Somera, NC Kyrpides, I Anderson, MS Gelfand, A Bhattacharya, V Kapatral, M D'Souza, MV Baev, Y Grechkin, F Mseeh, MY Fonstein, R Overbeek, AL Barabasi, ZN Oltvai, and AL Osterman. "Experimental Determination and System Level Analysis of Essential Genes in *Escherichia coli* MG1655". In: *Journal of Bacteriology* 185.19 (2003), pp. 5673–5684. DOI: 10.1128/JB.185.19.5673-5684.2003.
- [78] GD Geske, JC O'Neill, and HE Blackwell. "Expanding dialogues: from natural autoinducers to non-natural analogues that modulate quorum sensing in Gram-negative bacteria". In: *Chemical Society Reviews* 37 (7 2008), pp. 1432–1447. DOI: 10.1039/B703021P.
- [79] BR Glick. "Metabolic load and heterologous gene expression". In: *Biotechnology Advances* 13.2 (1995), pp. 247–261. DOI: [https://doi.org/10.1016/0734-9750\(95\)00004-A](https://doi.org/10.1016/0734-9750(95)00004-A).
- [80] AG Godin, B Lounis, and L Cognet. "Super-resolution Microscopy Approaches for Live Cell Imaging." In: *Biophysical Journal* 107(8):1777-1784 (2014). DOI: 10.1016/j.bpj.2014.08.028.
- [81] S Gottesman, E Roche, YN Zhou, and RT Sauer. "The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system." In: *Genes and Development* 12:1338-1347 (1998).
- [82] TA Gould, J Herman, J Krank, RC Murphy, and MEA Churchill. "Specificity of Acyl-Homoserine Lactone Synthases Examined by Mass Spectrometry". In: *Journal of Bacteriology* 188.2 (2006), pp. 773–783. DOI: 10.1128/JB.188.2.773-783.2006.
- [83] MR Green and J Sambrook. *Molecular Cloning: A Laboratory Manual*. Academic Press, 2012.
- [84] EP Greenberg. "Quorum sensing in gram-negative bacteria." In: *ASM News* 63:371-377 (1997).
- [85] F Grenier, D Matteau, V Baby, and S Rodrigue. "Complete Genome Sequence of *Escherichia coli* BW25113." In: *Genome Announcements* 2(5):e01038-14 (2014). DOI: 10.1128/genomeA.01038-14.
- [86] C Grilly, J Stricker, WL Pang, MR Bennett, and J Hasty. "A synthetic gene network for tuning protein degradation in *Saccharomyces cerevisiae*". In: *Molecular Systems Biology* 3.1 (2007), 127–n/a. DOI: 10.1038/msb4100168.

- [87] TM Gruber and CA Gross. "Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space." In: *Annual Review of Microbiology* 57(1):441-466 (2003). DOI: [10.1146/annurev.micro.57.030502.090913](https://doi.org/10.1146/annurev.micro.57.030502.090913).
- [88] L Guo, J Feng, Z Fang, J Xu, and X Lu. "Application of microfluidic lab-on-a-chip for the detection of mycotoxins in foods". In: *Trends in Food Science Technology* 46.2, Part A (2015), pp. 252–263. DOI: <https://doi.org/10.1016/j.tifs.2015.09.005>.
- [89] MT Guo, A Rotem, JA Heyman, and DA Weitz. "Droplet microfluidics for high-throughput biological assays". In: *Lab Chip* 12 (12 2012), pp. 2146–2155. DOI: [10.1039/C2LC21147E](https://doi.org/10.1039/C2LC21147E).
- [90] LM Guzman, D Belin, MJ Carson, and Beckwith J. "Tight Regulation, Modulation, and High-Level Expression by Vectors Containing the Arabinose PBAD Promoter." In: *Journal of Bacteriology* 177(14):4121–4130 (1995).
- [91] A Haldimann, LL Daniels, and BL Wanner. "Use of New Methods for Construction of Tightly Regulated Arabinose and Rhamnose Promoter Fusions in Studies of the *Escherichia coli* Phosphate Regulon". In: *Journal of Bacteriology* 180.5 (1998), pp. 1277–1286.
- [92] S Halldorsson, E Lucumi, R Gómez-Sjöberg, and RMT Fleming. "Advantages and challenges of microfluidic cell culture in polydimethylsiloxane devices." In: *Biosensors and Bioelectronics* 63:218-231 (2015). DOI: <https://doi.org/10.1016/j.bios.2014.07.029>.
- [93] D Hanahan. "Studies on transformation of *Escherichia coli* with plasmids." In: *Journal of Molecular Biology* 166(4):557-580 (1983). DOI: [https://doi.org/10.1016/S0022-2836\(83\)80284-8](https://doi.org/10.1016/S0022-2836(83)80284-8).
- [94] BL Hanzelka, MR Parsek, DL Val, PV Dunlap, JE Cronan, and EP Greenberg. "Acylhomoserine Lactone Synthase Activity of the *Vibrio fischeri* AinS Protein". In: *Journal of Bacteriology* 181.18 (1999), pp. 5766–5770.
- [95] J Hasty, D McMillen, and JJ Collins. "Engineered gene circuits." In: *Nature* 420:224–230 (2002). DOI: [10.1038/nature01257](https://doi.org/10.1038/nature01257).
- [96] M Havlikova, M Jirgl, and Z Bradac. "Human Reliability in Man-machine Systems." In: *Procedia Engineering* 100:1207-1214 (2015). DOI: <https://doi.org/10.1016/j.proeng.2015.01.485>.

- [97] R Hazan, YA Que, D Maura, and LG Rahme. "A method for high throughput determination of viable bacteria cell counts in 96-well plates." In: *BMC Microbiology* 12(1):259 (2012). DOI: 10.1186/1471-2180-12-259.
- [98] JM Heather and B Chain. "The sequence of sequencers: The history of sequencing DNA." In: *Genomics* 107(1):1-8 (2016). DOI: 10.1016/j.ygeno.2015.11.003.
- [99] S Heeb, MP Fletcher, SR Chhabra, SP Diggle, P Williams, and M Camara. "Quinolones: from antibiotics to autoinducers". In: *Fems Microbiology Reviews* 35(2):247-274 (2011). DOI: 10.1111/j.1574-6976.2010.00247.x.
- [100] L Heinz and M Walter. "Promoters in the *E. coli* replication origin." In: *Nature* 294:376-378 (1981).
- [101] H. Hillborg, JF. Ankner, UW. Gedde, GD. Smith, HK. Yasuda, and K. Wikstrom. "Crosslinked polydimethylsiloxane exposed to oxygen plasma studied by neutron reflectometry and other surface specific techniques." In: *Polymer* 18(41):6851-6863 (2000).
- [102] CRT Hillyar. "Genetic recombination in bacteriophage lambda". In: *Bioscience Horizons: The International Journal of Student Research* 5 (2012), hzs001. DOI: 10.1093/biohorizons/hzs001.
- [103] TT Hoang, R Ann, RR Karkhoff-Schweizer, AJ Kutchma, and HP Schweizer. "A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants". In: *Gene* 212.1 (1998), pp. 77-86. DOI: [https://doi.org/10.1016/S0378-1119\(98\)00130-9](https://doi.org/10.1016/S0378-1119(98)00130-9).
- [104] BP Hodkinson and EA Grice. "Next-Generation Sequencing: A Review of Technologies and Tools for Wound Microbiome Research." In: *Advances in Wound Care* 4(1):50-58 (2015). DOI: 10.1089/wound.2014.0542.
- [105] IG Hook-Barnard and DM Hinton. "Transcription Initiation by Mix and Match Elements: Flexibility for Polymerase Binding to Bacterial Promoters." In: *Gene Regulation Systems Biology* 1:275-293 (2007).
- [106] RM Horton, Z Cai, SN Ho, and LR Pease. "Gene Splicing by Overlap Extension: Tailor-Made Genes Using the Polymerase Chain Reaction." In: *BioTechniques* 8(5):528-535 (1990).
- [107] PA Hoskisson and G Hobbs. "Continuous culture – making a comeback?" In: *Microbiology* 151.10 (2005), pp. 3153-3159.

- [108] L Hostettler, L Grundy, S Käser-Pébernard, C Wicky, WR Schafer, and DA Glauser. "The Bright Fluorescent Protein mNeonGreen Facilitates Protein Expression Analysis *In Vivo*." In: *Genes, Genomes and Genetics* 7(2):607-615 (2017).
- [109] *How to create a wiki page?* https://en.wikipedia.org/wiki/Wikipedia:How_to_create_a_page. Accessed: 2019-01-12.
- [110] SE Hulme, WR DiLuzio, SS Shevkoplyas, L Turner, M Mayer, HC Bergc, and GM Whitesides. "Using ratchets and sorters to fractionate motile cells of *Escherichia coli* by length." In: *Lab on a chip* 11 (2008).
- [111] CA Hutchison, RY Chuang, VN Noskov, N Assad-Garcia, TJ Deerinck, MH Ellisman, J Gill, K Kannan, BJ Karas, L Ma, JF Pelletier, ZQ Qi, RA Richter, EA Strychalski, L Sun, Y Suzuki, B Tsvetanova, KS Wise, HO Smith, JI Glass, C Merryman, DG Gibson, and JC Venter. "Design and synthesis of a minimal bacterial genome." In: *Science* 351(6280) (2016). DOI: 10.1126/science.aad6253.
- [112] H Hutter. "Chapter 3 - Fluorescent Protein Methods: Strategies and Applications". In: 107 (2012). Ed. by Joel H. Rothman and Andrew Singson, pp. 67–92. DOI: <https://doi.org/10.1016/B978-0-12-394620-1.00003-5>.
- [113] BR Jack, SP Leonard, DM Mishler, BA Renda, D Leon, GA Suarez, and JE Barrick. "Predicting the Genetic Stability of Engineered DNA Sequences with the EFM Calculator". In: *ACS Synthetic Biology* 4.8 (2015), pp. 939–943. DOI: 10.1021/acssynbio.5b00068.
- [114] D Jackson and B Launder. "Osborne Reynolds and the Publication of His Papers on Turbulent Flow." In: *Annual Reviews of Fluid Mechanics* 39:19–35 (2007). DOI: 10.1146/annurev.fluid.39.050905.110241.
- [115] J Janda and S Abbott. "16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls." In: *Journal of Clinical Microbiology* 45:9 (2007).
- [116] M Juhas, DR Reuss, B Zhu, and FM Commichau. "Bacillus subtilis and *Escherichia coli* essential genes and minimal cell factories after one decade of genome engineering." In: *Microbiology* 160.11 (2014), pp. 2341–2351.
- [117] M Kaczanowska and M Rydén-Aulin. "Ribosome Biogenesis and the Translation Process in *Escherichia coli*." In: *Microbiology and Molecular Biology Reviews* 71(3):477-494 (2007). DOI: 10.1128/MMBR.00013-07.
- [118] JA Kahana and PA Silver. *Use of the A. Victoria Green Fluorescent Protein to Study Protein Dynamics In Vivo*. 2001. DOI: 10.1002/0471141755.ph0606s05.

- [119] P Kanigowska, Y Shen, Y Zheng, S Rosser, and Y Cai. "Smart DNA Fabrication Using Sound Waves: Applying Acoustic Dispensing Technologies to Synthetic Biology". In: *Journal of Laboratory Automation* 21.1 (2016), pp. 49–56. DOI: 10.1177/2211068215593754.
- [120] HB Kaplan and EP Greenberg. "Diffusion of autoinducer is involved in regulation of the *Vibrio fischeri* luminescence system." In: *Journal of Bacteriology* 163(3):1210-1214 (1985).
- [121] J Kato and M Hashimoto. "Construction of consecutive deletions of the *Escherichia coli* chromosome." In: *Molecular Systems Biology* 3.1 (2007), 132–n/a. DOI: 10.1038/msb4100174.
- [122] SH Kim. "Mechanisms of DNA Recombination and Genome Rearrangements: Methods to Study Homologous Recombination". In: *Methods in Enzymology* 600 (2018), pp. 233–253. DOI: <https://doi.org/10.1016/bs.mie.2017.11.012>.
- [123] SH Kim, AM Cavaleiro, M Rennig, and MHH Nørholm. "SEVA Linkers: A Versatile and Automatable DNA Backbone Exchange Standard for Synthetic Biology". In: *ACS Synthetic Biology* 5.10 (2016), pp. 1177–1181. DOI: 10.1021/acssynbio.5b00257.
- [124] TF Knight Jr and GJ Sussman. "Cellular gate technology." In: *Proceedings of UCM98* 257-272 (1998).
- [125] K Kobayashi, SD Ehrlich, A Albertini, G Amati, KK Andersen, M Arnaud, K Asai, S Ashikaga, S Aymerich, P Bessieres, F Boland, SC Brignell, S Bron, K Bunai, J Chapuis, LC Christiansen, A Danchin, M Débarbouillé, E Dervyn, E Deuerling, K Devine, SK Devine, O Dreesen, J Errington, S Fillinger, SJ Foster, Y Fujita, A Galizzi, R Gardan, C Eschevins, T Fukushima, K Haga, CR Harwood, M Hecker, D Hosoya, MF Hullo, H Kakeshita, D Karamata, Y Kasahara, F Kawamura, K Koga, P Koski, R Kuwana, D Imamura, M Ishimaru, S Ishikawa, I Ishio, D Le Coq, A Masson, C Mauël, R Meima, RP Mellado, A Moir, S Moriya, E Nagakawa, H Nanamiya, S Nakai, P Nygaard, M Ogura, T Ohanan, M O'Reilly, M O'Rourke, Z Pragai, HM Pooley, G Rapoport, JP Rawlins, LA Rivas, C Rivolta, A Sadaie, Y Sadaie, M Sarvas, T Sato, HH Saxild, E Scanlan, W Schumann, JFML Seegers, J Sekiguchi, A Sekowska, SJ Séror, M Simon, P Stragier, R Studer, H Takamatsu, T Tanaka, M Takeuchi, HB Thomaidis,

- V Vagner, JM van Dijl, K Watabe, A Wipat, H Yamamoto, M Yamamoto, Y Yamamoto, K Yamane, K Yata, K Yoshida, H Yoshikawa, U Zuber, and N Ogasawara. "Essential *Bacillus subtilis* genes". In: *Proceedings of the National Academy of Sciences* 100.8 (2003), pp. 4678–4683. DOI: 10.1073/pnas.0730515100.
- [126] B Koch and O Nybroe. "Initial characterization of a bolA homologue from *Pseudomonas fluorescens* indicates different roles for BolA-like proteins in *P. fluorescens* and *Escherichia coli*". In: *FEMS Microbiology Letters* 262.1 (2006), pp. 48–56. DOI: 10.1111/j.1574-6968.2006.00359.x.
- [127] MA Kohanski, DJ Dwyer, and JJ Collins. "How antibiotics kill bacteria: from targets to networks." In: *Nature Reviews Microbiology* 8(6):423–435 (2010).
- [128] S Konur, M Gheorghe, C Dragomir, L Mierla, F Ipate, and N Krasnogor. "Qualitative and quantitative analysis of systems and synthetic biology constructs using P systems." In: *ACS Synthetic Biology* 4(1):83-92 (2015). DOI: doi : 10.1021/sb500134w.
- [129] J Kozyra, A Ceccarelli, E Torelli, A Lopiccolo, JY Gu, H Fellermann, U Stimming, and N Krasnogor. "Designing Uniquely Addressable Bio-orthogonal Synthetic Scaffolds for DNA and RNA Origami." In: *ACS Synthetic Biology* 6(7):1140–1149 (2017).
- [130] TA Kunkel. "DNA Replication Fidelity". In: *Journal of Biological Chemistry* 279.17 (2004), pp. 16895–16898. DOI: 10.1074/jbc.R400006200.
- [131] N Kupp. "L-Edit: quick start guide". In: (2006).
- [132] E Kutter. *Concatemer (Genomes)*. 2001, pp. 435–436.
- [133] BH Kvitko, S Bruckbauer, J Prucha, I McMillan, EJ Breland, S Lehman, K Mladinich, KH Choi, RA Karkhoff-Schweizer, and HP Schweizer. "A simple method for construction of pir+ Enterobacterial hosts for maintenance of R6K replicon plasmids". In: *BMC Research Notes* 5.1 (2012), p. 157. DOI: 10.1186/1756-0500-5-157.
- [134] ES Lander. "The heroes of CRISPR". In: *Cell* 164 (2016).
- [135] TE Landrain, J Carrera, B Kirov, G Rodrigo, and A Jaramillo. "Modular model-based design for heterologous bioproduction in bacteria". In: *Current Opinion in Biotechnology* 20.3 (2009), pp. 272–279. DOI: <https://doi.org/10.1016/j.copbio.2009.06.003>.

- [136] R Lange and R Hengge-Aronis. "Growth phase-regulated expression of *bolA* and morphology of stationary-phase *Escherichia coli* cells are controlled by the novel sigma factor sigma S." In: *Journal of Bacteriology* 173.14 (1991), pp. 4474–4481. DOI: 10.1128/jb.173.14.4474-4481.1991.
- [137] JB Langebrake, GE Dilanji, SJ Hagen, and P De Leenheer. "Traveling waves in response to a diffusing quorum sensing signal in spatially-extended bacterial colonies". In: *Journal of Theoretical Biology* 363 (2014), pp. 53 –61. DOI: <https://doi.org/10.1016/j.jtbi.2014.07.033>.
- [138] N Lapique and Y Benenson. "Digital switching in a biosensor circuit via programmable timing of gene availability." In: *Nature Chemical Biology* 10:1020–1027 (2014).
- [139] MH Larson, WJ Greenleaf, R Landick, and SM Block. "Applied Force Reveals Mechanistic and Energetic Details of Transcription Termination." In: *Cell* 132(6):971–982 (2008).
- [140] J Lee, J Wu, Y Deng, J Wang, C Wang, J Wang, C Chang, Y Dong, P Williams, and LH Zhang. "A cell-cell communication signal integrates quorum sensing and stress response". In: *Nature Chemistry Biology* 9(5):339–43.1 (2013), pp. 261–286. DOI: 10.1038/nchembio.1225.
- [141] L Lesueur, L Mir, and F André. "Overcoming the Specific Toxicity of Large Plasmids Electrotransfer in Primary Cells *In Vitro*." In: *Molecular Therapy - Nucleic Acids* 5(3):e291 (2016).
- [142] JC. Lotters, O. Wouter, PH. Veltink, and P. Bergveld. "The mechanical properties of the rubber elastic polymer polydimethylsiloxane for sensor applications." In: *Journal of micromechanics and microengineering* 7(3):145–147 (1997). DOI: 10.1088/0960-1317/7/3/017.
- [143] HD Lu, AC Spiegel, A Hurley, LJ Perez, K Maisel, LM Ensign, J Hanes, BL Bassler, MF Semmelhack, and RK Prud'homme. "Modulating *Vibrio cholerae* Quorum-Sensing-Controlled Communication Using Autoinducer-Loaded Nanoparticles". In: *Nano Letters* 15.4 (2015), pp. 2235–2241. DOI: 10.1021/acs.nanolett.5b00151.
- [144] R Lutz and H Bujard. "Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements". In: *Nucleic Acids Research* 25(6):1203–1210 (1997).
- [145] *Lux promoter backwards transcription*. http://parts.igem.org/Part:BBa_R0062. Accessed: 2018-12-23.

- [146] E Maguin, H Prevost, SD Ehrlich, and A Gruss. "Efficient insertional mutagenesis in *lactococci* and other gram-positive bacteria." In: *Journal of Bacteriology* 178.3 (1996), pp. 931–5. DOI: 10.1128/jb.178.3.931-935.1996.
- [147] RM Maier. *Chapter 3 - Bacterial Growth*. Second Edition. 2009, pp. 37–54. DOI: <https://doi.org/10.1016/B978-0-12-370519-8.00003-1>.
- [148] N Marchand and CH Collins. "Synthetic Quorum Sensing and Cell–Cell Communication in Gram-Positive *Bacillus megaterium*". In: *ACS Synthetic Biology* 5.7 (2016), pp. 597–606. DOI: 10.1021/acssynbio.5b00099.
- [149] E Martinez-Garcia, T Aparicio, A Goni-Moreno, S Fraile, and V de Lorenzo. "SEVA 2.0: an update of the Standard European Vector Architecture for de-/reconstruction of bacterial functionalities". In: *Nucleic Acids Research* 43.D1 (2015), pp. D1183–D1189. DOI: 10.1093/nar/gku1114.
- [150] A Mata, C Boehm, AJ Fleischman, G Muschler, and S Roy. "Analysis of Connective Tissue Progenitor Cell Behavior on Polydimethylsiloxane Smooth and Channel Micro-Textures." In: *Biomedical microdevices* 4(4):267-275 (2002).
- [151] JC. McDonald, DC. Duffy, JR. Anderson, DT. Chiu, H. Wu, OJ. Schueller, and GM. Whitesides. "Fabrication of microfluidic systems in poly(dimethylsiloxane)." In: *Electrophoresis* 21(1):27-40 (2000).
- [152] DB Menendez, VR Senthivel, and M Isalan. "Sender–receiver systems and applying information theory for quantitative synthetic biology". In: *Current Opinion in Biotechnology* 31 (2015), pp. 101–107. DOI: <https://doi.org/10.1016/j.copbio.2014.08.005>.
- [153] K Michalodimitrakis and M Isalan. "Engineering prokaryotic gene circuits". In: *FEMS Microbiology Reviews* 33.1 (2009), pp. 27–37. DOI: 10.1111/j.1574-6976.2008.00139.x.
- [154] *Microfluidics tubing and connectors*. <https://www.elveflow.com/microfluidic-tutorials/microfluidic-reviews-and-tutorials/microfluidic-fittings-and-tubing-resources/the-basics-of-microfluidic-tubing-sleeves>. Accessed: 2018-12-23.
- [155] MB Miller and BL Bassler. "Quorum Sensing in Bacteria." In: *Annual Review of Microbiology* 55(1):165-199 (2001). DOI: 10.1146/annurev.micro.55.1.165.

- [156] M Mimee, AC Tucker, CA Voigt, and TK Lu. "Programming a Human Commensal Bacterium, *Bacteroides thetaiotaomicron*, to Sense and Respond to Stimuli in the Murine Gut Microbiota". In: *Cell systems* 1(1):62-7 (2015). DOI: 10.1016/j.cels.2015.06.001.
- [157] JR Moffitt, JB Lee, and P Cluzel. "The single-cell chemostat: an agarose-based, microfluidic device for high-throughput, single-cell studies of bacteria and bacterial communities." In: *Lab on a chip* 12(8):1487-1494 (2012). DOI: 10.1039/c2lc00009a.
- [158] J Monod. "The Growth of Bacterial Cultures." In: *Annual Reviews of Microbiology* 3:371-394 (1949).
- [159] J Monod and F Jacob. "General Conclusions: Teleonomic Mechanisms in Cellular Metabolism, Growth, and Differentiation." In: *Cold Spring Harbor Symposia on Quantitative Biology* (1961), 26:389-401.
- [160] SJ Moore, HE Lai, RJR Kelwick, SM Chee, DJ Bell, KM Polizzi, and PS Freemont. "EcoFlex: A Multifunctional MoClo Kit for *E. coli* Synthetic Biology". In: *ACS Synthetic Biology* 5.10 (2016), pp. 1059-1069. DOI: 10.1021/acssynbio.6b00031.
- [161] T Morimoto, K Ara, K Ozaki, and N Ogasawara. "A new simple method to introduce marker-free deletions in the *Bacillus subtilis* genome". In: *Genes and Genetic Systems* 84.4 (2009), pp. 315-318. DOI: 10.1266/ggs.84.315.
- [162] T Morimoto, K Ara, K Ozaki, and N Ogasawara. "A Simple Method for Introducing Marker-Free Deletions in the *Bacillus subtilis* genome". In: *Methods in Molecular Biology (Methods and Protocols)* 765 (2011).
- [163] KB Mullis. "The unusual origin of the polymerase chain reaction." In: *Sci Am* 262(4):56-61 (1990).
- [164] A Munteanu, M Constante, M Isalan, and RV Solé. "Avoiding transcription factor competition at promoter level increases the chances of obtaining oscillation". In: *BMC Systems Biology* 4.1 (2010), p. 66. DOI: 10.1186/1752-0509-4-66.
- [165] H Murray. "Connecting chromosome replication with cell growth in bacteria." In: *Current Opinion in Microbiology* 34:13-17 (2016). DOI: <https://doi.org/10.1016/j.mib.2016.07.013>.
- [166] H Murray and J Errington. "Dynamic Control of the DNA Replication Initiation Protein DnaA by Soj/ParA." In: *Cell* 135(1):74-84 (2008). DOI: 10.1016/j.cell.2008.07.044.

- [167] N Nandagopal and MB Elowitz. "Synthetic biology: integrated gene circuits." In: *Science* 333(6047):1244-8 (2011). DOI: doi:10.1126/science.1207084.
- [168] NASA publications. <https://gameon.nasa.gov/publications/>. Accessed: 2018-03-07.
- [169] Jonathan Naylor, Harold Fellermann, Yuchun Ding, Waleed K. Mohammed, Nicholas S. Jakubovics, Joy Mukherjee, Catherine A. Biggs, Phillip C. Wright, and Natalio Krasnogor. "Simbiotics: A Multiscale Integrative Platform for 3D Modeling of Bacterial Populations". In: *ACS Synthetic Biology* 6.7 (July 2017), pp. 1194–1210. DOI: 10.1021/acssynbio.6b00315.
- [170] KS Nilgiriwala, J Jimenez, PM Rivera, and D Del Vecchio. "Synthetic Tunable Amplifying Buffer Circuit in *E. coli*". In: *ACS Synthetic Biology* 4.5 (2015), pp. 577–584. DOI: 10.1021/sb5002533.
- [171] K Nishimura, S Tsuru, H Suzuki, and T Yomo. "Stochasticity in Gene Expression in a Cell-Sized Compartment". In: *ACS Synthetic Biology* 4.5 (2015), pp. 566–576. DOI: 10.1021/sb500249g.
- [172] RP Novick. "Plasmid Incompatibility." In: *Microbiological Reviews* 51(4):381-395 (1987).
- [173] PH Oliveira, F Lemos, GA Monteiro, and DMF Prazeres. "Recombination frequency in plasmid DNA containing direct repeats—predictive correlation with repeat and intervening sequence length". In: *Plasmid* 60.2 (2008), pp. 159–165. DOI: <https://doi.org/10.1016/j.plasmid.2008.06.004>.
- [174] A Orosz, I Boros, and P Venetianer. "Analysis of the complex transcription termination region of the *Escherichia coli* rrnB gene." In: *FEBS* 201(3):653–659 (1991). DOI: 10.1111/j.1432-1033.1991.tb16326.x.
- [175] G O'Toole, HB Kaplan, and R Kolter. "Biofilm Formation as Microbial Development". In: *Annual Review of Microbiology* 54.1 (2000), pp. 49–79. DOI: 10.1146/annurev.micro.54.1.49.
- [176] EM Ozbudak. "Noise and Multistability in Gene Regulatory Networks." In: *Thesis* (2004).
- [177] S Park, PM Wolanin, EA Yuzbashyan, H Lin, NC Darnton, JB Stock, P Silberzan, and R Austin. "Influence of topology on bacterial social interaction". In: *Proceedings of the National Academy of Sciences* 100.24 (2003), pp. 13910–13915. DOI: 10.1073/pnas.1935975100.

- [178] JC Pierce, D Kong, and W Masker. "The effect of the length of direct repeats and the presence of palindromes on deletion between directly repeated DNA sequences in bacteriophage T7". In: *Nucleic Acids Research* 19(14):3901-3905 (1991). DOI: 10.1038/msb4100050.
- [179] A Piruska, S Branagan, DM Cropek, JV Sweedler, and PW Bohn. "Electrokinetically driven fluidic transport in integrated three-dimensional microfluidic devices incorporating gold-coated nanocapillary array membranes." In: *Lab On A Chip* 8(10):1625-1631 (2008). DOI: 10.1039/B805768K.
- [180] G Pósfai, G Plunkett, T Fehér, D Frisch, GM Keil, K Umenhoffer, V Kolisnychenko, B Stahl, SS Sharma, M de Arruda, V Burland, SW Harcum, and FR Blattner. "Emergent Properties of Reduced-Genome *Escherichia coli*". In: *Science* 312.5776 (2006), pp. 1044–1046. DOI: 10.1126/science.1126439.
- [181] L Pray. "Discovery of DNA Structure and Function: Watson and Crick." In: *Nature Education* 1(1):100 (2008).
- [182] C Probst, A Grunberger, W Wiechert, and D Kohlheyer. "Microfluidic growth chambers with optical tweezers for full spatial single-cell control and analysis of evolving microbes". In: *Journal of Microbiological Methods* 95.3 (2013), pp. 470–476. DOI: <https://doi.org/10.1016/j.mimet.2013.09.002>.
- [183] J Radeck, K Kraft, J Bartels, T Cikovic, F Durr, J Emenegger, S Kelterborn, C Sauer, G Fritz, S Gebhard, and T Mascher. "The *Bacillus* BioBrick Box: generation and evaluation of essential genetic building blocks for standardized work with *Bacillus subtilis*". In: *Journal of Biological Engineering* 7.1 (2013), p. 29. DOI: 10.1186/1754-1611-7-29.
- [184] PH Rampelotto. "Extremophiles and Extreme Environments". In: *Life: Open Access Journal* 3(3):482-485 (2013). DOI: 10.3390/life3030482.
- [185] A Randall, P Guye, S Gupta, X Duportet, and R Weiss. "Design and connection of robust genetic circuits." In: *Methods in enzymology* 497:159–86 (2011). DOI: doi:10.1016/B978-0-12-385075-1.00007-X.
- [186] AZ Reeves, WE Spears, J Du, KY Tan, AJ Wagers, and CF Lesser. "Engineering *Escherichia coli* into a Protein Delivery System for Mammalian Cells". In: *ACS Synthetic Biology* 4.5 (2015), pp. 644–654. DOI: 10.1021/acssynbio.5b00002.
- [187] *Registry of standard biological parts plasmid backbones*. http://parts.igem.org/Plasmid_backbones. Accessed: 2018-03-16.

- [188] S Regot, J Macia, N Conde, K Furukawa, J Kjellen, T Peeters, S Hohmann, E de Nadal, F Posas, and R Sole. "Distributed biological computation with multicellular engineered networks". In: *Nature* 469(7329):207-11 (2011). DOI: 10.1038/nature09679.
- [189] B Reller, M Weinstein, and C Petti. "Detection and identification of microorganisms by gene amplification and sequencing." In: *Clinical Infectious Diseases* 44:8 (2007).
- [190] SJ. Remington. "Fluorescent proteins: maturation, photochemistry and photophysics." In: *Current Opinion in Structural Biology* 16(6):714-721 (2006). DOI: <https://doi.org/10.1016/j.sbi.2006.10.001>.
- [191] C Ren, Q Yan, and Z Zhang. "Minimum length of direct repeat sequences required for efficient homologous recombination induced by zinc finger nuclease in yeast". In: *Molecular Biology Reports* 41.10 (2014), pp. 6939–6948. DOI: 10.1007/s11033-014-3579-6.
- [192] TT Richardson, O Harran, and H Murray. "The bacterial DnaA-trio replication origin element specifies single-stranded DNA initiator binding." In: *Nature* 534:412–416 (2016).
- [193] G Rodrigo, B Kirov, S Shen, and A Jaramillo. "Theoretical and experimental analysis of the forced LacI-AraC oscillator with a minimal gene regulatory model". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 23.2 (2013), pp. 025109–025109.
- [194] N Roehner, J Beal, K Clancy, B Bartley, G Misirli, R Grunberg, E Oberortner, M Pocock, M Bissell, C Madsen, T Nguyen, M Zhang, Zh Zhang, Za Zundel, D Densmore, JH Gennari, A Wipat, HM Sauro, and CJ Myers. "Sharing Structure and Function in Biological Design with SBOL 2.0." In: *ACS Synthetic Biology* 5(6):498–506 (2016). DOI: 10.1021/acssynbio.5b00215.
- [195] RE Rose. "The nucleotide sequence of pACYC184". In: *Nucleic Acids Research* 16(1):355 (1988).
- [196] A Rotem, O Ram, N Shores, RA Sperling, A Goren, DA Weitz, and BE Bernstein. "Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state". In: *Nature Biotechnology* 33(11):1165-72 (2015). DOI: 10.1038/nbt.3383.
- [197] EG Ruby and KH Lee. "The *Vibrio fischeri*-*Euprymna scolopes* Light Organ Association: Current Ecological Paradigms." In: *Applied and Environmental Microbiology* 64(3):805-812 (1998).

- [198] EG Ruby and KH Neelson. "Symbiotic association of photobacterium *fischeri* with the marine luminous fish *Monocentris japonica*: a model of symbiosis based on bacterial studies." In: *The Biological Bulletin* 151:574-586 (1976).
- [199] R Rusconi, M Garren, and R Stocker. "Microfluidics Expanding the Frontiers of Microbial Ecology." In: *Annual review of biophysics* 43:65-91 (2014). DOI: 10.1146/annurev-biophys-051013-022916.
- [200] HM Salis, EA Mirsky, and CA Voigt. "Automated Design of Synthetic Ribosome Binding Sites to Precisely Control Protein Expression." In: *Nature Biotechnology* 27(10):946-950 (2009). DOI: 10.1038/nbt.1568.
- [201] F Sanger, S Nicklen, and AR Coulson. "DNA sequencing with chain-terminating inhibitors." In: *PNAS* 74(12):5463-5467 (1977).
- [202] *SBOL: The Synthetic Biology Open Language*. <http://sbolstandard.org/>. Accessed: 2018-01-16.
- [203] Y Schaerli, A Munteanu, M Gili, J Cotterell, J Sharpe, and M Isalan. "A unified design space of synthetic stripe-forming networks". In: *Nature Communications* 5:4905 (2014). DOI: 10.1038/ncomms5905.
- [204] R Schleif. "AraC protein, regulation of the l-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action." In: *FEMS Microbiology Reviews* 34(5):779-796 (2010). DOI: 10.1111/j.1574-6976.2010.00226.x.
- [205] Particle Sciences. "Protein structure." In: *Technical brief* 8 (2009).
- [206] SR Scott and J Hasty. "Quorum Sensing Communication Modules for Microbial Consortia". In: *ACS Synthetic Biology* 5.9 (2016), pp. 969-977. DOI: 10.1021/acssynbio.5b00286.
- [207] G Selzer and JI Tomizawa. "Specific cleavage of the p15A primer precursor by ribonuclease H at the origin of DNA replication." In: *PNAS* 79(23):7082-7086 (1982).
- [208] H Senn, U Lendenmann, M Snozzi, G Hamer, and T Egli. "The growth of *Escherichia coli* in glucose-limited chemostat cultures: a re-examination of the kinetics." In: *Biochimica et Biophysica Acta* 1201(3):424-436 (1994). DOI: [https://doi.org/10.1016/0304-4165\(94\)90072-8](https://doi.org/10.1016/0304-4165(94)90072-8).
- [209] GS Shadel, JH Devine, and Baldwin TO. "Control of the lux regulon of *Vibrio fischeri*." In: *Journal of Bioluminescence and Chemiluminescence* 5(2):99-106 (1990).
- [210] N Shaner, P Steinbach, and R Tsien. "A guide to choosing fluorescent proteins." In: *Nature Methods* 2:905-909 (2005).

- [211] SCC Shih, G Goyal, PW Kim, N Koutsoubelis, JD Keasling, PD Adams, NJ Hillson, and AK Singh. "A Versatile Microfluidic Device for Automating Synthetic Biology". In: *ACS Synthetic Biology* 4.10 (2015), pp. 1151–1164. DOI: 10.1021/acssynbio.5b00062.
- [212] T Shimada, Y Yamazaki, K Tanaka, and A Ishihama. "The Whole Set of Constitutive Promoters Recognized by RNA Polymerase RpoD Holoenzyme of *Escherichia coli*." In: *PLOS ONE* 9(6):e100908 (2014). DOI: 10.1371/journal.pone.0090447.
- [213] BA Shirley. "Protein Stability and Folding: Theory and Practice." In: *Methods in Molecular Biology*. 40 (1995).
- [214] SK Sia and GM Whitesides. "Microfluidic devices fabricated in Poly(dimethylsiloxane) for biological studies". In: *Electrophoresis* 24.21 (2003), pp. 3563–3576. DOI: 10.1002/elps.200305584.
- [215] MH Sibley and EA Raleigh. "A versatile element for gene addition in bacterial chromosomes". In: *Nucleic Acids Research* 40.3 (2012), e19. DOI: 10.1093/nar/gkr1085.
- [216] M Sim, S Koirala, D Picton, H Strahl, PA Hoskisson, CV Rao, CS Gillespie, and PD Aldridge. "Growth rate control of flagellar assembly in *Escherichia coli* strain RP437". In: *Scientific Reports* 7: 4118 (2017). DOI: 10.1038/srep41189.
- [217] Hagen SJ, M Son, JT Weiss, and JH Young. "Bacterium in a box: sensing of quorum and environment by the LuxI/LuxR gene regulatory circuit". In: *Journal of Biological Physics* 36(3):317–327 (2010). DOI: 10.1007/s10867-010-9186-4.
- [218] AL Slusarczyk, A Lin, and R Weiss. "Foundations for the design and implementation of synthetic genetic circuits." In: *Nature Reviews Genetics* 13(6):406–20 (2012). DOI: doi:10.1038/nrg3227.
- [219] AM Stevens and EP Greenberg. "Transcriptional Activation by LuxR." In: *Cell-Cell Signaling in Bacteria* (1999).
- [220] J Stricker, S Cookson, MR Bennett, WH Mather, LS Tsimring, and J Hasty. "A fast, robust and tunable synthetic gene oscillator." In: *Nature* 456(7221):516–9 (2008). DOI: 10.1038/nature07389.
- [221] M Strong. "Protein Nanomachines". In: *PLOS Biology* 2.3 (2004). DOI: 10.1371/journal.pbio.0020073.

- [222] SA Sundberg. "High-throughput and ultra-high-throughput screening: solution- and cell-based approaches." In: *Current Opinion in Biotechnology* 11(1):47-53 (2000). DOI: [https://doi.org/10.1016/S0958-1669\(99\)00051-8](https://doi.org/10.1016/S0958-1669(99)00051-8).
- [223] A Tamsir, JJ Tabor, and CA Voigt. "Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'." In: *Nature* 469(7329):212-5 (2011). DOI: [10.1038/nature09565](https://doi.org/10.1038/nature09565).
- [224] K Tanaka, CS Henry, JF Zinner, E Jolivet, MP Cohoon, F Xia, V Bidnenko, SD Ehrlich, RL Stevens, and P Noiro. "Building the repertoire of dispensable chromosome regions in *Bacillus subtilis* entails major refinement of cognate large-scale metabolic model." In: *Nucleic Acids Research* 41.1 (2013), pp. 687-699. DOI: [10.1093/nar/gks963](https://doi.org/10.1093/nar/gks963).
- [225] MA TerAvest, Z Li, and LT Angenent. "Bacteria-based biocomputing with Cellular Computing Circuits to sense, decide, signal, and act". In: *Energy and Environmental Science* 4 (12 2011), pp. 4907-4916. DOI: [10.1039/C1EE02455H](https://doi.org/10.1039/C1EE02455H).
- [226] K Tsuchiya and T Kimura. "Kinetics of substrate utilization in adenine-limited chemostat cultures of *Bacillus subtilis* KYA741." In: *Applied and Environmental Microbiology* 43(4):794-799 (1982).
- [227] A Turing. "The chemical basis of morphogenesis". In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 237.641 (1952), pp. 37-72. DOI: [10.1098/rstb.1952.0012](https://doi.org/10.1098/rstb.1952.0012).
- [228] A Typas, M Banzhaf, CA Gross, and W Vollmer. "From the regulation of peptidoglycan synthesis to bacterial growth and morphology." In: *Nature Reviews Microbiology* 10(2):123-136 (2011). DOI: [10.1038/nrmicro2677](https://doi.org/10.1038/nrmicro2677).
- [229] LK Ursell, JL Metcalf, LW Parfrey, and R Knight. "Defining the Human Microbiome." In: *Nutrition reviews* (2012), 70(Suppl 1):S38-S44. DOI: [10.1111/j.1753-4887.2012.00493.x](https://doi.org/10.1111/j.1753-4887.2012.00493.x).
- [230] Y Vervoort, A Gutiérrez-Linares, M Roncoroni, C Liu, J Steensels, and KJ Verstrepen. "High-throughput system-wide engineering and screening for microbial biotechnology." In: *Current Opinion in Biotechnology* 46:120-125 (2017). DOI: <https://doi.org/10.1016/j.copbio.2017.02.011>.
- [231] D Volfson, S Cookson, J Hasty, and LS Tsimring. "Biomechanical ordering of dense cell populations." In: *PNAS* 105(40):15346-51 (2008). DOI: [10.1073/pnas.0706805105](https://doi.org/10.1073/pnas.0706805105).

- [232] DS Wade, MW Calfee, ER Rocha, EA Ling, E Engstrom, JP Coleman, and EC Pesci. "Regulation of *Pseudomonas* Quinolone Signal Synthesis in *Pseudomonas aeruginosa*". In: *Journal of Bacteriology* 187.13 (2005), pp. 4372–4380. DOI: 10.1128/JB.187.13.4372-4380.2005.
- [233] P Wang, L Robert, J Pelletier, WL Dang, F Taddei, A Wright, and S Jun. "Robust Growth of *Escherichia coli*." In: *Cell* 20(12):1099-1103 (2010). DOI: 10.1016/j.cub.2010.04.045.
- [234] Y Wang, YY Yau, D Perkins-Balding, and JG Thomson. "Recombinase technology: applications and possibilities." In: *Plant Cell Rep* 30(3):267-85 (2011).
- [235] K Ward and Z Hugh Fan. "Mixing in microfluidic devices and enhancement methods." In: *Journal of Micromechanics and Microengineering* 25(9) (2015).
- [236] JD Watson and FHC Crick. "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid." In: *Nature* 171 (1953), pp. 737–38. DOI: 10.1038/171737a0.
- [237] JD Watson, M Gilman, J Witkowski, and M Zoller. "Recombinant DNA." In: *Scientific American Books* (1992). DOI: 10.1002/food.19940380229.
- [238] N Watson. "A new revision of the sequence of plasmid pBR322". In: *Gene* 70.2 (1988), pp. 399–403. DOI: [https://doi.org/10.1016/0378-1119\(88\)90212-0](https://doi.org/10.1016/0378-1119(88)90212-0).
- [239] E Weber, C Engler, R Gruetzner, S Werner, and S Marillonnet. "A Modular Cloning System for Standardized Assembly of Multigene Constructs". In: *PLoS ONE* 6(2):e16765 (2011). DOI: 10.1371/journal.pone.0016765.
- [240] M Weber and J Buceta. "Dynamics of the quorum sensing switch: stochastic and non-stationary effects". In: *BMC Systems Biology* 7.1 (2013), p. 6. DOI: 10.1186/1752-0509-7-6.
- [241] HU Weier and JW Gray. "A programmable system to perform the polymerase chain reaction." In: *DNA* 7(6):441-7 (1988).
- [242] R Weiss, G Homsy, and TF Knight Jr. "Toward *in-vivo* digital circuits." In: *Di-macs Workshop on Evolution as Computation* (1999).
- [243] HG Weller, G Tabor, H Jasak, and C Fureby. "A tensorial approach to computational continuum mechanics using object-oriented techniques". In: *Computers in Physics* 12(6) (1998).
- [244] *Which UUID version to use?* <https://stackoverflow.com/questions/20342058/which-uuid-version-to-use>. Accessed: 2019-01-11.

- [245] WB Whitman, DC Coleman, and WJ Wiebe. "Prokaryotes: the unseen majority." In: *PNAS* 95(12): 6578–83 (1998). DOI: 10.1073/pnas.95.12.6578.
- [246] JL Wiggs, JW Bush, and MJ Chamberlin. "Utilization of promoter and terminator sites on bacteriophage T7 DNA by RNA polymerases from a variety of bacterial orders." In: *Cell* 16(1):97-109 (1979).
- [247] G Wilson. "Software Carpentry: Getting Scientists to Write Better Code by Making Them More Productive". In: *Computing in Science and Engineering* (2006).
- [248] MK Winson, S Swift, L Fish, JP Throup, F Jorgensen, SR Chhabra, BW Bycroft, P Williams, and GSAB Stewart. "Construction and analysis of luxCDABE-based plasmid sensors for investigating N-acyl homoserine lactone-mediated quorum sensing". In: *FEMS Microbiology Letters* 163.2 (1998), pp. 185–192. DOI: 10.1111/j.1574-6968.1998.tb13044.x.
- [249] I Wong, S Atsumi, WC Huang, TY Wu, T Hanai, ML Lam, P Tang, J Yang, JC Liao, and CM Ho. "An agar gel membrane-PDMS hybrid microfluidic device for long term single cell dynamic study." In: *Lab on a chip* 10(20):2710-9 (2010). DOI: 10.1039/c004719h.
- [250] M Yaffe. "Reproducibility in science." In: *Science Signaling* 8:371 (2015).
- [251] P Yakovchuk, E Protozanova, and MD Frank-Kamenetskii. "Base-stacking and base-pairing contributions into thermal stability of the DNA double helix." In: *Nucleic Acids Research* 34(2):564-574 (2006). DOI: 10.1093/nar/gkj454.
- [252] X Yan, HJ Yu, Q Hong, and SP Li. "Cre/lox System and PCR-Based Genome Engineering in *Bacillus subtilis*". In: *Applied and Environmental Microbiology* 74.17 (2008), pp. 5556–5562. DOI: 10.1128/AEM.01156-08.
- [253] C Yanisch-Perron, J Vieira, and J Messing. "Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors." In: *Gene* 33(1):103-119 (1985). DOI: [https://doi.org/10.1016/0378-1119\(85\)90120-9](https://doi.org/10.1016/0378-1119(85)90120-9).
- [254] EA Yates, B Philipp, C Buckley, S Atkinson, SR Chhabra, RE Sockett, M Goldner, Y Dessaux, M Camara, H Smith, and P Williams. "N-Acylhomoserine Lactones Undergo Lactonolysis in a pH-, Temperature-, and Acyl Chain Length-Dependent Manner during Growth of *Yersinia pseudotuberculosis* and *Pseudomonas aeruginosa*". In: *Infection and Immunity* 70.10 (2002), pp. 5635–5646. DOI: 10.1128/IAI.70.10.5635-5646.2002.

- [255] TB Yehezkel, A Rival, O Raz, R Cohen, Z Marx, M Camara, JF Dubern, B Koch, S Heeb, N Krasnogor, C Delattre, and E Shapiro. "Synthesis and cell-free cloning of DNA libraries using programmable microfluidics". In: *Nucleic Acids Research* 44.4 (2016), e35. DOI: 10.1093/nar/gkv1087.
- [256] Y Yuzenkova, VR Tadigotla, K Severinov, and N Zenkin. "A new basal promoter element recognized by RNA polymerase core enzyme." In: *EMBO journal* 30(18):3766-3775 (2011).
- [257] N Zenkin and K Severinov. "The role of RNA polymerase subunit in promoter-independent initiation of transcription." In: *PNAS* 101(13):4396–4400 (2004). DOI: 10.1073/pnas.0400886101.
- [258] XZ Zhang, X Yan, ZL Cui, Q Hong, and SP Li. "mazF, a novel counter-selectable marker for unmarked chromosomal manipulation in *Bacillus subtilis*". In: *Nucleic Acids Research* 34.9 (2006), e71. DOI: 10.1093/nar/gkl358.
- [259] Y Zhang, F Buchholz, JPP Muyrers, and AF Stewart. "A new logic for DNA engineering using recombination in *Escherichia coli*". In: *Nature Genetics* 20 (1998), pp. 123–.
- [260] L Zhiwei, D Bin, J Zhihua, W Bingbing, X Xin, Y Dongyou, and L Weifen. "A versatile mini-mazF-cassette for marker-free targeted genetic modification in *Bacillus subtilis*". In: *Journal of Microbiological Methods* 95.2 (2013), pp. 207–214. DOI: <https://doi.org/10.1016/j.mimet.2013.07.020>.
- [261] WBJ Zimmerman. *Multiphysics modeling with finite element methods*. Vol. 18. World Scientific Publishing Co Inc, 2006.
- [262] N Ziv, NJ Brandt, and D Gresham. "The Use of Chemostats in Microbial Systems Biology." In: *Journal of Visualized Experiments* 80:50168 (2013). DOI: 10.3791/50168.
- [263] M Zuker. "Mfold web server for nucleic acid folding and hybridization prediction." In: *Nucleic Acids Research* 31(13):3406-15 (2003).
- [264] M Zwietering, I Jongenburger, F Rombouts, and K Van'T Riet. "Modeling of the Bacterial Growth Curve." In: *Applied and Environmental Microbiology* 56(6):1875-1881 (1990).