Design and Functional Assembly of Synthetic Biological Parts and Devices

Baojun Wang

A thesis submitted to the Imperial College London for the degree of Doctor of Philosophy

Department of Bioengineering
Imperial College London
Nov 2010
Declaration of Originality

I certify that the work presented in this thesis is my own and original. The work or any part of the work has not been previously submitted anywhere for any degree or diploma. I confirm that any ideas or work of others have been clearly referenced in the thesis and all main sources of help have been acknowledged.

Baojun Wang
Abstract

Programming living cells with synthetic gene circuits to perform desired tasks has been a major theme in the emerging field of synthetic biology. However, gene circuit engineering currently lacks the same predictability and reliability as seen in other mature engineering disciplines. This thesis focuses on the design and engineering of novel modular and orthogonal biological devices, and the predictable functional assembly of modular biological elements (BioParts) into customisable larger biological devices.

The thesis introduces the design methodology for engineering modular and orthogonal biological devices. A set of modular biological devices with digital logic functions, including the AND, NOT and combinatorial NAND gates, were constructed and quantitatively characterised. In particular, a novel genetic AND gate was engineered in Escherichia coli by redesigning the natural HrpR/HrpS heteroregulation motif in the hrp system of Pseudomonas syringae. The AND gate is orthogonal to E. coli chassis, and employs the alternative $\sigma^{54}$-dependent gene transcription to achieve tight transcriptional control. Results obtained show that context has a large impact on part and device behaviour, established through the systematic characterisation of a series of biological parts and devices in various biophysical and genetic contexts. A new, effective strategy is presented for the assembly of BioParts into functional customised systems using engineered ‘in-context’ characterised modules aided by modelling, which can significantly increase the predictability of circuit construction by characterising the component parts and modules in the same biophysical and genetic contexts as anticipated in their final systems. Finally, the thesis presents the design and construction of an application-oriented integrated system – the cell density-dependent microbe-based biosensor. The in vivo biosensor was programmed to be able to integrate its own cell density signal through an engineered cell-cell communication module and a second environmental signal through an environment-responsive promoter in the logic AND manner, with GFP as the output readout.
Acknowledgements

I am indebted to my advisor, Professor Richard Kitney, for the opportunity to work on this exciting project, his continued support and guidance whenever I met difficulty and needed help. I appreciate his vision and courage in this field. I would like to express my deep gratitude to my co-advisor and collaborator, Professor Martin Buck, for his excellent supervision and great enthusiasm in my work; for allowing me to be a member of his dynamic lab to complete my experiments and inspiring me by the long discussion we usually had.

Special thanks to Dr. Nicolas Joly for his guidance and generous help all the time, and the numerous fruitful discussion between us beside or away of the lab bench. I am honoured to have such a great friend who brings me lots of humour and joy.

Great thanks to Ellen James for teaching me the basic of molecular biology and helping me have transformed from an engineer to a biological engineer; Kirsten Jensen for the help and advice on the design and implementation of my wet lab experiment, and for the many orders placed for me. I am grateful to Dr. Vincent Rouilly, Dr. Matthieu Bultelle and Dr. Chueh-Loo Poh, who are the past members of Kitney’s group, for their initial looking after as well as introducing me into the attractive field of synthetic biology. I want to thank all other members in the two labs of Prof. Buck and Prof. Kitney, particularly Dr. Jörg Schumacher, Dr. Milija Jovanovic, James Chappell, Dr. Christoph Engl and Dr. Goran Jovanovic, for their kind help and advice in and out of the lab. Thanks to Dr. Robert Weinzierl for the generous help on the use of his productive liquid handling machines; Dr. Tom Ellis for the help on the use of flow cytometer; Dr. Peter Cashman and Professor Robert Krams for having examined my PhD transfer and provided valuable suggestions.

Special thanks to my parents, brother and love Lei, who have supported and encouraged me throughout my studies here.

Lastly, I would like to thank the sponsors of my PhD scholarship (UK/China Scholarships for Excellence) – the UK Department for Education and Skills, Imperial College London and China Scholarship Council.
Contents

Declaration of Originality ....................................................................................................... 2
Abstract .................................................................................................................................... 3
Acknowledgements .................................................................................................................. 4
Contents.................................................................................................................................... 5
List of Figures .......................................................................................................................... 9
List of Tables.......................................................................................................................... 12

Chapter 1 Introduction ......................................................................................................... 13
  1.1 Thesis Statement and Summary of Contributions ......................................................... 14
  1.1.1 Thesis statement..................................................................................................... 14
  1.1.2 Summary of contributions ..................................................................................... 14
  1.1.3 Thesis outline......................................................................................................... 16
  1.2 Synthetic Biology and the State-of-the-Art................................................................. 17
  1.2.1 Synthetic biology – an emerging discipline ......................................................... 17
  1.2.2 The state-of-the-art of synthetic biology............................................................... 20
  1.3 Current Design Principles and Challenges of Engineering Gene Circuits ............... 29
    1.3.1 Current design principles for engineering synthetic gene circuits ................. 29
    1.3.2 Challenges for engineering gene circuits ......................................................... 31
  1.4 Objectives of This Study ........................................................................................... 35

Chapter 2 Materials and Methods ....................................................................................... 36
  2.1 Materials .................................................................................................................... 37
  2.1.1 Media ..................................................................................................................... 37
  2.1.2 Antibiotics .............................................................................................................. 37
  2.1.3 Kits .......................................................................................................................... 38
  2.1.4 Dyes, chemicals and enzymes ............................................................................... 38
  2.1.5 Bacterial strains....................................................................................................... 39
  2.2 DNA Methods ............................................................................................................ 39
    2.2.1 DNA purification ................................................................................................. 40
    2.2.2 Polymerase Chain Reaction (PCR) ........................................................................ 40
    2.2.3 Agarose gel electrophoresis ................................................................................. 41
2.2.4 Gel extraction and purification of DNA ................................................................. 42
2.2.5 Restriction digest .................................................................................................. 42
2.2.6 Ligation ................................................................................................................ 42
2.2.7 Assembly strategy for DNA constructs ................................................................. 43
2.2.8 DNA sequencing and synthesis ............................................................................. 45
2.2.9 Competent cell preparation ................................................................................... 46
2.2.10 Heat shock transformation .................................................................................. 46
2.3 In Vivo Assay Methods .......................................................................................... 46
2.3.1 Growth of bacteria cells ......................................................................................... 46
2.3.2 Fluorescence assay for gene expression in living cells .......................................... 47
2.4 Modelling and Data Analysis Methods .................................................................... 51
2.4.1 Deterministic approach using ODE-based rate equations .................................... 51
2.4.2 Basic biochemical kinetics .................................................................................... 52
2.4.3 Modelling a constitutive single gene expression .................................................. 53

Chapter 3 The Design and Engineering of Modular Biological Devices ...................... 56
3.1 Approaches to Design Modular and Orthogonal Biological Devices ....................... 57
3.2 The Two Paradigms of Bacterial Gene Transcription and the $hpr$ Regulatory System in $Pseudomonas syringae$ ................................................................. 61
3.2.1 The two mechanistic paradigms of bacterial gene transcription............................ 61
3.2.2 The $hpr$ gene regulation system in $Pseudomonas syringae$ ............................... 63
3.3 Designing a Set of Modular Logic Devices and Functional Assembly of BioParts .... 68
3.3.1 The design of a modular logic AND gate ............................................................. 68
3.3.2 The design of a set of modular NOT gates and the composite NAND gate .......... 70
3.3.3 Functional assembly of BioParts using engineered in-context characterised modules .................................................................................................................. 71
3.4 Deriving Transfer Functions of the Designed Logic Devices ................................... 74
3.4.1 Deriving the transfer function of environment-responsive promoters .................. 74
3.4.2 Deriving the transfer function of the AND gate .................................................... 75
3.4.3 Deriving the transfer function of the NOT gate ..................................................... 76
3.4.4 Deriving the transfer function of the composite NAND gate ............................... 77

Chapter 4 Part Characterisation and the Engineering of a Modular Genetic AND Gate ................................................................. 79
4.1 The Initial Implementation of the AND Gate ............................................................ 80
List of Figures

Figure 1.1 An analogy between synthetic biology and computer engineering .................. 18
Figure 1.2 The architectures of first constructed synthetic gene circuits ..................... 20
Figure 1.3 Spatial pattern formation using a band-detect circuit based on intercellular communication ................................................................. 22
Figure 1.4 Genetic logic using microRNAs to control the translation of target mRNA transcripts ........................................................................................................... 24
Figure 1.5 The assembly of a synthetic M. mycoides genome .................................. 26
Figure 1.6 Environmentally controlled invasion of cancer cells by engineered bacteria ... 27
Figure 2.1 BioBrick Standard Assembly method .................................................. 44
Figure 2.2 Three way ligation method .................................................................... 45
Figure 2.3 A single gene expression modelling .................................................. 53
Figure 2.4 Modelling of a single gene expression dynamics ................................ 54
Figure 3.1 Simplified architecture of a prokaryotic transcriptional regulatory module ... 58
Figure 3.2 Modular design supports standard inputs and outputs of the designed biological modules ................................................................. 59
Figure 3.3 The two paradigms of gene transcription in bacteria ............................... 62
Figure 3.4 The σ^54-dependent transcription activation ......................................... 63
Figure 3.5 hrp gene regulation in the type III secretion system of P. syringae .......... 65
Figure 3.6 The extracted hrp gene regulation model in P. syringae ........................... 66
Figure 3.7 The rational design of the modular logic AND gate ............................... 69
Figure 3.8 The two- and three-input modular AND gates with one input inverted .... 70
Figure 3.9 The design of a set of modular NOT gates and the composite NAND gate ... 71
Figure 3.10 The physical and functional assemblies of synthetic biological modules .... 72
Figure 3.11 The architecture of an environment-responsive negatively regulated promoter (P1). ........................................................................................................... 74
Figure 3.12 The architecture of the genetic logic AND gate circuit ......................... 75
Figure 3.13 The architecture of the genetic logic NOT gate circuit ......................... 76
Figure 3.14 The architecture of the genetic logic NAND gate circuit ....................... 77
Figure 4.1 Schematic diagram for the construction of the genetic AND gate ............. 81
Figure 4.2 Plasmid maps showing the circuit constructs used for the initial version of the AND gate ................................................................................................. 81
Figure 4.3 The systematic characterisation of the AND gate................................................................. 82
Figure 4.4 Schematic diagram for the characterisation of the three inducible promoters: P_{lac}, P_{BAD} and P_{lux}................................................................................................................. 84
Figure 4.5 Plasmid maps showing some of the circuit constructs used for the characterisation of the three promoter inputs................................................................. 85
Figure 4.6 Dynamics of the P_{lac} promoter in response to various levels of induction.............. 86
Figure 4.7 The dose response of the P_{lac} (a), P_{BAD} (b) and P_{lux} (c) promoters to various induction levels in the two cell chassis in M9 media................................................................. 87
Figure 4.8 The characterisation results of P_{lac} (a), P_{BAD} (b) and P_{lux} (c) using 6 versions of RBS in response to various inductions ................................................................. 88
Figure 4.9 The core promoter regions and 5' UTR (five prime untranslated region) sequences of the three characterised promoters................................................................. 89
Figure 4.10 Normalised dose response of the three characterised promoters with 6 RBSs of various strengths ................................................................................................................. 90
Figure 4.11 The characterised responses of P_{lac} (a), P_{BAD} (b) and P_{lux} (c) under two temperatures................................................................................................................................. 92
Figure 4.12 Plasmid maps showing the circuit constructs used for the characterisation of the AND gate ................................................................................................................................. 94
Figure 4.13 The systematic characterisation of the modular AND gate........................................... 95
Figure 4.14 Fitted transfer function of the AND gate........................................................................ 96
Figure 4.15 The engineering and systematic characterisation of the modular AND gate with alternative inputs................................................................................................................... 98
Figure 4.16 Predictions of the AND gate behaviour using fitted transfer functions. ...................... 99
Figure 4.17 FACS assays of the promoter P_{lac} (a), P_{BAD} (b) and P_{lux} (c) in E. coli MC1061 after 5 hours growth in M9-glycerol at 37 °C................................................................. 101
Figure 4.18 FACS assays of the engineered AND gate using P_{lac} and P_{BAD} as the two inputs in E. coli MC1061 after 5 hours growth in M9-glycerol at 30 °C............................................. 102
Figure 4.19 FACS assays of the engineered AND gate using P_{lux} and P_{BAD} as the two inputs in E. coli MC1061 after 7 hours growth in M9-glycerol at 37 °C............................................. 103
Figure 4.20 Growth curves of E. coli MC1061 harbouring various circuit constructs in M9-glycerol at 30 °C. ......................................................................................................................... 104
Figure 4.21 Qualitative assays of the functionality of the engineered AND gate using P_{lac} and P_{BAD} as the two inputs in seven E. coli strains............................................................................. 106
Figure 4.22 Qualitative assays of the functionality of the engineered AND gate using P_{lux} and P_{BAD} as the two inputs in seven E. coli strains............................................................................. 107
Figure 5.1 Design of the modular genetic NOT gates. .......................................................... 113
Figure 5.2 Plasmid maps showing circuit constructs used for the characterisation of the NOT gates. .......................................................................................................................... 114
Figure 5.3 The characterisation of the lacI/P_{lac} based NOT gate. ........................................ 115
Figure 5.4 The characterisation of the tetR/P_{tet} based NOT gate. ....................................... 116
Figure 5.5 The characterisation of the cI/P_{lam} based NOT gate. ......................................... 117
Figure 5.6 Two-input NAND gate using P_{lac} and P_{BAD} inputs, and the cI/P_{lam} based NOT gate. ............................................................................................................................ 121
Figure 5.7 Two-input NAND gate using P_{lux} and P_{BAD} inputs, and the cI/P_{lam} based NOT gate. ............................................................................................................................ 122
Figure 5.8 The systematic characterisation of the first composite NAND gate. ................... 124
Figure 5.9 The systematic characterisation of the second composite NAND gate. ............... 125
Figure 5.10 FACS assays of the engineered composite NAND gate using P_{lac} and P_{BAD} as the two inputs in *E. coli* MC1061 after 5 hours (a-c) or 7 hours (d) growth in M9-glycerol at 30 °C. ................................................................................................ 127
Figure 6.1 The quorum sensing circuit in *Vibrio fischeri* and the regulatory region of *lux* operon. ................................................................. 133
Figure 6.2 The engineering of a cell density-dependent module.......................................... 136
Figure 6.3 The engineering of a cell density-dependent microbe-based biosensor.............. 139
Figure 6.4 Characterisation of the cell density-dependent biosensor.................................. 141
Figure 6.5 FACS side-scattering and forward-scattering graphs of the cells (*E. coli* MC1061) harbouring the engineered biosensor. ................................................................. 143
Figure 6.6 Potential improved architectures for the cell density-dependent module. ........... 146
Figure 7.1 The design of the distributed bacterial optimiser................................................ 154
List of Tables

Table 2.1 Bacterial growth media used in this study. .............................................................. 37
Table 2.2 Antibiotics used in this study. ................................................................................ 37
Table 2.3 Kits used in this study. ........................................................................................... 38
Table 2.4 Dyes, chemicals and enzymes used in this study. ................................................. 38
Table 2.5 Bacterial strains used in this study. ....................................................................... 39
Table 2.6 Reaction mix for PCR amplification used in this study. ........................................ 40
Table 2.7 Operating parameters for PCR reactions in this study. ........................................ 41
Table 4.1 The RBS sequences used for the promoter characterisation. ............................... 84
Table 4.2 The model fits for promoter characterisation using various RBSs in the standard condition with 95% confidence bounds otherwise fixed at bound. ........................ 91
Table 4.3 Summary of chassis compatibility assays of the AND gate in Figure 4.21. .......... 106
Table 4.4 Summary of chassis compatibility assays of the AND gate in Figure 4.22. .......... 108
Table 5.1 The model fits of the three types of NOT gates using various RBSs in the standard condition with 95% confidence bounds otherwise fixed at bound. ................. 118
Table A.1 The plasmid DNAs used in this study. ................................................................. 155
Table B.1 The oligo DNAs (primers) used in this study. ..................................................... 159
Chapter 1

Introduction
Chapter 1 - Introduction

1.1 Thesis Statement and Summary of Contributions

1.1.1 Thesis statement
The thesis focuses on the design, construction and characterisation of novel modular and orthogonal biological devices with digital logic functions, and the assembly of modular biological elements (BioParts) into customisable larger biological systems. One central goal is to be able to rationally engineer organisms with desired properties in a more predictable manner using these synthetic biological circuits. The thesis makes the following claim:
The functional assembly of synthetic biological parts and devices into customised biological systems can be performed more predictably using engineered ‘in-context’ (i.e. in the same biophysical and genetic contexts as anticipated in their final system) characterised modules.

1.1.2 Summary of contributions
In this thesis, I make the following main contributions:

(1) Design and engineering of a set of novel modular and orthogonal devices:
I present the design principles for constructing novel modular (interchangeable) and orthogonal (i.e. no cross talk) biological devices with a particular focus on the transformation of natural biological discoveries into useful functional modules. I engineer a set of such digital logic devices including the AND gates, the NOT gates and the composite NAND gates in Escherichia coli with novel gene regulation components from Pseudomonas syringae. These fundamental synthetic devices represent new contributions made to expand the currently limited toolkit of synthetic biology for constructing larger scale systems with complex circuitry.

(2) Determining the effect of context on part and device behaviour:
I systematically characterise a set of component biological parts (i.e. promoters and ribosome binding sites etc) and devices in various contexts (both biophysical and genetic) and determine the effect of context (both biophysical and genetic) on the part
and device behaviour. I show that the context has a large impact on the behaviour of the involved parts and modules, suggesting that we have to take the context effect into account to increase the likelihood of producing functional circuits from the assembly of individual components. The current scope for characterising synthetic parts and devices are also extended by including homogeneity, metabolic load and chassis compatibility characterisation.

(3) **An effective strategy for the functional assembly of BioParts:**

I develop an effective strategy for the assembly of BioParts into functional customised systems using engineered ‘in-context’ characterised modules. By characterising the candidate component parts and modules in the same biophysical or abiotic environment (e.g. media, temperature, carbon source) and genetic (e.g. embedded sequence context, chassis background) contexts as anticipated for their final target system, I have successfully assembled a set of functional logic circuits as mentioned before in a more predictable manner aided by modelling. This approach minimises the unexpected or high-order effects which could occur during circuit construction by eliminating or reducing the variations arising from the difference of contexts, and thus greatly increase the predictability for the functional assembly of BioParts into synthetic circuits. This semi-empirical approach has merit in that an operational system is arrived at.

(4) **An in vivo biosensor for practical applications:**

I present the design and construction of a cell density-dependent module from standardised biological parts based on the lux quorum sensing mechanism. I then engineer a modular cell density-dependent *E. coli*-based biosensor assembled from multiple characterised parts and modules. The *in vivo* biosensor is programmed to be able to integrate its own cell density signal through a cell-cell communication module and a second environmental signal through an environment-responsive promoter in the logic AND manner, with GFP level as the output readout. This modular sensing device can be customised to have application in many areas such as biotechnology including fermentation. Thus the work demonstrates the great potential of engineering synthetic circuits to program organisms for real world applications.
1.1.3 Thesis outline

In the remainder of the thesis, Chapter 1 introduces the advances in synthetic biology, and the current design principles and existing challenges for engineering gene circuits, followed by a description of the main objectives of this study. Chapter 2 summarises the materials and methods that are used for both the experiment and modelling in this study. Chapter 3 describes the principles used to design and engineer modular and orthogonal biological devices as well as the designs and functional assembly strategy for the synthetic modules constructed in the following two chapters. Derivations of the transfer functions of designed modules are then described. Chapter 4 presents the systematic characterisation of component parts, i.e. promoters and RBSs, in various live cell biophysical and genetic contexts to illustrate the impact of context on their behaviour. The chapter next presents the forward engineering and thorough characterisation of the designed modular logic AND gate. Chapter 5 presents the engineering of the three types of modular logic NOT gates and the composite logic NAND gates. Chapter 6 presents the design and construction of a cell density-dependent module based on the lux quorum sensing system. The chapter next describes the engineering of a modular cell-density dependent in vivo biosensor using the modules that have been constructed with arabinose sensing as an example. Chapter 7 summarises the study and discusses possible future directions.
1.2 Synthetic Biology and the State-of-the-Art

This section introduces the emerging discipline of synthetic biology, which forms the basis of the study in this thesis, and reviews the latest advances in this field. Particularly, the studies on the engineering of synthetic gene circuits in living cells – the major theme in synthetic biology – are presented and discussed.

1.2.1 Synthetic biology – an emerging discipline

In the past 60 years we have seen enormous advances in biology, in particular molecular biology. Not only has there been a rapid accumulation of various genetic data and biological information (encompassing genes, proteins, and their interactions in many organisms) particularly following the sequencing of many microbial genomes as well as the human genome, but, also, the great advancement in the technologies that allows us to manipulate genetic materials. One such technology is genetic engineering, which has enabled us to use specific pieces of DNA to modify or improve existing cellular functions. This is achieved by modifying the endogenous gene regulation networks inside a cell using technologies such as DNA recombination, DNA sequencing and synthesis, DNA amplification and site-directed mutagenesis. Following these advances, it is now possible to design novel genetic circuits from scratch and embed them into living cells to enable them with desired properties in a nearly predictable manner. This has led to the birth of a new field called synthetic biology, which has been established as an engineering discipline and targets designing and constructing synthetic biological systems from genetic components with desired functionalities and characteristics.

Synthetic biology is defined as the design and construction of new biological parts, devices and systems, and the redesign of existing, natural biological systems for useful purposes (Synthetic Biology Organisation, 2008). The goal is to extend or modify the behaviour of organisms and engineer them to perform new customised tasks, such as producing drugs, chemicals and biofuels.

Figure 1.1 (Andrianantoandro et al., 2006) illustrates the goal and methods of synthetic biology by making analogy to computer engineering at the different levels
Figure 1.1 An analogy between synthetic biology and computer engineering. The diagram is from (Andrianantoandro et al., 2006).

of a hierarchy. In this hierarchy, they both take a bottom-up approach to build more
complex systems by integrating constituent parts. At the bottom are genes, proteins
and chemical molecules analogous to the physical layer of transistors, resistors and
capacitors in computer engineering. At the next layer, biochemical reactions regulate
the information flow, equivalent to logic gates in a computer. At the module layer,
complex pathways can be assembled from a library of biological devices with varying
functions and function like integrated circuits. By integrating these modules into cells,
the cells will be programmed with desired behaviour. More complex tasks can be
accomplished by using a cell population, in which cells communicate each other to
perform in a coordinated way, much like the case of computer networks.

A deep understanding of the role and significance of synthetic biology can be
obtained by comparing it to the well established disciplines of systems biology,
synthetic organic chemistry and electrical engineering.
Synthetic biology builds upon biology. It can be treated as a bottom up approach to study biological systems by assembling engineered modular components to form larger or more complex systems that mimic the natural ones. But the approach used in synthetic biology places it apart from other, existing biological studies; i.e. by the extensive use of engineering principles and tools. There are some similarities to systems biology, which also involves the application of engineering theory to biology studies. However, while systems biology focuses on the comprehensive study of natural biological systems, often within a biomedical context, synthetic biology seeks to build novel and artificial biological systems for specific purposes. It is described as the engineering application of the biomedical science, rather than the extension of bioscience research (The Academy of Medical Sciences and the Royal Academy of Engineering, 2007). It is better to see that they are two complementary ways, as the system design principles found by systems biology will guide the forward engineering of novel biological systems for synthetic biology, which in return help prove and support former discoveries.

By making the comparison to organic chemistry, synthetic biology is to biology what synthetic organic chemistry is to chemistry. It is the combination of the synthetic approach with the traditional analytical approach that revolutionized chemistry, leading to a deep understanding of the fundamental principles of chemical structure and reactivity (Yeh and Lim, 2007). Similarly, the synthetic approach taken by synthetic biology is a necessary complement to the analytical approach currently taken by most biologists to elucidate the design principles of biological systems.

Finally, from an engineering point of view, what synthetic biologists are doing now is quite like what electrical engineers have been doing for many years, designing electronic circuits using standard components, such as resistors, capacitors and transistors. The difference lies in the building blocks that are used. Synthetic biologists design genetic circuits with specified functions using standard engineered biological parts such as genes, promoters, ribosome binding sites and terminators. In this regard, synthetic biology is to biology what electrical engineering is to physics, which both deal with electrons but one focuses on the understanding of their nature and the other aims to make use of them to build useful applications.
1.2.2 The state-of-the-art of synthetic biology

The beginning of synthetic biology era is hallmarked by the three independent studies published in Nature journal in 2000 (Becskei and Serrano, 2000; Elowitz and Leibler, 2000; Gardner et al., 2000), which introduced the first synthetic gene circuits. As shown in Figure 1.2, they are all transcriptional regulation circuits constructed in \textit{E. coli}. The negative autoregulation circuit (Figure 1.2a) was shown to increase the stability of gene expression and reduce the noise by an artificially constructed negative feedback. The toggle switch circuit (Figure 1.2b) was shown to exhibit bistability through two mutually repressed gene regulation modules. The synthetic oscillatory circuit (Figure 1.2c) was shown to display damped oscillatory behaviour in terms of fluorescent output in individual cells by linking three repressing modules in a ring topology. After these successes, there has been continuous effort in designing switches (Ajo-Franklin et al., 2007; Atkinson et al., 2003; Becskei, 2001) and oscillators (Atkinson et al., 2003; Danino et al., 2010; Kitney et al., 2007; Stricker et al., 2008; Tigges et al., 2009) by incorporating various negative and positive feedbacks in the engineered synthetic circuits either in prokaryotic or eukaryotic cells.

**Figure 1.2 The architectures of first constructed synthetic gene circuits.** a, The negative autoregulation circuit (Becskei and Serrano, 2000). b, The genetic toggle switch (Gardner et al., 2000). c, The synthetic oscillator (Elowitz and Leibler, 2000).

Hence, it has been no more than ten years since the first pieces of proof-of-principle work came out in synthetic biology. However, a rapid growth has taken
place in this field in the past decade and many examples have been produced to demonstrate the principles of gene circuit engineering and the promising applications it allows to have. The following gives a brief review of the major advances in this field by grouping them into several aspects that are currently prevalent in the synthetic biology community.

**Intracellular vs. Intercellular or Interspecies circuits**

Although most of the circuits constructed so far are dependent on individual cellular behaviour alone, some pieces of work have applied cell-cell communication with synthetic intercellular circuits to achieve coordinated responses among a population of cells or different species. This is a major advance because it could not only achieve spatial-temporal control over a population of cells to fulfil complex tasks, but, also, help overcome cell to cell variations due to the stochastic nature of gene expression in individual cells (Elowitz et al., 2002) by synchronised cellular behaviour (Danino et al., 2010). Most of these intercellular systems used a quorum sensing circuit module derived from the symbiotic bacterium *Vibrio fischeri* (Bassler, 2002; Waters and Bassler, 2005). The module is composed of LuxI, an enzyme which catalyzes the synthesis of the freely diffusible signalling molecule – AHL, LuxR, a transcriptional activator dependent on AHL, and the LuxR-AHL complex responsive to LuxI promoter. In response to high concentrations of AHL (for example, due to a high cell density), genes located downstream of the LuxI promoter will be transcribed.

One classic example is the pattern formation system which has been constructed in *E. coli* by utilising such intercellular communication module to form specified spatial patterns over a bacterial lawn grown on an agar plate. As shown in Figure 1.3 (Basu et al., 2005), the system comprises the sender cells and the receiver cells. The sender cells placed in the plate centre contain LuxI and synthesise AHL molecules which diffuse into surrounding medium, while the plated receiver cells contain band-detect circuits and respond to the AHL signalling molecules secreted from the sender cells, but only express GFP at certain ranges of AHL level. Since the level of AHL forms a spatial gradient that decreases with the distance away from the sender cells, only the receiver cells that are located within the range of the band-detecting window
express fluorescent proteins and thus form a ring pattern. The bullseye spatial pattern formed by two strains of receiver cells with different band detecting windows and fluorescent reporters is displayed on the top right in Figure 1.3.

![Figure 1.3 Spatial pattern formation using a band-detect circuit based on intercellular communication.](image)

Other engineered intercellular circuits include the transient pulse generating circuit containing a feed forward network module (Basu et al., 2004), and the programmed population control circuit which autonomously controls the density of *E. coli* cell population (You et al., 2004). Kobayashi et al. (Kobayashi et al., 2004) have engineered cells that activated protein synthesis when only the cell population reaches a critical density. Tabor et al. (Tabor et al., 2009) have created a synthetic genetic edge detection circuit which enabled cells to have the capability of detecting the edges between light and dark regions based on intercellular communication and a synthetic light sensor. Synthetic ecosystems using two engineered populations of *E. coli* cells (Balagadde et al., 2008; Brenner et al., 2007; Chuang et al., 2010; Song et al., 2009) have also been constructed. Other synthetic ecosystems, allowing communication across different species or kingdoms, were shown to enable quorum sensing between
and among mammalian cells, bacteria, yeast, and even plants through multistep airborne communication (Weber et al., 2007). And the signals used for intercellular signalling in the study were expanded as well including volatile aldehydes, small vitamin-derived molecules, or antibiotics that diffuse either by gas or liquid phase.

Transcriptional vs. Post-transcriptional control

Gene expression control underlies the construction of synthetic gene circuits. So far transcription regulation is the dominating method to control gene expression in designing functional circuits, which normally uses trans-acting transcription factors or cis-acting DNA elements to regulate the mRNA transcription of the gene of interest. However, a series of studies carried out exhibits the potential of regulating gene expression at the post-transcriptional level for designing biological circuits, which mainly controls target protein synthesis from transcribed mRNAs (Isaacs et al., 2006).

One such example is the riboswitches (Winkler, 2005), which include aptamer domains in the mRNAs to control gene expression in a ligand-dependent manner. The specific binding of the ligand to the aptamer domain of the riboswitch induces a conformational change in the 5' UTR of its own mRNA, thereby regulating gene translation. An engineered theophylline-dependent riboswitch has been shown to control chemotactic gene expression in *E. coli* and thus guide the bacterial movement (Topp and Gallivan, 2007) toward the theophylline signal. The riboregulator designed by Bayer et al. (Bayer and Smolke, 2005) uses trans-acting antisense RNAs that bind target mRNA to regulate eukaryotic gene expression in a ligand-responsive manner. Ribozyme based synthetic RNA switches have also been shown to be able to tightly regulate gene expression in a programmable manner (Win and Smolke, 2007; Yen et al., 2004), such as the control of mammalian T-cell proliferation (Chen et al., 2010). More recently, a type of small regulatory RNA, named microRNAs, was exploited to control the transcription of the mRNA of interest via a RNA interference pathway in a programmable logic manner (Leisner et al., 2010). Figure 1.4 shows the logic AND gate constructed on the basis of this mechanism. The transcription factors regulate synthetic genes coding for the microRNAs, which in turn bind their corresponding
microRNA target sequences in the 3' UTR of the output gene transcript to inhibit its degradation caused by these microRNA target sequences.

Figure 1.4 Genetic logic using microRNAs to control the translation of target mRNA transcripts. The diagram is adapted from (Leisner et al., 2010).

Rackham et al. (Rackham and Chin, 2005b) designed orthogonal ribosome mRNA pairs by evolving natural ribosome to new versions with altered mRNA-binding sites (orthogonal ribosomes) which exclusively translate mRNAs that are not substrates for the endogenous ribosome (orthogonal mRNAs). These new pairs are orthogonal to their wide type, and can be used to programme cellular logic (Rackham and Chin, 2005a). A modular AND gate has been constructed by including amper stop codons into the T7 RNA polymerase gene under an inducible promoter (Anderson et al., 2007). Thereby the transcribed T7 mRNAs are not translated to polymerase until the nonsense amber codon suppressor tRNA \textit{supD} is also transcribed. There is also work on controlling gene expression both at the transcriptional and translational levels to achieve tight and tunable control. An engineered, tunable genetic switch that couples repressor proteins and an RNAi target design has been shown to effectively
turn any gene off in mouse and human cells (Deans et al., 2007). Post transaltional
regulation is demonstrated as well by engineering scaffold protein interactions to
systematically reshape signalling dynamics of the yeast mating MAP kinase pathway
(Bashor et al., 2008).

**Bottom up vs. Top down approaches**
Constructing artificial gene circuits with basic genetic elements from scratch can be
considered as the bottom up approach in synthetic biology. With the construction of
larger and more complex genetic systems, we could imagine a totally artificial
organism in the future, in which every part is synthetic, including the genome. Efforts
toward such a synthetic organism have been ongoing. Generally this work starts from
the top level or targets the top level, i.e. the whole genome and the organism, which is
considered as the top down approach at present.

Initial work has been carried out by Venter and his colleagues in identifying
essential genes of the minimal bacterium *Mycoplasma genitalium* (Glass et al., 2006;
Hutchison et al., 1999). Engineered *E. coli* with genome reduction up to 15% is
studied (Kolisnychenko et al., 2002; Posfai et al., 2006), which displayed some
emergent and unexpected beneficial properties. Forster et al. (Forster and Church,
2006) are using well-defined essential genes to synthesize a minimal cell which owns
replication and other essential biological pathways, and is fed only by small molecule
nutrients. The bacteriophage T7 genome has also been redesigned by replacing around
30% natural genetic elements with engineered DNA (Chan et al., 2005). Whole
genome transition between different bacterial species (Lartigue et al., 2007; Lartigue
et al., 2009) and the first assembled synthetic genome (Gibson et al., 2008) have been
accomplished by the pioneers in this field, which are the first two steps towards a final
whole synthetic organism. More recently, the researchers at the J. Craig Venter
Institute have successfully created a so-called synthetic bacterial cell, which is
controlled wholly by the chemically synthesised genome (Gibson et al., 2010). As
shown in Figure 1.5, they first ordered 1078 DNA fragments (around 1 kb each) of the
whole *M. mycoides* genome from a DNA synthesis company. The ordered cassettes
are 1 kb size around with overlaps to adjacent cassettes but there are designed
sequences like the watermarks within some of them. They next used a hierarchical approach to assemble the cassettes into the final chromosome in three steps in yeast by transformation and homologous recombination. The extracted synthetic genome was then transformed into recipient *M. capricolum* cells. The transformed colonies were screened to select the ones containing only the synthetic genome. The synthetic *M. mycoides* bacterium was confirmed by sequence analysis to contain only the designed synthetic DNA sequences and has expected phenotypic properties.

**Figure 1.5 The assembly of a synthetic *M. mycoides* genome.** The 1.08-Mbp synthetic *M. mycoides* genome was assembled from 1078 overlapping DNA cassettes in three steps by transformation and homologous recombination in yeast. The diagram is from (Gibson et al., 2010).

The results of these projects are all of great importance for synthetic biologists as they can not only offer insights on which components on the genome are essential, but also can help provide an engineered standard chassis with known features which synthetic biologists can use in the future. Other genome engineering examples include making automated multiple targeted changes to a genome to optimise a synthetic metabolic pathway (Wang et al., 2009b), and generating genome scale modified bacteria with modified expression of almost every gene in *E. coli* (Warner et al., 2010).
Fundamental vs. Application-oriented systems

Most of the systems that have been constructed in synthetic biology are fundamental, without a direct practical application and could be called “toy systems”. But one major aim of synthetic biology is to have diverse practical applications in many areas by engineering biology. Up to now there are already a few application-oriented synthetic systems that have been demonstrated although this field should be more driven by engineering applications.

**Figure 1.6 Environmentally controlled invasion of cancer cells by engineered bacteria.** Invasion is induced when bacteria are above a critical cell density or in a hypoxic environment. On the left is the engineered circuit working in a cell density-dependent manner. The diagrams are from (Anderson et al., 2006).

Kobayashi et al. (Kobayashi et al., 2004) have coupled the engineered genetic toggle switch to the cell’s natural SOS signalling pathway to form biofilms upon response to DNA-damaging agents. The engineered *E. coli* bacteria with a synthetic quorum sensing circuit are triggered to invade tumor cells under certain conditions as shown in Figure 1.6 (Anderson et al., 2006). A proof-of-principle bacterial photographing system is constructed in *E. coli*, which uses a synthetic protein made by fusing a light-sensing domain of one protein to a signal transduction domain of another. As a result, the engineered bacterium can express a certain enzyme to turn the media black when exposed to light (Levskaya et al., 2005).
Another major application area comes from the combination of synthetic biology with metabolic engineering. For instance, Keasling et al. have demonstrated how to produce the precursor chemical compounds for anti-malaria drug artemisinin in *E. coli* using an engineered metabolic pathway from the yeast strain (Martin et al., 2003) and later in engineered yeast (Ro et al., 2006). Bayer et al. (Bayer et al., 2009) used synthetic metagenomic approach to seek optimised versions of methyl halide transferase (MHT) enzyme from many natural sources for the synthesis of methyl halides from biomass in a symbiotic co-culture of the engineered yeast and the cellulolytic bacterium *Actinobacteria fermentans*. Other examples include using engineered cyanobacterium *Synechococcus elongatus* to produce isobutyraldehyde and isobutanol directly from carbon dioxide (Atsumi et al., 2009), and optimising the 1-deoxy-d-xylulose-5-phosphate (DXP) biosynthesis pathway in *E. coli* to improve lycopene production using a genome engineering approach (Wang et al., 2009b).

These application-oriented systems have shown the enormous potential of engineering synthetic biological systems for practical applications. The prospective application areas could span a wide range from microbial biosensors (Khalil and Collins, 2010), biofuels production (Lee et al., 2008; Savage et al., 2008), biopharmaceuticals and biomaterials synthesis (Bayer et al., 2009) to gene therapy, tissue engineering and bio-computing (Haynes et al., 2008).
1.3 Current Design Principles and Challenges of Engineering Gene Circuits

This section briefly reviews the design principles currently used in the synthetic biology community to engineer synthetic gene circuits with desired functions and discuss their advantages and disadvantages. The existing challenges in gene circuit design and construction are discussed subsequently.

1.3.1 Current design principles for engineering synthetic gene circuits

Lacking the predictability found in other mature engineering disciplines, constructing synthetic gene circuits with desired properties from discrete gene components is non-trivial and usually involves an iterative process of design, construction, testing and verification. Thus to increase the likelihood of producing functional gene circuits and to reduce the associated cost and effort, there are several design principles that is currently prevalent in this field to guide the process of gene circuit engineering. The following introduces these design principles and their successful applications.

Rational design

One such approach is iterative rational design (Purnick and Weiss, 2009), which involves building and analysing a computational model of a designed system, experimentally constructing the relevant circuit and verifying the system behaviour, fine-tuning the circuit guided by the modelling until the desired performance is achieved. The genetic toggle switch (Gardner et al., 2000), repressilator (Elowitz and Leibler, 2000), negative auto-regulation (Becskei and Serrano, 2000) and pattern formation circuits (Basu et al., 2005) are among the examples of the application of this design principle. This approach was further developed using well-characterised and modelled components to predict the target synthetic system behaviour (Guido et al., 2006; Rosenfeld et al., 2007; Tabor et al., 2009). However, in practice, this method normally involves an iterative and labourious process of trial-and-error to fine tune the target circuit because the model parameters of the components in the system are not always available, and the emergent effect (unexpected system behaviour) may occur even when the components are well characterised.
**Directed evolution**

Another method widely used in creating or fine tuning functional gene circuits is directed evolution (Haseltine and Arnold, 2007), which subjects a given component or circuit to random mutagenesis (by error-prone PCR or PCR with designer variant oligonucleotides), followed by a screen or selection to isolate mutants that meet the desired behavioural criteria (McDaniel and Weiss, 2005). Arnold et al. have applied directed evolution to evolve an original unmatched genetic circuit to the functional one by mutating the protein coding sequence of a gene and its RBS within the circuit (Yokobayashi et al., 2002), and to evolve the transcriptional activator LuxR leading to improved response to quorum sensing signals (Hawkins et al., 2007). Other examples include the engineered cancer cell invasion circuit (Anderson et al., 2006) and the genetic logic AND gate (Anderson et al., 2007), which were constructed by selecting the circuit variants generated with libraries of mutant RBS sequences for a certain gene in the systems. Thus this approach harnesses the advantages of genetic diversity and selection to engineer functional circuit without having detailed knowledge of the underlying components. But it may require many rounds of evolution and screening before the objective performance is achieved. It is also of low efficiency if the space for evolution is large, and the mutant library size increases combinatorially with the number of components to be evolved.

**Combinatorial synthesis**

The third approach which is used is known as combinatorial synthesis, which constructs circuit variants with parts in predefined different combinations and configurations and then selects out the functional ones with the desired performance (Michalodimitrakis and Isalan, 2009; Purnick and Weiss, 2009). This combinatorial approach was first demonstrated in a study of randomly shuffling the connectivity of the three transcriptional regulators (TetR, LacI and CI) and their respective promoters (Guet et al., 2002). The generated transcriptional networks displayed varying phenotypic behaviour resembling different logic functions of NAND, NOR and NOT IF. Thus by changing the underlying gene network architecture, combinatorial synthesis provides an alternative method for generating diverse phenotypes, as well as
for studying biological networks. In another study, researches combinatorially constructed around 600 recombinations of bacterial promoters with different transcription or sigma factors ORFs on plasmids in *E. coli* and thus added new links to the wild-type bacterial global transcription network (Isalan et al., 2008). They surprisingly found that the rewired bacteria can tolerate most of the added links without growth defects, and some even acquired survival benefits against certain selection pressures. Other examples include combinatorial promoter design by adding and rearranging multiple operator sites within a promoter (Cox et al., 2007; Murphy et al., 2007) to generate diverse promoter regulatory functions.

**Hybrid approaches**

Besides the design principles described, there are also hybrid approaches developed to improve the efficiency of gene circuit engineering by combining the advantages of the three approaches mentioned above. For example, the rational design-combinatorial approach (Dueber et al., 2009; Ellis et al., 2009; Ramalingam et al., 2009) improves the likelihood of successful construction using quantitative modelling to guide the choice of the component combinations in the circuit. The integrated rational design-directed evolution approach (Feng et al., 2004; Lou et al., 2010; Sayut et al., 2009; Zhan et al., 2010) uses the modelling results of rational design to help narrow down the number of targets in the system to be mutated and evolved, and thus increases the efficiency. Another one is the combinatorial-evolutionary approach (Atsumi and Little, 2006; Pfleger et al., 2006; Wang et al., 2009b), which normally subjects multiple sites in a gene network to be evolved in parallel to produce an optimised combination of the network components for the selected characteristics.

**1.3.2 Challenges for engineering gene circuits**

As introduced in the previous section, synthetic biologists have accomplished a great deal in a short time. A number of synthetic biological parts with basic functions (Voigt, 2006) have been developed and assembled to construct some proof-of-principle circuits, such as switches (Ajo-Franklin et al., 2007; Gardner et al., 2000;
Winkler, 2005), oscillators (Elowitz and Leibler, 2000; Fung et al., 2005; Stricker et al., 2008; Tigges et al., 2009), digital logic devices (Anderson et al., 2007; Bronson et al., 2008; Zhan et al., 2010), cell-cell communicators (Basu et al., 2005; Danino et al., 2010; You et al., 2004) and sensors (Anderson et al., 2006; Bayer and Smolke, 2005; Kobayashi et al., 2004; Tabor et al., 2009). However, the majority of these are constructed in an *ad hoc* way and our ability to predictably and reliably engineer synthetic biological systems remains quite limited. As summarised in a statistical analysis of the current publications in this field till 2009 (Purnick and Weiss, 2009), the complexity in terms of the number of regulatory promoters in the synthetic circuitries seems to have reached a plateau with a maximum of 6 promoters. The construction of synthetic biological systems with desired properties is still an expensive, and to some extent unpredictable and labourious research process at present (Endy, 2005; Kwok, 2010; Lu et al., 2009), and is far from an engineering routine with lots of trial-and-error. Here I summarize the major obstacles and challenges that remain to be overcome before the engineering of biology becomes routine.

(1) **Standardisation, characterisation and modularity:** The immediate challenges to be tackled are the modularity and reusability problems. Most of the parts and modules used for constructing biological systems are currently undefined, i.e. not standardised and quantitatively characterised, and thus lack modularity and reusability. The current standard for defining parts and modules needs to be revised to improve the modularity and to facilitate the physical and functional assemblies of them. Furthermore, a part or module characterised in one context may change its behaviour in a new working condition and configuration. New standard and method for characterisation need to be developed, which can reduce or eliminate such context dependency issue and thus increases the reusability of the characterisation data across different locations and conditions (Kelly et al., 2009).

(2) **Orthogonality:** Another challenge which needs to be addressed is the so called orthogonality problem. Orthogonality means that the newly added parts and modules should not interfere with existing ones in the designed biological systems as well as
the genetic background of the host chassis, i.e. no crosstalk. Although biological devices are expected to work in a similar manner to their electronic equivalents, a key difference is that unlike electronic digital circuits the individual components are not connected by wires. The parts and devices sit in a melange of different substances, and the interactions of biological components have to depend on the chemical specificity between them. Unfortunately, the toolkit of synthetic biology now contains only a small repertoire of orthogonal regulatory elements such as the LacI, TetR and phage lambda CI regulatory proteins and their cognate promoters. This greatly constrains the development of larger scale systems with high level functions, which normally comprise many components. Thus, a pressing need is to expand the currently limited number of orthogonal parts and modules available in the toolbox and to systematically characterise them in certain model chassis, like *E. coli* and *yeast*. Though most of genes among different species are homologous, the gene evolution process, for example, speciation, gene duplication, gene loss and horizontal gene transfer (Koonin, 2005), has led to the formation of various specialised functional genes or gene clusters in many species. These genes usually exist to enable the hosts to develop some exclusive functions or survive in their corresponding niches such as the virulence related genes in the plant pathogenic *Pseudomonas syringae* and the heat resistant genes in certain thermophilic bacteria. Therefore, these specialised genes are likely to be orthogonal to the chassis of species without the relevant exclusive characteristics. The diverse range of natural building blocks in the myriad of prokaryotic and eukaryotic species provides a rich source from which to engineer orthogonal devices with a variety functions.

(3) **Functional assembly of BioParts:** The functional assembly problem poses another great challenge. The current design principles including rational design, directed evolution, combinatorial synthesis and their hybrids have been shown to work in designing some small scale biological systems, but they still involves a large amount of tinkering work and there are lots of trial-and-error. Our ability to assemble individual biological parts and modules into a customised system with desired functionality remains of low efficiency and lacks enough predictability. To transform gene circuit design and construction into a predictable and reliable engineering
process, more efficient strategies need to be developed to guide the assembly of BioParts into functional biological circuits and the fine-tuning of them.

(4) **Variation:** The fourth great challenge may be the variation problem, which is brought by evolution and noise inherent in biological systems. The engineered biological circuits are subject to potential mutations during the evolution process of the host organisms, which can crash the whole system completely. In addition, the function of the engineered circuit is affected by the cellular context, which continuously evolves according to the environmental conditions as shown in a study (Dekel and Alon, 2005). Thus the variation of the cellular context may lead to the change of circuit behaviour. Another issue is noise, which can make the engineered circuit behave unstably and non-homogeneously. The gene expression is a stochastic process and protein concentrations fluctuate in cells. There is also cell to cell variation. Therefore, new solutions are required to combat the effect of evolution and noise to make synthetic circuits function stably and robustly (Andrianantoandro et al., 2006; Weiss, 2004).

(5) **Supporting technologies and platforms:** Another major challenge is the supporting technologies and platforms associated with the construction and testing of synthetic biological systems. Construction is one bottleneck at present. Synthetic biologists spend a large portion of their time on the cloning and assembling of genetic parts and modules. Thus more efficient standards and methods for the rapid physical assembly of large gene fragments are required. The BioBrick standard assembly (Knight, 2003) and the recent Gibson assembly (Gibson et al., 2009) represent the efforts towards this goal. Technologies which lead to high throughput and cheaper DNA sequencing and synthesis will greatly facilitate the speed of construction and bring down the relevant cost. New measurement technologies and platforms are also required to rapidly and easily probe the circuit dynamics, such as using new gene reporters (Martin et al., 2009) and the microfluidic platforms (Balagadde et al., 2005; Bennett and Hasty, 2009; Cookson et al., 2005; Danino et al., 2010; Gulati et al., 2009) to test kinetic cellular processes and behaviour.
1.4 Objectives of This Study

As described in the previous sections, synthetic biology is still at a relatively early stage. Many challenges and obstacles remain to be overcome before the art of engineering gene circuits has the same predictability and reliability as seen in the design and construction of electronic circuits. Efforts both at the biological and engineering sides are required for the maturation of this interdisciplinary discipline. Novel genetic parts and modules, and efficient design principles for assembling them are needed at present. New gene regulation modes are also expected at each step of the whole gene expression process, from DNAs, RNAs, transcription factors, to sigma factors, RNA polymerases and ribosomes, to achieve tight and flexible control of the genetic information flow in cells. An extensive use of engineering concepts, theories and tools, such as predictive modelling, will be of great value for constructing larger and more complex synthetic gene circuits.

The objectives of this study are to overcome some of the aforementioned major challenges for engineering gene circuits. First, as a way to expand the currently limited toolkit of synthetic biology, the study presents the design and engineering of some novel modular and orthogonal logic devices by redesigning a natural biological module (i.e. the HrpR/HrpS hetero-regulation motif in the \textit{hrp} system of \textit{Pseudomonas syringae}) and using an alternative regulation mode (i.e. the \(\sigma^{54}\)-dependent gene transcription) to achieve tight gene expression control (see Chapter 3). Second, this study examines the effect of context (both biophysical and genetic contexts) on part and device behaviour and determines new methods and standards for the characterisation of parts and devices (see Chapter 4). Third, through the engineering of the designed logic circuits, the study aims to seek and verify a new strategy for functionally assembling synthetic parts and modules in a more predictable manner (see Chapter 3, 4 & 5). Last, the study also aims to design and construct an application-oriented system (i.e. the cell density-dependent \textit{in vivo} biosensor) by systematically integrating the parts and modules engineered in the thesis (see Chapter 6).
Chapter 2

Materials and Methods
2.1 Materials

2.1.1 Media

The media used for growing the bacteria strains in this study are shown in Table 2.1. All media were autoclaved (121 °C for 15 minutes) or filtered (0.22 µm filter) for sterilisation purposes.

Table 2.1 Bacterial growth media used in this study.

<table>
<thead>
<tr>
<th>Medium</th>
<th>Contents</th>
<th>Used for</th>
</tr>
</thead>
<tbody>
<tr>
<td>LB (Luria-Bertani Broth)</td>
<td>10 g/L peptone, 5 g/L NaCl, 5 g/L yeast extract</td>
<td>General growth in liquid media</td>
</tr>
<tr>
<td>LB agar</td>
<td>10 g/L peptone, 5 g/L NaCl, 5 g/L yeast extract, 18 g/L agar</td>
<td>General growth in solid media</td>
</tr>
<tr>
<td>M9 minimal media with 0.4% glycerol</td>
<td>11.28 g/L M9 salts, 1mM thiamine, hydrochloride, 0.4% (v/v) glycerol, 0.2% (w/v) casamino acids, 2mM MgSO4, 0.1mM CaCl2</td>
<td>Characterisation</td>
</tr>
<tr>
<td>M9 minimal media with 0.01% glucose</td>
<td>11.28 g/L M9 salts, 1mM thiamine, hydrochloride, 0.01% (v/v) glucose, 0.2% (w/v) casamino acids, 2mM MgSO4, 0.1mM CaCl2</td>
<td>Characterisation</td>
</tr>
</tbody>
</table>

2.1.2 Antibiotics

The Antibiotics used are summarised in Table 2.2. They were filter sterilised and used at the following target concentrations in media.

Table 2.2 Antibiotics used in this study.

<table>
<thead>
<tr>
<th>Antibiotics</th>
<th>Concentrations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ampicillin</td>
<td>25 µg/ml</td>
</tr>
<tr>
<td>Kanamycin</td>
<td>25 µg/ml</td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>25 µg/ml</td>
</tr>
<tr>
<td>Tetracycline</td>
<td>10 µg ml</td>
</tr>
</tbody>
</table>
2.1.3 Kits

The commercially available kits in Table 2.3 were used in this study.

Table 2.3 Kits used in this study.

<table>
<thead>
<tr>
<th>Kit</th>
<th>Company</th>
<th>Used for</th>
</tr>
</thead>
<tbody>
<tr>
<td>QIAprep Spin Miniprep Kit</td>
<td>QIAgen</td>
<td>Purification of plasmid DNA</td>
</tr>
<tr>
<td>QIAquick PCR Purification Kit</td>
<td>QIAgen</td>
<td>Gel extraction of DNA</td>
</tr>
<tr>
<td>E.Z.N.A. Plasmid Miniprep Kit</td>
<td>Omega Bio-tek</td>
<td>Purification of plasmid DNA</td>
</tr>
</tbody>
</table>

2.1.4 Dyes, chemicals and enzymes

The commercially available dyes, chemicals and enzymes in Table 2.4 were used in this study.

Table 2.4 Dyes, chemicals and enzymes used in this study.

<table>
<thead>
<tr>
<th>Dyes, chemicals and enzymes</th>
<th>Company</th>
<th>Used for</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYBR Safe</td>
<td>Invitrogen</td>
<td>DNA gel stain</td>
</tr>
<tr>
<td>5× loading buffer</td>
<td>Sigma</td>
<td>DNA gel loading</td>
</tr>
<tr>
<td>100 bp/1 kb DNA ladder</td>
<td>Invitrogen</td>
<td>DNA gel electrophoresis</td>
</tr>
<tr>
<td>Isopropyl-D-1-thiogalactopyranoside (IPTG)</td>
<td>Fisher Scientific</td>
<td>Induction of gene expression</td>
</tr>
<tr>
<td>arabinose</td>
<td>Sigma</td>
<td>Induction of gene expression</td>
</tr>
<tr>
<td>N-(3-Oxohexanoyl)-L-homoserine lactone (AHL)</td>
<td>Sigma</td>
<td>Induction of gene expression</td>
</tr>
<tr>
<td>5× M9 minimal Salts</td>
<td>Sigma</td>
<td>Making M9 media</td>
</tr>
<tr>
<td>Thiamine hydrochloride</td>
<td>Sigma</td>
<td>Making M9 media</td>
</tr>
<tr>
<td>Casamino acids</td>
<td>Fisher Scientific</td>
<td>Making M9 media</td>
</tr>
<tr>
<td>D-glucose</td>
<td>Sigma</td>
<td>M9 media carbon source</td>
</tr>
<tr>
<td>Glycerol</td>
<td>VWR BDH</td>
<td>M9 media carbon source</td>
</tr>
<tr>
<td>Dulbecco’s Phosphate Buffered Saline, Ca²⁺ and Mg²⁺ free (D-PBS buffer)</td>
<td>Invitrogen GIBCO</td>
<td>Cell washing and dilution</td>
</tr>
<tr>
<td>Cloned Pfu DNA polymerase</td>
<td>Stratagene</td>
<td>for FACS assay</td>
</tr>
<tr>
<td>Restriction enzymes; T4 DNA ligase</td>
<td>New England Biolabs</td>
<td>DNA amplification in PCR</td>
</tr>
</tbody>
</table>
2.1.5 Bacterial strains

All strains used in this study are listed in Table 2.5. Bacterial strains were streaked for single colonies from glycerol stocks onto solid agar media and grown overnight at 37 °C. Liquid cultures (see Section 2.3.1) were inoculated with single bacterial colonies and grown overnight at 37 °C with shaking (200 rpm). The day liquid cultures were then inoculated from the overnight cultures and grown at 37 °C or 30 °C with shaking (200 rpm).

Table 2.5 Bacterial strains used in this study.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>XL1-Blue</td>
<td>endA1 gyrA96 thi-1 recA1 lac glnV44 F' [::Tn10 (Tet') proAB' lacI'] (\Delta(lacZ)M15) hsdR17((\tau^c) mK(^+))</td>
<td>Stratagene</td>
</tr>
<tr>
<td>MG1655</td>
<td>F(^+) ihG- rfb-50 rph-1</td>
<td>Laboratory</td>
</tr>
<tr>
<td>MC4100</td>
<td>F [araD139] (\Delta(argF-lac)169) (\lambda^c) e14- flhD5301 (\Delta(fruK-yeiR)725) (araA25) relA1 rpsL150(strR) rbsR22 (\Delta(fimB-fimE)632) (::IS1) deoC1</td>
<td>Laboratory</td>
</tr>
<tr>
<td>MC1061</td>
<td>F (\Delta(ara-leu)7697) [araD139] (\Delta(codB-lacI)3) galK16 galE15 (\lambda^c) e14- mcrA0 relA1 rpsL150(strR) spoT1 mcrB1 hsdR2((\tau^c) m(^+))</td>
<td>Laboratory</td>
</tr>
<tr>
<td>TOP10</td>
<td>F mcrA (\Delta(mrr-hsdRMS-mcrBC)) (\phi80lacZAM15) (\Delta lacX74) nupG recA1 araD139 (\Delta(ara-leu)7697) galE15 galK16 rpsL(St(^R)) endA1 (\lambda^c)</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>DH5(\alpha)</td>
<td>F endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG</td>
<td>Laboratory</td>
</tr>
<tr>
<td>BW25113</td>
<td>(rrnB) DE(lacZ4787 HsdR514 DE(araBAD)567 DE(rhaBAD)568 rph-1)</td>
<td>Keio</td>
</tr>
<tr>
<td>BL21-Gold(DE3)</td>
<td>E. coli B F(^+) ompT hsdS(rB(^-) mB(^-)) T7 gene 1 ind1 sam7 nin5) (\text{Tet'}) gal (\lambda)(DE3 [lacI lacUV5-]) endA4 Hte</td>
<td>Stratagene</td>
</tr>
</tbody>
</table>

2.2 DNA Methods

An overview of DNA manipulations

The target genes used in this study were first either synthesised (by GENEART; see Section 2.2.8) according to their sequenced sequences in the public database (e.g. http://pseudomonas.com) or amplified from chromosomal or plasmid DNA (see
Section 2.2.1 for plasmid DNA purification) via Polymerase Chain Reaction (PCR; see Section 2.2.2) using gene-specific sets of primers harbouring desired restriction sites and modifications. DNAs obtained from PCR were subjected to agarose gel electrophoresis (see Section 2.2.3) and the target DNA bands were extracted and gel purified (see Section 2.2.4). The purified DNA from either plasmid or PCR was digested with appropriate restriction enzymes (see Section 2.2.5) and ligated into the cloning or other desired vector (see Section 2.2.6). The DNA construct was then verified by DNA sequencing (see Section 2.2.8). Following transformation into the target cell strain using the heat-shock method (see Section 2.2.10), several clones were selected for antibiotic resistance and re-streaked on fresh LB agar plate under antibiotic selection. The purified plasmids from these clones were again checked by restriction digestion before being used in subsequent studies. Typically, DNA was run on 1.0% (w/v) agarose (BDH) gels in 1× TBE supplemented with 10 µl SYBR Safe DNA Stain for 100 ml gel (see Section 2.2.3). For short DNAs below 200 bp, 1.2-1.5% (w/v) agarose gels were used for visualising them. For BioBrick constructs, specific assembly strategy was used (see Section 2.2.7). All plasmids and oligo DNAs (primers) used in this study are listed in Appendix A and B.

2.2.1 DNA purification

Plasmid DNAs were purified from target bacterial cells grown in a 5 ml overnight LB culture supplemented with the appropriate antibiotic using the QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer’s instructions. The plasmids were eluted in a final volume of 50 µl deionized water.

2.2.2 Polymerase Chain Reaction (PCR)

Specific primers were used to amplify the desired DNA fragment from plasmid DNA using *Pfu* DNA polymerase (Stratagene). *Pfu* DNA polymerase is a proofreading DNA polymerase and used here to increase the fidelity of DNA synthesis during PCR compared to the normal *Taq* DNA polymerase. The reaction mixture for PCR amplification from a plasmid DNA used in this study is listed in Table 2.6.
Table 2.6 Reaction mix for PCR amplification used in this study.

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount per reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distilled H₂O</td>
<td>38 µl</td>
</tr>
<tr>
<td>10× pfu buffer</td>
<td>5 µl</td>
</tr>
<tr>
<td>dNTPs (10 mM each)</td>
<td>1 µl</td>
</tr>
<tr>
<td>Forward primer (10 µM)</td>
<td>2 µl</td>
</tr>
<tr>
<td>Reverse primer (10 µM)</td>
<td>2 µl</td>
</tr>
<tr>
<td>DNA template (10× diluted miniprep DNA)</td>
<td>1 µl</td>
</tr>
<tr>
<td>Cloned Pfu DNA polymerase (2.5 U/µl)</td>
<td>1 µl</td>
</tr>
<tr>
<td>Total reaction volume</td>
<td>50 µl</td>
</tr>
</tbody>
</table>

The PCR reaction was performed in an Eppendorf Mastercycler gradient thermal cycler PCR machine with the following parameters:

Table 2.7 Operating parameters for PCR reactions in this study.

<table>
<thead>
<tr>
<th>PCR Step</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>94 °C for 2 min</td>
</tr>
<tr>
<td>Denaturation</td>
<td>94 °C for 30 sec</td>
</tr>
<tr>
<td>Annealing</td>
<td>Tm – 4 °C (Tm is the calculated melting temperature of applied primers)</td>
</tr>
<tr>
<td>Extension</td>
<td>72 °C for 60 sec (for 1 kb around fragment)</td>
</tr>
<tr>
<td>Final extension</td>
<td>72 °C for 10 min</td>
</tr>
<tr>
<td>Cycle number</td>
<td>30 (for the middle 3 cycling steps only)</td>
</tr>
</tbody>
</table>

2.2.3 Agarose gel electrophoresis

DNA fragments were separated and analysed using agarose gel electrophoresis. The agarose gels were prepared according to the size of the DNA fragments to be separated. Typically, 1.0% (w/v) agarose (BDH) were dissolved by heating in 100 ml 1× TBE buffer (Sigma). SYBR Safe DNA Gel Stain was added to the melted solution with a final concentration of 1:10,000. The solution was then poured into a gel tank with a comb and left to cool down. DNA samples were mixed with 5× loading buffer (Sigma) and loaded onto the gel together with the appropriate DNA ladder
(Invitrogen). The gel was run at 100 volts in 1× TBE buffer until the dye front of the ladder reached the bottom of the gel or the bands were clearly separated. The gel was then visualised under blue light using a Bio-Rad gel imaging system.

2.2.4 Gel extraction and purification of DNA

If needed, the DNA fragments were extracted from agarose gels and purified using a QIAquick PCR Purification Kit (Qiagen) according to the manufacturer’s instructions. The DNA was finally eluted in 30 µl distilled water.

2.2.5 Restriction digest

7-10 µl of purified plasmid DNA were digested in a 20 µl reaction volume supplemented with the corresponding buffer (New England Biolabs (NEB) or Fermentas) at 37 °C for 1.5 hours using 1 µl per restriction enzyme (NEB or Fermentas). The digest was then analyzed on an agarose gel and the desired DNA fragments were excised and purified using a QIAGEN PCR Purification Kit for subsequent cloning.

2.2.6 Ligation

Following restriction digest, the cut vector and insert DNA fragments (ideal molar ratio of vector to insert = 1:3) were ligated in a reaction volume of 20 µl either at 4 °C overnight using 1µl T4 DNA ligase and the 10× reaction buffer (NEB) or at room temperature for 30 min using 1µl Quick T4 DNA ligase and the 2× reaction buffer. The reaction mix was brought up to a final volume of 20 µl with distilled water and mixed thoroughly. Different ratios of DNA insert to vector were tried if the first ligation failed. After ligation, 10 µl of the reaction solution was used to transform the chemically competent *E. coli* cells and transformants were obtained using the appropriate antibiotic selection.
2.2.7 Assembly strategy for DNA constructs

Besides the normal cloning strategy based on restriction enzyme digest, gel purification and ligation, the BioBrick standard assembly method was used for producing some of the DNA constructs in this study. BioBrick is a standard for interchangeable DNA parts and is invented as a way of applying engineering principles of abstraction and standardisation to the design of composable biological components for easy and standard assembly of them (Knight, 2003; Shetty et al., 2008). BioBrick parts are designed to be flanked with the same BioBrick prefix (carrying EcoRI and XbaI restriction sites) and BioBrick suffix (carrying SpeI and PstI restriction sites) sequences as listed below.

**BioBrick prefix:**  \texttt{GAATTCGCGCCGCTTCTAGAG}

**BioBrick suffix:**  \texttt{TACTAGTAGCCGCGCTGCAG}

The scar between two joint BioBrick parts:  \texttt{TACTAGAG}

The BioBrick standard assembly is based on conventional cloning methods and utilises the compatible sticky sequence (\texttt{CTAG}) generated from the SpeI (\texttt{ACTAGT}) and XbaI (\texttt{TCTAGA}) digested sites. When two BioBrick parts join together, the assembly results in a scar between them, as indicated above, which cannot be cut by XbaI or SpeI restriction enzyme. As shown in Figure 2.1, the BioBrick standard assembly is used to piece two BioBrick parts together, one blue and one green. The two BioBrick parts are buried in their corresponding BioBrick plasmid backones which have the same BioBrick flanking sequences. The blue part was first cut out using EcoRI and SpeI restriction enzymes. The plasmid harbouring the green part was digested with EcoRI and XbaI restriction enzymes to generate a gap and to be ligated with the digested blue part. Following gel electrophoresis and purification, the insert of the blue part and the cut plasmid containing the green part were ligated under appropriate conditions to allow the EcoRI sticky ends to join together and the SpeI sticky end to join together with the XbaI sticky end. In this way, the two BioBrick parts are combined together and a new composite part is generated though the assembly results in a scar in the middle. The composite part itself is a new BioBrick
part, flanked with standard BioBrick prefix and suffix, and can be used to join with other BioBrick parts in the same way to form a larger construct. Thus, BioBrick standard assembly is a simple and easy method to assemble as many BioBrick parts in a desired order. The disadvantage is that the parts and plasmid backbones have to conform to BioBrick standard. In some cases the natural parts themselves contain the four restriction sites used here and so have to undergo synonymous codon exchange either by site directed mutagenesis or *de novo* synthesis. Further, ligating BioBrick parts leads to an 8 bp (base pairs) scar between two joined parts preventing protein fusion formation.

![Figure 2.1 BioBrick Standard Assembly method](Registry of Standard Biological Parts, 2010a).

To increase the cloning efficiency, a three way ligation method was used in this study to assemble two separate BioBrick constructs into one desired vector in a single reaction. As shown in Figure 2.2, the two BioBrick parts (Part 1 in blue plasmid and Part 2 in green plasmid) and the construction plasmid (in red) was cut with
appropriate restriction enzymes and gel purified in parallel. The two parts and the construction vector were then mixed and ligated in one reaction to save time and effort, instead of using the two way ligation approach in two serial steps, to complete the assembly.

![Diagram](image)

**Figure 2.2 Three way ligation method.** The method is used for assembling two separately digested BioBrick parts and one cut BioBrick plasmid backbone in one ligation reaction (Registry of Standard Biological Parts, 2010b).

### 2.2.8 DNA sequencing and synthesis

For sequencing, DNA samples were prepared in a total volume of 15 µl comprising 13.5 µl purified plasmid DNA and 1.5 µl sequencing primer (10 pmol/µl). The samples were sent to and sequenced by Eurofins MWG Operon. The sequences were then analysed using the BioEdit Sequence Alignment Edit software.

For gene synthesis, the DNA sequences were designed according to the specification and flanked with BioBrick standard restriction sites. The designed DNA fragments were sent to and synthesised by GENEART. The synthesised DNAs came back in GENEART commercial cloning vectors and were cut out for subsequent usage. The primers used in this study were ordered from Eurofins MWG Operon. The sequences of the essential gene elements used in this study are listed in Appendix C.
Chapter 2 - Materials and Methods

2.2.9 Competent cell preparation
Chemically competent cells were used in this study and prepared using the calcium chloride method for heat shock transformation. A single colony of the required bacterial strain was inoculated in 5 ml LB media within a 14 ml Falcon tube containing the appropriate antibiotic for overnight growth at 37 °C with shaking (200 rpm). The overnight culture was diluted 100 fold to 200 ml fresh LB media supplemented with appropriate antibiotic and grown to mid log phase (OD$_{600}$ between 0.3 – 0.4) at 37 °C. The culture was then transferred to pre-cooled 50 ml Falcon tubes and incubated on ice for 10 min before being centrifuged (5578 × g, 6 min, 0 °C). The harvested cells were washed gently with cold CaCl$_2$ solution (50 mM) and pelleted by centrifugation (5578 × g, 6 min, 0 °C). The washed pellet was resuspended in cold CaCl$_2$ solution (50 mM), left on ice for 30 min and pelleted again. The pellet was resuspended in 1 ml cold solution (50 mM CaCl$_2$, 15% (v/v) glycerol) and incubated on ice for 2 hours. The cells were then dispensed into 1.5 ml microtubes in 100 µl aliquots and stored at -80 °C for later heat shock transformation.

2.2.10 Heat shock transformation
Either 1 µl plasmid DNA (around 20 ng/µl) or 10 µl ligation mix was added to 50 µl chemically competent cells which had been thawed on ice. The mix was gently mixed and incubated on ice for 30 min before being heat shocked at 42 °C for 90 sec. The cells were then incubated on ice for 2 min and recovered in 0.5 ml LB for 1 hour with shaking (100 rpm) at 37 °C. 100 µl of the recovered cells was spread onto a LB agar plate containing the appropriate antibiotic to select transformants harbouring the plasmid of interest. The plate was incubated overnight at 37 °C.

2.3 In Vivo Assay Methods

2.3.1 Growth of bacteria cells
Unless otherwise stated, the cells were grown in 5 ml LB in 30 ml universal tubes for general purpose and in 4 ml M9 minimal media (glycerol or glucose as the carbon
source) in 14 ml Falcon tubes for characterisation purpose at 37 °C and 200 rpm in an Innova shaking incubator (New Brunswick Scientific). All media were supplemented with appropriate antibiotics related to the plasmids residing in the bacterial cells. The overnight cultures were inoculated with a single colony from a freshly streaked LB agar plate containing the appropriate antibiotics. Day-cultures were inoculated from overnight cultures and diluted to OD$_{600}$ = 0.05, and grown to mid log phase (OD$_{600}$ ~ 0.4 – 0.7) or for required length of time prior to analysis.

For characterisation using a fluorometric assay of GFP synthesis, day-cultures were grown and monitored in 96 well micro-assay-plates. Diluted day-cultures were initially loaded into wells and induced with various levels of inducers in a final volume of 200 µl per well. The plate was then immediately incubated in the microplate reader (BMG POLARstar Omega) at 30 °C or 37 °C as required with programmed shaking (200 rpm, linear mode) between each 20 min cycle of absorbance and fluorescence readings.

2.3.2 Fluorescence assay for gene expression in living cells

As this study involves intensive assay of gene expression and regulation, it is important to choose the appropriate gene reporter. There are several gene reporters that have been widely used for gene expression quantification. For example, the β-galactosidase (β-gal) enzymatic assay has enabled sensitive quantification and high throughput screening of gene expression strength. This reporter requires the addition of exogenous reagents to lyse and penetrate living cells to produce the enzymatic reactions required for the colorimetric enzyme assay. Thus, it is better suited as cell population-averaged assay and cannot be easily adapted for quantification of gene expression within single cells and for long time dynamic assay unless used in a microfluidic set up. Another popular reporter is the luciferase derived from bioluminescent organisms like Vibrio fischeri. However, bacterial luciferase assay either needs the addition of an exogenous substrate by expressing only the luxA and luxB genes or without exogenous substrate by expressing a set of five genes (luxCDABE) of the lux operon to result in the bioluminescence – the blue-green light
Chapter 2 - Materials and Methods

with maximum intensity at 490 nm (Amos, 2004). This enables it can be easily quantified by measuring the brightness of the emitted light. But the expression of several genes imposes extra burden on living cells and the stability of lux-encoded products is not high. In this study, Green Fluorescent Protein (GFP) is used as needed. GFP emits a detectable fluorescent signal after excited with light of specific wavelengths and the addition of exogenous substrate or cofactors is not required. GFP is very stable after being expressed and works well across a wide range of organisms including bacteria, yeast, mammals, etc. Therefore, GFP can be used in vivo for population level measurement of expression strength using fluorometry and for high throughput single-cell level quantification using fluorescence activated cell sorting (FACS). In addition, it normally does not interfere with the growth of the host and can be dynamically assayed using a time-lapse microscopy or programmed robotic fluorometer.

2.3.2.1 The GFP reporter gene – gfpmut3b (encoding GFPmut3b)

Green fluorescent protein was originally isolated from jellyfish Aequorea Victoria and contains an intrinsic chromophore to emit the fluorescence after excitation. It only needs oxygen to mature and have little or very low toxicity on the cell host. The wild type (WT) GFP has an excitation spectrum peak at 395 nm with a minor peak at 475 nm, and an emission peak at 509 nm with a minor peak at 475 nm (Amos, 2004). In order to be adapted for the various research purposes, there are many variants derived from the WT GFP, like GFPmut2, GFPmut3, EGFP, etc. These mutants usually have mutations around the chromophore region (amino acids 65-67: Ser-Tyr-Gly) of the protein, which results in the change of maturation time, the shift of excitation and emission spectrum. The mutant GFP used in this study is GFPmut3b carrying triple mutations from the WT one (S2R, S65G, S72A), which is 20 times more fluorescent than the WT one when excited at 488 nm and folds more efficiently when expressed in E. coli (Andersen et al., 1998; Cormack et al., 1996). Thus, GFPmut3b can be excited by the normally used light source – Argon laser with excitation light wavelength of 488 nm, well suited for the fluorometry and FACS based assays used throughout this study.
2.3.2.2 Cell population-averaged assay using fluorometry

Population-averaged fluorescence measurement is used for characterising promoter and device output response in this study by a BMG POLARstar microplate fluorometer. 4 ml M9 cultures (0.4% glycerol or 0.01% glucose as the carbon source where appropriate) were inoculated with single colonies from freshly streak plates and supplemented with the appropriate antibiotics. The overnight cultures after 16 h growth were diluted into fresh pre-warmed M9 media containing the right antibiotics in a final volume of 4 ml with target concentration of OD$_{600}$ = 0.05. The diluted sample-cultures were dispensed as 195 µl per well (190 µl per well for culture requiring double inductions) in a 96 well micro-assay-plate (Bio-Greiner, chimney black, flat clear bottom). The loaded sample cultures were rapidly induced by addition of 5 µl per well of each appropriate inducer in various concentrations, which resulted in various final induction levels, using a multichannel pipette. The blank wells were loaded with 200 µl per well M9 media, and the negative and positive control wells were loaded with 200 µl per well appropriate diluted cultures of OD$_{600}$ = 0.05. The induced plate was then immediately placed into the fluorometer with a programmed protocol of repeated absorbance (OD at 600 nm) and fluorescence (485 nm for excitation, 520 ± 10nm for emission, Gain = 1000) readings. There was continuous shaking (200 rpm, linear mode) between repeated measurements (20 min/cycle) and the plate incubation was set at 30 °C or 37 °C as required. To compensate the effect of evaporation, absorbance readings at a specific time were calibrated according to the tested evaporation speed (4 µl/h per well at 30 °C, 8 µl/h per well at 37 °C). The medium backgrounds of absorbance and fluorescence were determined from the wells loaded with blank M9 media and were subtracted from the readings of other wells. The calculated fluorescence per OD$_{600}$ at a specific time for a sample culture was determined after subtracting its triplicate-averaged counterpart of the negative control cultures (GFP-free) at the same time. The data were first processed in BMG Omega Data Analysis Software (v1.10) and were analysed in Microsoft Excel 2007 and MathWorks Matlab (R2010a) after exported.
2.3.2.3 Single-cell level assay using FACS

To investigate the heterogeneity of gene expression amongst individual cells, FACS (fluorescence activated cell sorting) was used in this study for high throughput single-cell quantification of GFP expression. The BD (Becton-Dickinson) FACSCalibur flow cytometer was used to assay the bacterial samples here. BD FACSCalibur having an Argon blue laser for excitation at 488 nm and the FL1 detector centred at 530 nm with 30 nm bandpass filter fits well for the GFP fluorescence quantification.

Protocol:

The overnight cultures grown in 4 ml M9 media in Falcon tubes were diluted into fresh pre-warmed M9 media (supplemented with appropriate antibiotics) in a final volume of 4 ml with OD\textsubscript{600} = 0.05 in 14 ml Falcon tubes. The diluted day cultures were induced with appropriate inducers and placed in an Innova shaking incubator at 30 °C or 37 °C as required with 200 rpm shaking. After 5 hours when cells were grown to mid log phase (OD\textsubscript{600} around 0.3 – 0.5), cells were pelleted using centrifugation (5578 × g, 6 min). The supernatants were removed and the pellets were resuspended in 3 ml filtered dPBS buffer (0.22 µm filter) before being transferred to BD Falcon 5 ml round bottom tubes. The cultures were then placed on ice prior to be analysed by the BD FACSCalibur flow cytometer. The pre-warmed up flow cytometer (Argon laser 488 nm, FL1 detector 530/30 nm) were tuned with dPBS and negative control cultures (same strain but GFP free) till appropriate instrument settings settled using the BD CellQuest Pro software on a MAC workstation. The settings for forward and side scatter detectors were adjusted to place the acquired cells in the proper location on the scatter graph. The gain for FL1 detector was further tuned using the negative (GFP-free) control culture to let the cellular fluorescence of the negative control distributed within the first decade under log mode. Thus the FACS assay of the negative control under such configuration was treated as the background auto-fluorescence of the cells. Normally 20,000 total events were collected for each culture with low flow rate at room temperature. The acquired data were analyzed using the FlowJo software (v7.2 for windows platform) with an appropriate gate on the forward-scattering and side-scattering graph for all cultures.
2.4 Modelling and Data Analysis Methods

Modelling approaches and system analysis techniques have been widely used not only for the building of complex engineering systems, but, also, for studying the underlying mechanisms of many complicated biological systems, especially in the case of systems biology studies. Recently, computational modelling as an essential engineering tool has also been extensively applied to the construction and analysis of engineered gene circuits (Elowitz and Leibler, 2000; Gardner et al., 2000; Hasty et al., 2001) to increase the probability to succeed, improve the robustness of constructed systems and reduce the cost through quantitative prediction. There are numerous modelling approaches which can be used to model biological systems. The choice of modelling method is generally determined by the questions studied. The two modelling approaches widely used to model gene expression are the deterministic continuous model based on ODEs (ordinary differential equations) and the stochastic discrete model based on chemical master equation. Gene expression stochasticity has been demonstrated in some gene regulation networks (Choi et al., 2008; Elowitz et al., 2002; Kaern et al., 2005) and can lead to some interesting experimental phenomena like bistability (Becskei and Serrano, 2000; Veening et al., 2008). In this study it is mainly focused on the average behaviour of the prokaryotic cell (E. coli) populations to demonstrate the performance of our engineered circuits, both temporal and at a steady state; therefore the deterministic model based on ODEs is used for the mathematical modelling here. The model formulation and the essential biochemical kinetics for modelling biological reactions are introduced below before a simple gene expression modelling example is illustrated at the end.

2.4.1 Deterministic approach using ODE-based rate equations

By assuming the reactions occur in a well mixed system, the production and depletion of each species in the system could be described by an ordinary differential equation. The dynamics of the whole system then can be represented by a set of coupled ODEs:

\[
\frac{dx}{dt} = f(x, t, \theta)
\]  

(2.1)
where \( \mathbf{x} \) is a vector of the concentrations of reacting species, \( \mathbf{\theta} \) is a vector of model parameters, and \( f(\mathbf{x}, t, \mathbf{\theta}) \) is a vector of function describing the rate expressions of each species.

This approach is deterministic, as given the same parameter values and initial conditions, different rounds of simulations will generate the same time evolution for each species in the system.

2.4.2 Basic biochemical kinetics

As for the kinetic modelling of a biological system, it is treated that the system is composed of a series of biochemical reactions, whose kinetics can be described by rate expressions. The multiple reactions in living system span from elementary ones to enzyme catalysed ones, which can be modelled using biochemical kinetic laws. In the following, the essential kinetic laws for modelling biochemical reactions are described, i.e. the law of mass action, Michaelis-Menten enzyme kinetics and Hill kinetics.

**Law of Mass Action**

If the biochemical reactions in the system are elementary reactions, the rate of reactions can be described by the mass action law that the reaction rate is proportional to the product of the reactant concentrations, as shown in the following.

\[
\text{reaction: } A \xrightarrow{k} B \quad \text{reaction rate: } \frac{d[A]}{dt} = -k[A] = -\frac{d[B]}{dt} \tag{2.2}
\]

**Michaelis-Menten Kinetics**

Commonly, biological reactions are complicated and enzyme-catalysed. In this way, one reaction model used to describe the enzymatic reactions is the Michaelis-Menten kinetics. In this model, it assumes that the enzyme is not consumed and its total concentration level stays constant. The enzyme only interacts with the substrates to form an enzyme-substrate (E:S) complex, which leads to the synthesis of product as shown below.

\[
E + S \xrightleftharpoons[k_r]{k_f} E:S \xrightarrow{k_p} E + P
\]
Suppose that the intermediate complex (E·S) is at the quasi-steady state and the substrate is far more than the enzyme. The rate of product synthesis could be described by the Michaelis-Menten equation:

\[
\frac{d[P]}{dt} = \frac{V_{\text{max}}[S]}{[S]+K_M}
\]  

where \(V_{\text{max}} (= k_p[E]_T)\), where \([E]_T\) is the total enzyme concentration) is the maximum reaction rate and \(K_M\) is the Michaelis-Menten constant (\(K_M = (k_r + k_p)/k_f\)), the concentration of substrate needed for reaching half maximal expression.

**Hill Kinetics**

If the enzyme has multiple binding sites and can bind substrate simultaneously, the Hill kinetics is commonly used to derive the reaction rate:

\[
E + nS \xrightarrow{k_1} nS\cdot E \xrightarrow{k_2} E + P \quad \frac{d[P]}{dt} = \frac{v_{\text{max}}[S]^n}{[S]^n + K_M^n} \quad \text{(Hill equation)} \quad (2.4)
\]

where \(n\) is the Hill coefficient, which describes the cooperativity of the reacting substrate. For \(n = 1\), Hill kinetics is reduced to Michaelis-Menten kinetics. Generally, the larger the \(n\), the sharper the response curve described by Hill kinetics will be.

**2.4.3 Modelling a constitutive single gene expression**

\[
\begin{align*}
\frac{d[\text{protein}]}{dt} &= k_p[\text{mRNA}] - \gamma_p[\text{protein}] \\
\frac{d[\text{mRNA}]}{dt} &= k_R - \gamma_R[\text{mRNA}]
\end{align*}
\]  

\(\gamma_p\) and \(\gamma_R\) denote the degradation rates of protein and mRNA, respectively.

**Figure 2.3 A single gene expression modelling.** (The picture is by courtesy of (Ozbudak et al., 2002))
Taking a constitutive single gene expression as an example, mRNA molecules are transcribed from DNA template at rate $k_r$ and proteins are translated from each mRNA at the rate of $k_p$ (Figure 2.3). Proteins and mRNA degrade at rates of $\gamma_p$ and $\gamma_r$ respectively. Two coupled ODEs (Equation 2.5) can be used to represent this process in a deterministic manner as what displayed on the right of Figure 2.3. Using this deterministic model, a short time evolution of mRNA and protein is simulated, as shown in Figure 2.4.

![Figure 2.4 Modelling of a single gene expression dynamics.](image)

**Figure 2.4 Modelling of a single gene expression dynamics.** The gene expression was simulated using the deterministic approach based on the ODEs with typical parameter values of gene expression).

Referring to Figure 2.4, the deterministic simulation reveals mRNA synthesis reaches steady state much faster than protein production, the same as what the experiments have demonstrated. Therefore, the model for a single gene expression could be simplified by combing the transcription and translation as a single step using the following ODE equation.

$$\frac{d[protein]}{dt} = k - \gamma_p[protein]$$  \hspace{1cm} (2.6)

Equation 2.6 is derived from Equation 2.5 by following a steady state assumption for the $[mRNA]$ and defining the constant $k$ as $k = k_p k_r / \gamma_r$. This is reasonable if the response delay induced by mRNA transcription is not a concerned issue in the study and also because the mRNAs follow a much faster dynamics than
the proteins. In case that the promoter of the gene is under the regulation of other transcription factors, Michaelis-Menten or Hill kinetics can be applied to the model where appropriate ($k = \alpha \cdot k_{\text{max}} + k_{\text{max}} \cdot [A]^n / (K_M^n + [A]^n)$) for positive regulation by activators and $k = \alpha \cdot k_{\text{max}} + k_{\text{max}} \cdot K_M^n / (K_M^n + [A]^n)$) for negative regulation by repressors, where $[A]$ is the regulator’s concentration, $K_M$ is Hill constant relating to the regulator’s binding affinity, $n$ is Hill coefficient relating to the regulator’s cooperativity, $k_{\text{max}}$ is the maximum expression rate and $\alpha$ is the basal expression rate due to the promoter’s leakage). All modelling work in this study was implemented in MathWorks Matlab (R2010a).
Chapter 3

The Design and Engineering of Modular Biological Devices
In this chapter, the methods which were used to design and engineer novel synthetic biological devices as well as the design of the modular logic devices constructed in this study are presented. In Section 3.1, the current approaches to design modular and orthogonal biological devices, particularly the devices with logic functions, are discussed. In Section 3.2, the two paradigms for bacterial gene transcription, i.e. the $\sigma^{70}$-dependent and $\sigma^{54}$-dependent transcription, are introduced first with a focus on the potential advantages of the latter one for designing transcription-based logic devices. A $\sigma^{54}$-dependent hetero-regulation module in the hrp regulatory system in *Pseudomonas syringae pv tomato* DC3000 are next described. In Section 3.3, a set of modular logic devices are designed on the basis of the described natural biological modules. The transfer functions of the designed genetic devices are then derived in Section 3.4.

### 3.1 Approaches to Design Modular and Orthogonal Biological Devices

In principle, fully characterised modular biological devices have standardised interfaces and can be easily incorporated into larger and more complex systems. However, most of the synthetic gene circuits constructed currently are just for their own specific purposes in a particular context and they often lack modularity and reusability. In addition, these biological circuits are usually constructed from a limited number of commonly used regulatory components, like the LacI, TetR and phage lambda CI repressor proteins and their regulatory promoters (Lu et al., 2009). Thus, only a small repertoire of orthogonal regulatory elements are available in the current toolkit of synthetic biology. This constrains the development of more complicated systems that might comprise many components because the use of non-orthogonal components in one system is likely to lead to unintended interactions. Since biological systems lack the same physical isolation as occurs in electronic and mechanical engineering, the interactions of biological components have to depend on the chemical specificity between them. There is a pressing need to expand the synthetic biology toolkit of available parts and modules which are truly modular and orthogonal.
As for biological systems, natural genetic networks are typically described as circuits of interconnected functional modules consisting of interacting DNAs, RNAs, proteins and small molecules. Taking the prokaryotic transcriptional regulatory module as an example, the module generally consists of three major elements: a promoter region; the gene or genes expressed from that promoter (a gene sequence usually comprises several functional parts – ribosome binding site, protein coding sequence and terminator); and the transcriptional factor proteins that bind to their cognate sites in the promoter to positively or negatively regulate the expression of that gene(s), as shown in Figure 3.1.

![Figure 3.1 Simplified architecture of a prokaryotic transcriptional regulatory module](image)

**Figure 3.1 Simplified architecture of a prokaryotic transcriptional regulatory module** (TF stands for transcription factor protein. →, ←, ↔, → stand for the promoter, ribosome binding site (RBS), protein coding sequence and terminator respectively).

Owing to this inherent modularity in biology, a biological component can be treated as a module with a specific function, so long as its sensitivities to abiotic factors (e.g. heat, salt, pH) and biotic factors (e.g. proteases, chaperones, competing ligands) are sufficiently well catalogued and understood. Many functional sub-modules are then assembled together to construct a larger module with a more complicated function. Referring to Figure 3.1, a promoter, RBS, protein coding sequence and a terminator, connected in a row, construct a protein generator module, of which both the input and output are protein concentrations.

It would be ideal that the designed biological modules have standard inputs and outputs in order to promote their reusability and modularity. The inputs and outputs are best if they are generic. Proposed units supporting modular design in the synthetic
biology community currently include **PoPS** (polymerase per second) for transcription-based devices (Canton et al., 2008) and **RiPS** (ribosomes per second) for translation-based devices, as shown in Figure 3.2. PoPS represents the rate at which RNA polymerase moves past a given position in the DNA. RiPS is used as a standard unit to measure translational activity of an mRNA molecule and is related to rate of ribosome initiation and elongation. In some sense, they can be thought of as analogous to current flowing through a particular point in a wire. By using PoPS or RiPS as the inputs and outputs of designed modules, they could be arbitrarily hooked up together to compose more complex devices or systems. Thus, devices using PoPS or RiPS as the signal carrier are composable and modular. Modular devices are strongly suggested in this study for designing synthetic gene circuits with reusability and modularity. However, both PoPS and RiPS cannot be measured directly during an experiment. They are usually estimated from protein concentration, e.g. the GFP level, in practice. As the devices and modules designed in this study are transcription-based, the modular design based on PoPS was adopted here but using protein concentration level as the signal carrier for both the input and output, which can be directly derived from the corresponding characterisation data.

**Figure 3.2** Modular design supports standard inputs and outputs of the designed biological modules.

There are normally two ways to develop modules with orthogonality, i.e. the evolution and the genetically distinct natural modules. Here the orthogonality means
that the modules should not interfere with existing parts and modules in the designed biological systems as well as the genetic background circuit of the host. For the pathway using evolution, the original functional module is generally subject to mutations in positions with interaction specificity to generate alternative mutants which can keep the basic function of the original module but cannot interact with it, i.e. lose an energetically favourable interaction with the original partner molecules. As shown in a study of developing orthogonal variants of the LacI and its operator Olac interacting pair (Zhan et al., 2010), several of these non-interacting variant repression pairs can be used in one system to develop transcriptional logic gates. However, many biological modules are usually not well studied and the space of evolution is generally large. This method is time-consuming and of low efficiency and requires many rounds of screening. For the pathway using genetically distinct natural modules, the module components are generally derived from sources with different cellular background (i.e. different species). Like the tetR and cI genes commonly used in the E. coli chassis, they are not endogenous genetic elements of the host and thus likely orthogonal to the host background genetic circuit, whereas the lacI gene is endogenous and non-orthogonal to E. coli and should be used in specific E. coli strains which have mutated lac operon on the chromosome to avoid potential interference with the chassis background. This method has great advantages as the natural diversity has generated thousands of genetically distinct species and normally there are thousands of different genetic modules of specific functions within each species. Other successful examples of this method include the Vibrio fischeri LuxR/LuxI (Basu et al., 2005; You et al., 2004) and Pseudomonas aeruginosa LasR/LasI (Balagadde et al., 2008; Brenner et al., 2007) quorum sensing elements, and the T7 and T3 RNA polymerases and their cognate regulatory promoters (Anderson et al., 2007; Friedland et al., 2009).

For the study in this thesis, the second method using genetically distinct natural modules was applied to develop novel modular and orthogonal devices as a way to expand the limited toolkit of synthetic biology at the moment. In contrast to the currently dominated σ^{70}-dependent gene regulation used in constructing synthetic parts and modules, a different regulatory mechanism of gene activation (i.e. the σ^{54}-dependent gene transcription) was exploited to engineer transcriptional logic devices,
which allows more tight control of the underlying regulation since its transcription system is fully dependent on activation and is otherwise very much in an off state. The difference between these two paradigms of bacterial gene transcription is introduced in the following section.

3.2 The Two Paradigms of Bacterial Gene Transcription and the hrp Regulatory System in Pseudomonas syringae

3.2.1 The two mechanistic paradigms of bacterial gene transcription

Transcription is one of the fundamental processes in biology and a major control point for regulating gene expression to establish coordinated responses of living cells in an ever changing environment. The regulation of transcription is often achieved by the control of transcription initiation. The transcription initiation in bacteria is a multistep process which involves several components, i.e. the multisubunit RNA polymerase (RNAP), sigma factor (σ) protein and the target promoter DNA. The bacterial RNAP core enzyme comprises five conserved subunits (α₂ββ΄ω; E) and associates with a range of σ factors to form the RNAP holoenzyme (Eσ) which can bind specific promoters (Schumacher et al., 2006). The dissociable σ factor in the holoenzyme is responsible for promoter recognition and thus determines which genes are transcribed. There are two major classes of σ factors in bacteria, σ⁰ and σ⁵⁴, which confer different regulatory properties to the core RNAP enzyme for transcription initiation. This represents the two distinct paradigms of bacterial gene transcription (Figure 3.3).
Referring to Figure 3.3a, the $\sigma^{70}$-dependent gene transcription represents the first paradigm in which the $\sigma^{70}$-RNAP holoenzyme (E$\sigma^{70}$) binds the $\sigma^{70}$ specific promoter at conserved positions -35 (TTGACA) and -10 (TATAAT) from the transcription start site (TSS) at +1. The E$\sigma^{70}$-promoter DNA complex can spontaneously isomerise to form the open complex competent for transcription. The regulation of this type of gene transcription is usually achieved by regulatory proteins that either recruit (for activators) or obstruct (for repressors) the RNAP to the promoter sites. Referring to Figure 3.3b, the $\sigma^{54}$-dependent gene transcription represents the second paradigm in which the $\sigma^{54}$-RNAP holoenzyme (E$\sigma^{54}$) binds the $\sigma^{54}$ specific promoter at conserved positions -24 (GG) and -12 (TGC) from the TSS (Schumacher et al., 2006). The E$\sigma^{54}$-promoter DNA complex cannot spontaneously isomerise to form the open complex and is instead transcriptionally silent (Buck et al., 2006; Joly et al., 2010; Wigneshweraraj et al., 2008).
Figure 3.4 The $\sigma^{54}$-dependent transcription activation. The cartoon is adapted from (Buck et al., 2006).

The activation of E$\sigma^{54}$-promoter DNA complex requires the bacterial enhancer binding protein (bEBP), which hydrolyses ATP to catalyse the formation of the open promoter complex (Figure 3.4). The bEBP binds to the upstream enhancer DNA sequences located approximately 150 bp upstream from the TSS and contacts the E$\sigma^{54}$-promoter complex via DNA looping, which is facilitated by the DNA bending protein – integration host factor (IHF). The bEBPs are AAA+ (ATPase associated with various cellular activities) proteins that catalyse ATP hydrolysis and convert the derived chemical energy to a mechanical force needed to remodel the conformation of the closed complex and trigger the formation of the open complex (Buck et al., 2006). Therefore, the $\sigma^{54}$-dependent gene transcription allows more tight control of the underlying regulation, which absolutely requires an activator with the default activity close to zero in contrast to the $\sigma^{70}$-dependent gene transcription. This is particularly useful for developing transcription-based devices for which the low basal expression level is highly appreciated and enables tight and precise control of the output target gene expression.

3.2.2 The $hrp$ gene regulation system in *Pseudomonas syringae*

In this section, a $\sigma^{54}$-dependent hetero-regulation module in the $hrp$ gene regulatory system for Type III secretion in *Pseudomonas syringae* (*P. syringae*) is introduced in an intention to utilise it to design novel orthogonal transcription-based devices. The
Type III secretion system in *P. syringae* is described below before the extracted *hrp* gene regulation network in *Pseudomonas syringae pv tomato* DC3000 is illustrated.

Type III secretion systems (T3SS) underlie the pathogenicity of many Gram-negative bacteria, including both many important animal and plant pathogens. *Pseudomonas syringae* is a widely used and representative plant pathogen in studying plant-pathogen interactions. *P. syringae* elicits leaf spots and foliar necrosis in host plants (e.g. tomato and *Arabidopsis thaliana*) and the hypersensitive response (HR) in non-hosts, which is characterised by the defence-associated rapid programmed cell death of the plant cells in contact with the bacteria. The T3SS in *P. syringae* is encoded by the *hrp* (hypersensitive response and pathogenicity) genes and *hrc* (hypersensitive response and conserved) genes residing in pathogenic gene clusters in the chromosome (Collmer et al., 2000). As shown in Figure 3.5, the Type III secretion system is a complex proteinaceous injection machine that can translocate secreted intracellular effector proteins into the plant cell directly through the assembled secretion apparatus on the surface of the pathogen. The extracellular filamentous appendage named ‘Hrp pilus’ (hypersensitive response and pathogenicity pilus) spans the bacterial membranes, intercellular milieu, the plant cell wall and plasma membrane, and serves as a long-distance transport device (Buttner and Bonas, 2006).

The regulation of *hrp* gene expression in *P. syringae* is primarily through regulatory proteins HrpL (L), HrpR (R), HrpS (S) and HrpV (V) (Figure 3.5). An unknown plant derived signal activates a signalling transduction pathway, which consists of the HrpR, HrpS and HrpL. This leads to the transcription of both the *hrc/hrp* type III secretion genes and the effector genes. The two enhancer binding proteins, HrpR and HrpS, bind upstream of the σ54-dependent *hrpL* promoter to activate the expression of HrpL (a member of the extracytoplasmic function family of σ factors). *hrp* gene expression is negatively regulated by the Lon protease which degrades HrpR (Bretz et al., 2002) and by the negative regulator HrpV – this acts upstream of HrpR and HrpS. The alternative σ factor HrpL then drives the expression of *hrc/hrp* type III secretion genes, which leads to the assembly of the T3SS across the bacterial envelope, plant cell wall, and the plant plasma membrane. The assembled T3SS translocates effector proteins into the plant cell to initiate the hypersensitive
response, or to suppress plant defence, or, possibly, to release nutrients and/or water via unknown mechanisms (Jin et al., 2003).

Figure 3.5 hrp gene regulation in the type III secretion system of *P. syringae* (Jin et al., 2003).
Since the purpose in this study is to utilise transcriptional regulation to mimic the integration of digital logic devices, the focus here is only on the primary regulatory factors that are responsible for regulating the T3SS secretion. The extracted gene regulation model for the hrp regulatory system in *P. syringae* is shown in Figure 3.6. Based on the current knowledge of the hrp gene regulation, the HrpR, HrpS, HrpL and HrpV proteins have been identified as the primary regulatory components for controlling the expression of hrp gene clusters that encode the type III protein export complex. Evidence has shown that HrpR and HrpS are two homologous DNA binding proteins and expressed as an operon¹, and both are required to activate hrpL transcription from the $\sigma^{54}$-dependent hrpL promoter (Hutcheson et al., 2001). HrpL is the primary transcriptional factor controlling the expression of hrp regulon genes. HrpV is a negative regulator for hrp regulon and acts upstream of the HrpR/HrpS-HrpL regulatory cascade (Preston et al., 1998). A recent work reported that the HrpV

---

¹ operon – a genetic unit or cluster that consists of one or more genes that are transcribed as a unit and are expressed in a coordinated manner.
interacts specifically with HrpS to negatively regulate the \( hrpL \) promoter (Jovanovic et al., 2011).

The gene regulation model introduced above has several potential advantages for developing transcription-based genetic logic devices. Firstly, there are multiple regulation factors co-regulating one single promoter (\( hrpL \)) in the system, which increases the flexibility to rewire the network to create multiple-input transcriptional logic devices. Secondly, the central promoter \( hrpL \) is \( \sigma^{54} \)-dependent and therefore allows tight control of the output gene expression with default basal expression activity being very close to zero. Thirdly, the two activator proteins HrpR and HrpS are both enhancer binding proteins, which usually exist in high-order oligomeric forms at their active states and thus can potentially increase the cooperativity and sensitivity of the underlying transcriptional activation.
3.3 Designing a Set of Modular Logic Devices and Functional Assembly of BioParts

In this section, a set of modular transcription-based logic devices are designed on the basis of the \textit{hrp} hetero regulation module introduced in the previous section and other classic biological modules. Particularly, a modular tight-controlled and hypersensitive genetic circuit with digital logic AND function and a combinatorial modular NAND gate circuit are rationally designed. The proposed functional assembly approach for engineering gene circuits with predictable functions is described and discussed in the end.

3.3.1 The design of a modular logic AND gate

The logic AND gate is rationally designed based on the $\sigma^{54}$-dependent HrpR/HrpS hetero-regulation module in the \textit{hrp} gene regulatory system for Type III secretion in \textit{Pseudomonas syringae} (Wang et al., 2009a). As shown in Figure 3.7a, the two-input AND gate comprises two genes (i.e. \textit{hrpR} and \textit{hrpS}) and one regulatory promoter (\textit{hrpL}). The \textit{hrpR} and \textit{hrpS} genes are placed under the control of two separate $\sigma^{70}$-dependent promoters and a reporter gene acts as the measurable output (as GFP protein) in the experiments. Thus the output reporter can now be expressed only when both of the two input promoters are activated otherwise the default output remains close to the zero level. The $\sigma^{54}$-dependent hetero regulation allows the AND gate to be tightly regulated and sensitive. In addition, the design is modular because the inputs and output of the AND gate are both promoters, and the inputs can be reconnected to different input sensors and the output can be used to drive various cellular responses. As illustrated in Figure 3.7b, the function of the AND gate can be tested using two environment-responsive promoters as the inputs and the \textit{gfp} as the output reporter. There is high fluorescent output, i.e. the output promoter \textit{hrpL} is turned on, only when the logic AND combination of input inducers appears as the truth table shows. Figure 3.7c shows the functional sequence of \textit{hrpL} promoter of \textit{Pseudomonas syringae} \textit{pv. tomato} DC3000. The conserved -12 and -24 sites are where the $\sigma^{54}$ binds specifically. The sequence in red is the putative UAS (upstream activator sequence) where HrpR and HrpS bind, and the sequence in bold face is the IHF binding site.
Chapter 3 - The Design and Engineering of Modular Biological Devices

Figure 3.7 The rational design of the modular logic AND gate. The AND gate is devised by redesigning the natural HrpR/HrpS hetero-regulation motif.

More logic AND gate devices can be designed by including the negative regulatory \( \textit{hrpV} \) gene in the circuit (Wang et al., 2009a). A two-input logic AND gate with one-input inverted can be developed by placing \( \textit{hrpR} \) and \( \textit{hrpS} \) as an operon under one input promoter while \( \textit{hrpV} \) under the other input promoter (Figure 3.8a). A three-input logic AND gate with one-input inverted can be developed by placing \( \textit{hrpR} \), \( \textit{hrpS} \) and \( \textit{hrpV} \) under the control of three separate input promoters to co-regulate the output \( \textit{hrpL} \) promoter (Figure 3.8b).
3.3.2 The design of a set of modular NOT gates and the composite NAND gate

Biological modules mimicking digital logic gates are necessary to realise the practical potentials of being able to systematically and synthetically control the genetic information flow within living cells. Therefore here a set of modular genetic NOT gates and a genetic NAND gate, which implement the essential digital logic NOT and NAND functions respectively, were designed to expand the available transcriptional logic devices in the synthetic biology toolkit.

As shown in Figure 3.9a, the modular NOT gate is developed using a biological repressor module (e.g. the lacI/P_{lac}, tetR/P_{tet}, cI/P_{lam} pairs), in which the repressor I binds to its cognate promoter P_{NOT} and inhibit the transcription of the output promoter significantly. The NOT gate function can be characterised under an environment-responsive promoter (I_1-inducible P_1) using a reporter gene like the gfp as the output. As a result, the output level (e.g. fluorescence) is in reverse proportion to the amount of the inducer input. The modular NAND gate (Figure 3.9b) is designed to be a combinatorial circuit comprising an AND gate and a NOT gate. The component AND gate is the same as the one in Figure 3.7 and the NOT gate is the same as the one in Figure 3.9a. Thus it results from the direct coupling together of two gates by harnessing the modularity and reusability of the component modules. The composite NAND gate function can be characterised using two environment-responsive
promoters ($I_1$-inducible $P_1$ and $I_2$-inducible $P_2$) as the inputs and the $gfp$ reporter as the output. It is expected this device has the ‘NOT AND’ function, i.e. the logic NAND gate.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{not_gate_diagram.png}
\caption{The design of a set of modular NOT gates and the composite NAND gate. \textbf{a}, The modular logic NOT gate is composed of a biological repressor module (the $I/P_{\text{NOT}}$ pair). \textbf{b}, The composite logic NAND gate comprises a modular AND gate and a modular NOT gate, and is modular itself.}
\end{figure}

\subsection*{3.3.3 Functional assembly of BioParts using engineered in-context characterised modules}

Synthetic biology uses the idea that modular biological elements (BioParts) exist which can be combined and modified to construct novel devices. Figure 3.10
illustrates this approach, where two biological modules with the AND gate (module I) and NOT gate (module II) functions can be easily combined physically to form a larger composite module. It is expected that the composite module will naturally has the NAND gate function. Here the modularity of biological elements spans two aspects, i.e. physical level and functional level. The physical modularity can be achieved through imposing standard interface on the component modules, like PoPS or RiPS, allowing for the easy exchange and physical assembly of them. The functional modularity assumes that the functions of the modules persist when used in a new condition or configuration, while the functional assembly of them into a large customised system depends on the proper matching of the characteristics of the underlying component modules.

![Figure 3.10 The physical and functional assemblies of synthetic biological modules.](image)

However, rationally designed biological devices often do not initially function as intended and a large amount of trial-and-error are required to tinker with them before the desired performance is achieved. This is partly because the individual modules are not sufficiently well characterised for purpose and the functional assembly of them is not straightforward due to their different characteristics. Moreover, the component modules characterised in one context may vary quite
differently in another working condition (Klumpp et al., 2009; Tan et al., 2009) and configuration (Salis et al., 2009). This is illustrated by the difficulties encountered and the progress made in synthetic biology over the last decade (Lu et al., 2009; Purnick and Weiss, 2009). Although a number of synthetic biological parts with basic functions (Voigt, 2006) have been developed and assembled to construct some proof-of-principle circuits, the majority of them are constructed in an ad hoc way and the construction of synthetic gene circuits with desired functions is far from an engineering routine and requires lots of tweaking. The current design principles for engineering gene circuits, which include rational design (Gardner et al., 2000; Guido et al., 2006), directed evolution (Dougherty and Arnold, 2009; Haseltine and Arnold, 2007; Yokobayashi et al., 2002), combinatorial synthesis (Guet et al., 2002) and their hybrids (Atsumi and Little, 2004; Dueber et al., 2009), have been shown to work in designing some small gene circuits, but often require iterative refining and screenings, do not take into account the impact between cellular context and engineered circuits, and the scale of the networks is usually small. Thus, our ability for engineering gene circuits is limited by the lack of a number of well-characterised interchangeable parts that behave with predictable functions across various contexts and efficient strategies for the functional assembly of individual parts into large scale customisable systems.

To address this issue, a new approach, which uses engineered in-context quantitatively characterised modules, is proposed here for the functional assembly of BioParts into customisable larger systems. As both the biophysical, biological and genetic contexts in which the biological parts and modules behave might well be expected to have a large impact on their functionalities, the candidate component parts and modules for a target synthetic system ideally should be characterised in those biophysical (e.g. media, temperature, carbon source) and genetic (e.g. embedded sequence context, chassis background) contexts as anticipated for their final system(s). This approach minimises the unexpected or high-order effects which could occur during circuit construction by eliminating or reducing the variations arising from the difference of contexts. We can also characterise the synthetic parts and modules in various contexts and model their behaviour correspondingly. The behaviour of the assembled circuit comprising characterised parts and modules in a
certain context then can be predicted reliably from the individual models of the components in the same context. As a result, the functional assembly of synthetic parts into customisable larger systems can be performed more predictably using these engineered in-context characterised modules aided by modelling.

3.4 Deriving Transfer Functions of the Designed Logic Devices

One of the aims of this study is to develop models of individual parts and modules to allow the predictable assembly of them into customised systems. The deterministic model based on ODEs is used for the mathematical modelling of the gene expression and regulation in this study as described in Chapter 2. The following describes the derivation of the transfer function models for the modules designed in the previous section.

3.4.1 Deriving the transfer function of environment-responsive promoters

Since environment-responsive promoters function as the inputs of the designed genetic devices, it is necessary to derive their transfer function first. As the case in the following diagram shows (Figure 3.11), the promoter \( P_1 \) is negatively regulated by its constitutively expressed repressor \( R_1 \) and responsive to the exogenous inducer \( I_1 \) to turn on the transcription of the downstream reporter gene \( G \).

![Figure 3.11 The architecture of an environment-responsive negatively regulated promoter (\( P_1 \)).](image)
The reporter gene expression can be modelled by (Alon, 2007; Zoltan et al., 2006):

$$\frac{d[G]}{dt} = \alpha_i \cdot k_i + \frac{k_i \cdot [I_i]^{n_i}}{[I_i]^{n_i} + K_i^{n_i}} - d \cdot [G] \quad (3.1)$$

where $\alpha_i \cdot k_i$ is the basal constitutive activity of the promoter, $k_i \cdot [I_i]^{n_i}/([I_i]^{n_i} + K_i^{n_i})$ is the activity due to cooperative transcription activation by assuming the concentration of the repressor is constant to model the effect of varying the concentration of the inducer $I_i$, and $d \cdot [G]$ is the constitutive degradation activity of protein G. Here the concentrations of the inducer $I_i$ and reporter protein G are denoted by $[I_i]$ and $[G]$ respectively; $K_i$ and $n_i$ are the Hill constant and coefficient relating to the promoter-regulator/inducer interaction; $k_i$ is the maximum expression rate of the promoter due to induction and $\alpha_i$ is a constant relating to the basal level of the promoter due to leakage ($0 \leq \alpha_i < 1$); and $d$ is the degradation rate of G.

The steady state solution of Equation 1 is given by:

$$f([I_i]) = [G]_s = k_i' ([\alpha_i + [I_i]^{n_i}/(K_i^{n_i} + [I_i]^{n_i})]$$

where $k_i' = k_i/d$ represents the maximum expression level due to induction. Equation 3.2 gives the reporter protein level at steady state for the inducible promoter $P_1$ and is also the transfer function of $P_1$.

### 3.4.2 Deriving the transfer function of the AND gate

![Figure 3.12 The architecture of the genetic logic AND gate circuit.](image-url)
As shown in Figure 3.12, the σ^{54}-dependent hrpL promoter is synergistically co-activated by the hetero proteins HrpR and HrpS, which mimics the two-input logic AND function. Based on the currently known mechanism underlying this hetero-regulated module, both the bacterial enhancer-binding proteins are required to bind the UAS of hrpL to remodel the conformation of σ^{54}-RNAP-hrpL closed complex to an open one for the activation of transcription. The hrpL regulation in the designed AND gate is represented by the product of two Hill function curves as the following model of the output reporter gene expression shows.

\[
\frac{d[G]}{dt} = \frac{k_i \cdot ([R]/K_R)^n_R \cdot ([S]/K_S)^n_S}{1+(([R]/K_R)^n_R \cdot ([S]/K_S)^n_S)} - d \cdot [G] \quad (3.3)
\]

Thus the normalised transfer function of the AND gate is derived as:

\[
f([R]_s, [S]_s) = \frac{[G]_{s,\text{max}}}{[G]_{s,\text{max}}} = \frac{([R]_s/K_R)^n_R \cdot ([S]_s/K_S)^n_S}{1+(([R]_s/K_R)^n_R \cdot ([S]_s/K_S)^n_S)} \quad (3.4)
\]

where \(K_R, K_S\) and \(n_R, n_S\) are the Hill constants and coefficients for HrpR and HrpS respectively. \([R]_s\) and \([S]_s\) are the steady levels of HrpR and HrpS, whose levels are under the control of two separate inducible promoters (P_1 and P_2) as indicated by Equation 3.2. \([G]_{s,\text{max}} = k_i/d\) is the maximum output level of the AND gate at steady state, in which \(k_i\) is the maximum expression rate due to activation and \(d\) is the degradation rate of G.

### 3.4.3 Deriving the transfer function of the NOT gate

The NOT gate in this study is designed on the basis of biological repressor modules. The R_3/P_3 based NOT gate shown in Figure 3.13 is characterised under the inducible promoter P_1 in response to inducer I_1 with gene G as the output reporter.

![Figure 3.13 The architecture of the genetic logic NOT gate circuit.](image-url)
The reporter gene expression in this NOT gate circuit can be modelled by:

$$\frac{d[G]}{dt} = \alpha \cdot k + \frac{k \cdot K^n}{K^n + [R]^n} - d \cdot [G]$$  \hspace{1cm} (3.5)

Thus the transfer function of the NOT gate is derived as:

$$f([R]_s) = [G]_s = k \cdot (\alpha \cdot K^n + [R]_s^n)$$  \hspace{1cm} (3.6)

where $K$ and $n$ are the Hill constant and coefficient relating to $R/P$ interaction, $k$ represents the maximum expression level of $P$ due to repression relief and $[R]_s$ is the steady levels of $R$, whose level is under the control of the inducible promoter as indicated by Equation 3.2.

3.4.4 Deriving the transfer function of the composite NAND gate

Figure 3.14 shows the architecture of the composite logic NAND gate circuit, which comprises the two environment-responsive promoters $P_1$ and $P_2$ as the inputs, the modular logic AND gate and NOT gate modules, and the reporter $G$ as the output. The transfer function of the NAND gate can be derived by directly coupling the transfer functions of the individual modules, i.e. the NOT gate, AND gate and environment-responsive promoters, in the system. The output of a forward module acts as the input of the next module in the system cascade. Thus the NAND gate transfer function is obtained as:

$$[G]_{\text{NAND}} = f([R]_s) = k \cdot (\alpha \cdot K^n + [R]_s^n)$$  \hspace{1cm} (3.7)
where $k_3'$ represents the maximum expression level of $P_3$ due to repression relief and

$$[R_3]_{ss} = f([R]_{ss}, [S]_{ss}) = [G]_{ANDmax} \frac{([R]_{ss}/K_R)^{\alpha_1} ([S]_{ss}/K_S)^{\alpha_2}}{(1 + ([R]_{ss}/K_R)^{\alpha_1})(1 + ([S]_{ss}/K_S)^{\alpha_2})}$$

where $[G]_{ANDmax}$ is the maximum output level of the AND gate at steady state; $[R]_{ss} = f([I_1]) = k_1'(\alpha_1 + [I_1]^n/(K_1^n + [I_1]^n))$ and $I_1$ is the inducer of the inducible promoter for the regulation of $hrpR$; $[S]_{ss} = f([I_2]) = k_2'(\alpha_2 + [I_2]^n/(K_2^n + [I_2]^n))$ and $I_2$ is the inducer of the inducible promoter for the regulation of $hrpS$ in the system. All other parameters in the model have the same meaning as described in sections 3.4.1-3.4.3.
The objective in this chapter is to verify the functional assembly approach proposed in this thesis by engineering the novel genetic AND gate that was designed in Chapter 3. Both the rational design and the functional assembly approaches have been applied to construct this modular AND gate. By quantitatively characterising a set of parts and modules in various live cell bio-physical and genetic contexts, the impact of the context on their behaviour was illustrated. The engineered AND gate was then subject to thorough characterisation including the homogeneity, metabolic load and chassis compatibility assays.

4.1 The Initial Implementation of the AND Gate

Once the modular genetic AND gate had been rationally designed based on the $\sigma^{54}$-dependent HrpR/HrpS hetero-regulation module in the *hrp* regulatory system in *Pseudomonas syringae*, the next step was to select the appropriate genetic elements to construct it for experimental characterisation and validation. In this section, the rational design approach was applied to implement the AND gate without quantitatively characterising the component elements on route.

Firstly, the two well studied inducible promoters P$_{lac}$ and P$_{BAD}$ were chosen as the driving inputs for the designed AND gate, and the gfp reporter gene as the output. In conforming to the BioBrick standard, the core components *hrpR*, *hrpS* and the promoter *hrpL* were sent for synthesis by GENEART to eliminate the four restriction sites (i.e. EcoRI, XbaI, SpeI and PstI) used in this assembly standard through synonymous codon exchange. As an initial trial, the two genes were flanked with the same ribosome binding site rbsH (Table 4.1) and the BioBrick double terminator BBa_B0015. As shown in Figure 4.1, the *hrpR* and *hrpS* fragments are placed under the control of the IPTG inducible P$_{lac}$ and the arabinose inducible P$_{BAD}$ promoters respectively and the rbs30-gfp is placed downstream of the *hrpL* as the output reporter. The AND gate functions such that the output is on (in terms of high fluorescence) only when both of the inputs are high, i.e. P$_{lac}$ and P$_{BAD}$ are both highly induced by IPTG and arabinose respectively.
Figure 4.1 Schematic diagram for the construction of the genetic AND gate. The engineered AND gate uses inducible promoters $P_{\text{lac}}$ and $P_{\text{BAD}}$ as the inputs, the same RBS (rbsH) upstream the $hrpR$ and $hrpS$ genes, and rbs30-gfp as the output reporter.

Two-input AND gate using $P_{\text{lac}}$ and $P_{\text{BAD}}$ as the inputs

Input 1  
Input 2  
Output

Figure 4.2 Plasmid maps showing the circuit constructs used for the initial version of the AND gate as designed in Figure 4.1.

To construct the AND gate, the synthesised $hrpR$ gene fragment was cloned into plasmid pAPT110 (p15A ori, Kan$^\text{r}$) (Polard and Chandler, 1995) under the IPTG inducible $P_{\text{lac}}$ as the first input; the $hrpS$ fragment was cloned into plasmid pBAD18-cm (pBR322 ori, Cm$^\text{r}$) (Guzman et al., 1995) under the arabinose inducible $P_{\text{BAD}}$ as the second input; while the $hrpL$-rbs30-gfp construct was carried on the pSB4A3 (pSC101 ori, Amp$^\text{r}$) (Shetty et al., 2008) vector as the output of the AND gate. Figure 4.2 shows the plasmid maps for the constructs used for this AND gate. The plasmid constructs were next transformed into $E.\ coli$ MC1061. The result of full characterisation of the device in this cell chassis is shown in Figure 4.3, where the cells were assayed for 64 combinations of the two-input inductions in the fluorometer.
The data are the normalised average of three repeats in *E. coli* MC1061 in M9-glycerol at 30 °C with variations less than 10% between biological replicates. The detailed experimental protocols for the characterisation are described in Chapter 2.

**Figure 4.3 The systematic characterisation of the AND gate.** The response of the AND gate for 64 combinations of input inductions. The data are shown for induction by (left to right) 0, 3.9 × 10^{-4}, 1.6 × 10^{-3}, 6.3 × 10^{-3}, 2.5 × 10^{-2}, 0.1, 0.4 and 1.6mM IPTG, and induction by (bottom to top) 0, 3.3 × 10^{-4}, 1.3 × 10^{-3}, 5.2 × 10^{-3}, 2.1 × 10^{-2}, 8.3 × 10^{-2}, 0.33 and 1.3 mM arabinose.

Since the component parts were not characterised, perhaps not surprisingly, this rationally designed AND gate failed to work precisely as intended as shown in Figure 4.3, where a distorted device behaviour was observed. Difficulties in engineering the functional AND gate circuit are largely due to that the individual components are not sufficiently characterised in the relevant context and the functional assembly of them is not straightforward due to their different characteristics. In addition, the behaviour of components may vary in different working conditions (Klumpp et al., 2009; Tan et al., 2009) and configurations (Salis et al., 2009) because there are many factors which can affect gene expression in living cells, for example, cell chassis, growth medium including carbon source, embedded sequence context like 5' UTR (five prime untranslated region) and RBS (ribosome binding site), temperature. As examining the impact of each of these factors on the behaviour of the engineered circuit is a non-trivial undertaking, it was decided to characterise each basic part and sub-module of the circuit in various contexts (both physical and genetic), which will provide a
4.2 Characterising a Set of Inducible Promoters and RBSs

In this section, a set of basic parts, i.e. the inducible promoters and RBSs, were characterised in various biophysical (e.g. temperature variations) and genetic contexts in order to see to what extent these contexts could affect their behaviour. The characterisation work here serves as the basis for the rational forward engineering of the designed genetic AND gate in Section 4.3.

4.2.1 Promoter characterisation setting up

Due to the requirement of modularity, both the inputs and the output of the AND gate are designed to be promoters. Thus, the inputs can be any interchangeable promoters and the output promoter can be connected to various genes to drive different cellular responses. To verify the required integrative behaviour of the AND gate, three environment-responsive promoters, i.e. the IPTG inducible $P_{\text{lac}}$, the arabinose inducible $P_{\text{BAD}}$ and the N-(3-Oxohexanoyl)-L-homoserine lactone (AHL) inducible $P_{\text{lux}}$ (Figure 4.4), were chosen as the candidate inputs and systematically characterised. The $P_{\text{lac}}$ and $P_{\text{BAD}}$ are endogenous $E.\ coli$ promoters, while the $P_{\text{lux}}$ is a synthetic promoter built from the quorum sensing $lux$ operon in $Vibrio fischeri$ and the tet promoter (Canton et al., 2008) and can be potentially used as the receiver for the AHL signalling molecule secreted by other cells. The two cell strains $E.\ coli$ MC4100 and $E.\ coli$ MC1061, which have destructive mutations of both the $P_{\text{lac}}$ and $P_{\text{BAD}}$ promoters on their genome (Table 2.5, Chapter 2), were chosen as the candidate hosts. The chemically well-defined M9 medium was used and supplemented with either 0.01% glucose (M9-glucose) or 0.4% glycerol (M9-glycerol) as the carbon source. The detailed compositions of the two media are as described in Chapter 2 (Materials and Methods). Two normally used temperatures for $E.\ coli$ culture, 30 °C and 37 °C, were used to evaluate the effect of temperature variations. In addition, for the purpose of the experiments, 6 versions of RBS of various strengths (Table 4.1) were used to
characterise each promoter in a way to find balanced promoter/RBS pairs acting as the driving inputs for the AND gate. Referring to Table 4.1, the first 5 RBSs are from the Registry of Standard Biological Parts (i.e. rbs30, rbs31, rbs32, rbs33, rbs34) and are initially reported to have distinct strengths by the community (Registry of Standard Biological Parts, 2010c). The last one, rbsH, is designed here as a homemade version.

![Diagram of promoters](image)

**Figure 4.4** Schematic diagram for the characterisation of the three inducible promoters: $P_{lac}$, $P_{BAD}$ and $P_{lux}$. The gfp reporter gene ($gfp$mut3b) bearing a set of RBSs with various strengths was used to characterise the three inducible promoters: the IPTG-responsive $P_{lac}$ promoter (top), the arabinose-responsive $P_{BAD}$ promoter (middle) and the AHL-responsive synthetic $P_{lux}$ promoter (bottom). The sequences of RBSs used are listed in Table 4.1. The BioBrick double terminator BBa_B0015 following gfp was used to terminate the transcription in all cases.

**Table 4.1** The RBS sequences used for the promoter characterisation.

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Sequence of RBS</th>
<th>Strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>rbs30 (BBa_B0030)</td>
<td>TCTAGAGTTAAAGAGGAGAGAATACTAGATG</td>
<td>Strong</td>
</tr>
<tr>
<td>rbs31 (BBa_B0031)</td>
<td>TCTAGAGTCACACAGGAGAACTAGATG</td>
<td>Weak</td>
</tr>
<tr>
<td>rbs32 (BBa_B0032)</td>
<td>TCTAGAGTCACACAGGAGAACTAGATG</td>
<td>medium weak</td>
</tr>
<tr>
<td>rbs33 (BBa_B0033)</td>
<td>TCTAGAGTCACACAGGAGAACTAGATG</td>
<td>weakest</td>
</tr>
<tr>
<td>rbs34 (BBa_B0034)</td>
<td>TCTAGAGAAAGAGGAGAGAATACTAGATG</td>
<td>very strong</td>
</tr>
<tr>
<td>rbsH (Homemade)</td>
<td>TCTAGAGAGGAGATATACCAGT</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 4.5 shows the three constructs used for the characterisation of $P_{\text{lac}}$, $P_{\text{BAD}}$ and $P_{\text{lux}}$ with the RBSs of rbsH, rbs33 and rbs33 shown for each. Plasmid pBW105lac-gfp (p15A ori, Kan$^r$) containing the IPTG inducible $P_{\text{lac}}$ was used for $P_{\text{lac}}$ promoter characterisation. Plasmid pBW203ara-gfp (pBR322 ori, Cm$^r$) containing the arabinose inducible $P_{\text{BAD}}$ was used for $P_{\text{BAD}}$ promoter characterisation. The synthetic AHL-inducible $P_{\text{lux}}$ promoter (BBa_F2620 (Canton et al., 2008)) was cloned into pSB3K3 (p15A ori, Kan$^r$) (Shetty et al., 2008) and characterised with plasmid pBW303lux-gfp. All data (fluorescence/OD$_{600}$) were acquired 5 hours after induction from the cells grown in 96 well microplates, when cells were in the phase of exponential growth and the steady state assumption for protein expression is applied. Cells were grown and subjected to continuous shaking (200 rpm, linear mode) and repeated absorbance and fluorescence readings were made. Figure 4.6 shows the cell growth curves and dynamic responses of $P_{\text{lac}}$ promoter in one experimental characterisation over 20 hours. It can be seen that the exponential growth phase lasts several hours, i.e. from the 2 to 5 hours. The fluorescent responses of the promoter first reached to a plateau between the 5 and 8 hours and then decrease slowly over time. Thus the fluorescence/OD$_{600}$ value after 5 hours growth was used to determine the response level of the cells at steady state.

**Promoter characterisation for the input transfer functions**

![Plasmid maps showing some of the circuit constructs used for the characterisation of the three promoter inputs.](image)

**Figure 4.5** Plasmid maps showing some of the circuit constructs used for the characterisation of the three promoter inputs.
Figure 4.6 Dynamics of the P_{lac} promoter in response to various levels of induction. a, Growth curves of the P_{lac}-rbs30-gfp characterisation strain with various IPTG inductions. The host is E. coli MC1061 grown in a 96 well plate in the fluorometer set at 30 °C and with repeating absorbance and fluorescence readings (20 min/cycle). b, The fluorescence/OD_{600} values over time.

4.2.2 Characterisation results and analysis

As a first step, to evaluate the promoter responses in different chassis and media, the three promoters were characterised in response to various induction levels in the two cell chassis (E. coli MC4100 and E. coli MC1061) in the two media (M9-glucose and M9-glycerol) at 30 °C. The gfp with a strong RBS (rbs30-gfp) was used as the output reporter. As shown in Figure 4.7, promoter P_{lac} is nearly open and does not produce the desired inducer dependent switching characteristic from the low to high induction levels in E. coli MC4100 grown in M9-glycerol. The reason behind this phenomenon is complicated and it might be due to an unintended interaction of the P_{lac} with the endogenous genetic background of this chosen host. While in E. coli MC4100 grown in M9-glucose, it can be seen (Figure 4.7b) that promoter P_{BAD} is inhibited a lot due to the catabolite repression effect of glucose although low (0.01%) in the medium. However, the P_{lac} is not sensitive to this effect at this level of glucose and is inducible (Figure 4.7a), and P_{lux} is slightly inhibited (Figure 4.7c). While in E. coli MC1061 grown in M9-glycerol, all the three promoters responded with the desired switching
Figure 4.7 The dose responses of the $P_{\text{lac}}$ (a), $P_{\text{BAD}}$ (b) and $P_{\text{lux}}$ (c) promoters to various induction levels in the two cell chassis ($E. coli$ MC4100 or $E. coli$ MC1061) in M9 media (M9-glycerol or M9-glucose). a, $P_{\text{lac}}$ induced by (left to right) $0$, $3.9 \times 10^{-4}$, $1.6 \times 10^{-3}$, $6.3 \times 10^{-3}$, $2.5 \times 10^{-2}$, $0.1$, $0.4$ and $1.6$ mM IPTG. b, $P_{\text{BAD}}$ induced by (left to right) $0$, $3.3 \times 10^{-4}$, $1.3 \times 10^{-3}$, $5.2 \times 10^{-3}$, $2.1 \times 10^{-2}$, $8.3 \times 10^{-2}$, $0.33$ and $1.3$ mM arabinose. c, $P_{\text{lux}}$ induced by (left to right) $0$, $2.4 \times 10^{-2}$, $9.8 \times 10^{-2}$, $3.9 \times 10^{-1}$, $1.6$, $6.3$, $25$ and $100$ nM AHL. The data (fluorescence/OD$_{600}$) were the average of three independent repeats with error bars denoting the standard deviations.
Figure 4.8 The characterisation results of $P_{lac}$ (a), $P_{BAD}$ (b) and $P_{lux}$ (c) using 6 versions of RBS in response to various inductions (by 0, $3.9 \times 10^{-4}$, $1.6 \times 10^{-3}$, $6.3 \times 10^{-3}$, $2.5 \times 10^{-2}$, 0.1, 0.4, 1.6, 6.4 and 12.8 mM IPTG; 0, $3.3 \times 10^{-4}$, $1.3 \times 10^{-3}$, $5.2 \times 10^{-3}$, $2.1 \times 10^{-2}$, $8.3 \times 10^{-2}$, 0.33, 1.3, 5.3 and 10.7 mM arabinose; and 0, $1.5 \times 10^{-3}$, $6.1 \times 10^{-3}$, $2.4 \times 10^{-2}$, $9.8 \times 10^{-2}$, $3.9 \times 10^{-1}$, 1.6, 6.3, 25 and 100 nM AHL respectively), and the fits to a Hill function model (solid lines) of the promoter response. All data (fluorescence/OD$_{600}$) were the average of three independent repeats in *E. coli* MC1061 in M9-glycerol at 30 °C. Error bars, s.d. (n = 3).
characteristics although their maximum response levels in terms of fluorescence/OD<sub>600</sub> are different. Thus, this context (E. coli MC1061, M9-glycerol, 30 °C) was selected as the standard condition for the characterisation in the subsequent stages.

In the second step, in search of the appropriate promoter/RBS pairs for each inducible promoter, the 6 RBSs spanning a large range of translation efficiencies, i.e. rbs30-34 plus rbsH (Table 4.1), were used to characterise the three promoters under the selected standard condition. It can be seen that (Figure 4.8a–c), the maximum output dependent upon each promoter varied widely using different versions of the RBS. This provides valuable information for choosing matched promoter/RBS pairs as the driving inputs for the designed AND gate. Strikingly, it shows that the order of the strengths of the 6 RBSs across the three promoters varies. This is largely due to the different 5' UTR following each promoter as shown in Figure 4.9, which can vary the secondary structure and the stability of the gene transcript. The observation suggests that the same part (like RBS) used in different sequence contexts might lead to varying behaviour, as also shown by another study (Salis et al., 2009). Figure 4.10 shows that the responses of each promoter using the 6 RBSs become similar after being normalised, which establishes that the RBS can be used like a linear amplifier to adjust the expression level of proteins in cells at steady state.

![Figure 4.9 The core promoter regions and 5' UTR (five prime untranslated region) sequences of the three characterised promoters.](image)

The 5' UTR shown starting from +1 site is the sequence between the core promoter region and the RBS used for the characterisation.
Figure 4.10 Normalised dose responses of the three characterised promoters with 6 RBSs of various strengths: the IPTG-responsive $P_{\text{lac}}$ promoter (a), the arabinose-responsive $P_{\text{BAD}}$ promoter (b) and the AHL-responsive $P_{\text{lux}}$ promoter (c). All curves have very similar Hill coefficients apart from the $P_{\text{lac}}$-rbs33-gfp construct due to no response.

The data were fitted to the Hill function model for the promoter steady state response in the form $f([I_i]) = k_i' \left( \alpha_i + [I_i]^n \right) \left( K_i^n + [I_i]^n \right)$ as derived in Chapter 3, where $[I_i]$ is the concentration of the inducer, $K_i$ and $n_i$ are the Hill constant and coefficient respectively relating to the promoter-regulator/inducer interaction, $k_i'$ is the maximum expression level due to induction and $\alpha_i$ is a constant relating to the basal level of the promoter due to leakage. The nonlinear least square curve fitting function in Matlab was used for the data fitting here and the obtained best fit coefficients are listed in Table 4.2 as well as the plotted fitted curves shown in Figure 4.8. It can be seen that the Hill constant ($K_i$) and coefficient ($n_i$) tend to be close to each other for the same promoter characterised using different RBSs although the maximum response level ($k_i'$) varies.
Table 4.2 The model fits for promoter characterisation using various RBSs in the standard condition with 95% confidence bounds otherwise fixed at bound. Note the last three were the model fits for the characterisation at 37 °C. $R^2$ represents the goodness of the fitting.

<table>
<thead>
<tr>
<th>Promoter/RBS</th>
<th>$k_1$ (au)</th>
<th>$\alpha$</th>
<th>$n_1$</th>
<th>$K_I$ (mM)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plac/rbs30</td>
<td>9456 ± 487</td>
<td>0.0012 ± 0.0276</td>
<td>1.37 ±0.27</td>
<td>0.228 ± 0.039</td>
<td>0.9983</td>
</tr>
<tr>
<td>Plac/rbs31</td>
<td>525.8 ± 36.2</td>
<td>0.1335 ± 0.0411</td>
<td>1.363 ± 0.366</td>
<td>0.294 ± 0.065</td>
<td>0.9972</td>
</tr>
<tr>
<td>Plac/rbs32</td>
<td>195.5 ± 17.7</td>
<td>0.1487 ± 0.0609</td>
<td>1.637 ± 0.62</td>
<td>0.218 ± 0.065</td>
<td>0.994</td>
</tr>
<tr>
<td>Plac/rbs33</td>
<td>27.61 ± 1.2e9</td>
<td>0.975 ± 6.5e7</td>
<td>1.933e-10</td>
<td>0.100</td>
<td>-2.4e-10</td>
</tr>
<tr>
<td>Plac/rbs34</td>
<td>7648 ± 152</td>
<td>1.472e-9</td>
<td>1.369 ± 0.124</td>
<td>0.259 ± 0.021</td>
<td>0.9991</td>
</tr>
<tr>
<td>Plac/rbsH</td>
<td>2071 ± 100</td>
<td>0.0037 ± 0.0241</td>
<td>1.282 ± 0.231</td>
<td>0.287 ± 0.045</td>
<td>0.9987</td>
</tr>
<tr>
<td>PBAD/rbs30</td>
<td>1.048e5 ± 0.026e5</td>
<td>0.0055 ± 0.0104</td>
<td>1.228 ± 0.104</td>
<td>0.370 ± 0.028</td>
<td>0.9997</td>
</tr>
<tr>
<td>PBAD/rbs31</td>
<td>8.521e4 ± 0.434e4</td>
<td>0.0051 ± 0.0229</td>
<td>1.512 ± 0.314</td>
<td>0.417 ± 0.062</td>
<td>0.9986</td>
</tr>
<tr>
<td>PBAD/rbs32</td>
<td>5.208e4 ± 0.098e4</td>
<td>0.0026 ± 0.0072</td>
<td>1.268 ± 0.078</td>
<td>0.516 ± 0.029</td>
<td>0.9999</td>
</tr>
<tr>
<td>PBAD/rbs33</td>
<td>1.29e4 ± 0.046e4</td>
<td>0.0013 ± 0.014</td>
<td>1.323 ± 0.161</td>
<td>0.513 ± 0.055</td>
<td>0.9994</td>
</tr>
<tr>
<td>PBAD/rbs34</td>
<td>1.411e5 ± 0.063e5</td>
<td>0.0050 ± 0.0163</td>
<td>1.173 ± 0.162</td>
<td>0.524 ± 0.072</td>
<td>0.9992</td>
</tr>
<tr>
<td>PBAD/rbsH</td>
<td>9.229e4 ± 0.189e4</td>
<td>0.0033 ± 0.0086</td>
<td>1.415 ± 0.107</td>
<td>0.480 ± 0.029</td>
<td>0.9998</td>
</tr>
<tr>
<td>Plux/rbs30</td>
<td>1.221e5 ± 0.084e5</td>
<td>0.0095 ± 0.032</td>
<td>1.584 ± 0.422</td>
<td>3.073e-6</td>
<td>0.9955</td>
</tr>
<tr>
<td>Plux/rbs31</td>
<td>7.693e4 ± 0.21e4</td>
<td>0.0126 ± 0.013</td>
<td>1.771 ± 0.193</td>
<td>2.955e-6</td>
<td>0.9993</td>
</tr>
<tr>
<td>Plux/rbs32</td>
<td>5.143e4 ± 0.161e4</td>
<td>0.0113 ± 0.0143</td>
<td>1.655 ± 0.202</td>
<td>3.509e-6</td>
<td>0.9991</td>
</tr>
<tr>
<td>Plux/rbs33</td>
<td>1970 ± 61</td>
<td>0.0308 ± 0.0148</td>
<td>1.742 ± 0.212</td>
<td>3.366e-6</td>
<td>0.9991</td>
</tr>
<tr>
<td>Plux/rbs34</td>
<td>1.349e5 ± 0.103e5</td>
<td>0.0149 ± 0.0373</td>
<td>1.898 ± 0.059</td>
<td>2.890e-6</td>
<td>0.9944</td>
</tr>
<tr>
<td>Plux/rbsH</td>
<td>3.194e4 ± 0.137e4</td>
<td>0.0114 ± 0.0192</td>
<td>1.616 ± 0.269</td>
<td>3.784e-6</td>
<td>0.9984</td>
</tr>
<tr>
<td>Plux/rbsH at 37 °C</td>
<td>2946 ± 32</td>
<td>0.0160 ± 0.019</td>
<td>1.697 ± 0.271</td>
<td>0.116 ± 0.011</td>
<td>0.9993</td>
</tr>
<tr>
<td>PBAD/rbs33 at 37 °C</td>
<td>1.203e4 ± 0.017e4</td>
<td>4.69e-10</td>
<td>1.396 ± 0.078</td>
<td>0.525 ± 0.026</td>
<td>0.9998</td>
</tr>
<tr>
<td>Plux/rbs33 at 37 °C</td>
<td>1987 ± 142</td>
<td>2.2e-7 ± 0.0381</td>
<td>1.776 ± 0.682</td>
<td>5.27e-6 ± 2.9e-6</td>
<td>0.9964</td>
</tr>
</tbody>
</table>
Figure 4.11 The characterised responses of $P_{\text{lac}}$ (a), $P_{\text{BAD}}$ (b) and $P_{\text{lux}}$ (c) under two temperatures $30 \, ^\circ\text{C}$ and $37 \, ^\circ\text{C}$ respectively, and the model fits. The inducer concentrations used for the inductions are the same as in Figure 4.8. All data (fluorescence/OD$_{600}$) were the average of three independent repeats in *E. coli* MC1061 in M9-glycerol. Error bars, s.d. (n = 3).
Finally, to test the effect of temperature upon each promoter’s behaviour, the three inducible promoters were characterised at two temperatures 30 °C and 37 °C under the otherwise standard culturing condition. Here rbsH was chosen for P_{lac}, and rbs33 for P_{BAD} and P_{lux}. Referring to Figure 4.11a–c, it can be seen that the effect of temperature shift on the three promoter responses are different. In the case of P_{BAD}, a change in temperature from 30 °C to 37 °C has only a small effect. But P_{lux} becomes leakier at 30 °C than at 37 °C, and P_{lac} has a higher response at 37 °C than at 30 °C as reflected by the results of data fitting to the Hill function model (Table 4.2). One reason behind these different variations might be due to the different effect of temperature on the binding affinities between the transcription factor proteins and their cognate DNA binding sites of these three inducible promoters. The result shown here suggests that a change of the physical context can have varying impacts on different parts in the system.

4.3 Engineering the Functional AND Gates

4.3.1 Construction and characterisation of the AND gate

Based on the results of part characterisation (i.e. the promoters and RBSs), it proceeded to apply them to the systematic design process of the AND gate. In the initial trial version of the AND gate (Figure 4.1 and Figure 4.3), the input of P_{BAD}/rbsH pair, maximum response level around 90000 A.U. (Figure 4.8b), is much stronger than the input of P_{lac}/rbsH pair, maximum response level around 2000 A.U. (Figure 4.8a), in the selected context. Thus the distorted asymmetric response of this AND gate (Figure 4.3) is largely due to the unmatched two inputs. In order to balance the two inputs, the P_{BAD}/rbsH pair was replaced with the P_{BAD}/rbs33 pair as rbs33 is a much weaker RBS compared to rbsH and can tune the input down to a much lower level as shown in Figure 4.8b. Figure 4.12 are the plasmids maps showing the constructs built for the circuit. This engineered version of the AND gate produced the
desired classic AND gate response in the standard condition as the characterisation data shown in Figure 4.13, where the output is turned on only when both inputs are highly induced. It can be seen that the output response is very sharp in the transition from the “off” (low fluorescence) to “on” (high fluorescence) state and is close to the one of a digital logic AND gate. The engineered AND gate has many merits of an ideal biological AND logic gate, e.g. rapid output state switching across a narrow transition region of the inputs, a relatively large dynamic range and almost zero level of the output at “off” state. The device behaviour was also subject to FACS assay as described in Chapter 2 and the results are shown in Figure 4.13c, where the entire population of cells is turned on in the presence of both inducers (1.6 Mm IPTG and 1.3 mM arabinose respectively) and the entire population is off when either inducer is not added. Comparing Figure 4.13b to Figure 4.3, it can be seen that the forward engineered AND gate based on quantitatively characterised parts can lead to more predictable system behaviour than the rationally designed AND gate.

Two-input AND gate using $P_{\text{lac}}$ and $P_{\text{BAD}}$ as the inputs

![Plasmid maps showing the circuit constructs used for the characterisation of the AND gate](image)

Figure 4.12 Plasmid maps showing the circuit constructs used for the characterisation of the AND gate as the results shown in Figure 4.13.
Figure 4.13 The systematic characterisation of the modular AND gate. **a.** The schematic of the engineered AND gate. **b.** The fluorescent response of the AND gate for 72 combinations of input inductions measured in the fluorometer. The data are shown for the AND gate induction by (left to right) 0, 3.9 × 10^{-4}, 1.6 × 10^{-3}, 6.3 × 10^{-3}, 2.5 × 10^{-2}, 0.1, 0.4 and 1.6 mM IPTG, and induction by (bottom to top) 0, 3.3 × 10^{-4}, 1.3 × 10^{-3}, 5.2 × 10^{-3}, 2.1 × 10^{-2}, 8.3 × 10^{-2}, 0.33, 1.3 and 5.3 mM arabinose. The data were the average of three independent repeats in *E. coli* MC1061 in M9-glycerol at 30 °C with variations less than 10%. **c.** The flow cytometry assay of the AND gate to determine the fluorescent responses at individual cells for the four logic combinations of input inductions.
4.3.2 Parameterisation of the AND gate transfer function

The data of this functional device was used to parameterise the normalised transfer function model of the AND gate as derived in Chapter 3:

\[ f([R]_{arb}+[S]_{arb}) = \frac{[G]_{arb}}{[G]_{max}} = \frac{([R]_{arb}/K_R)^{n_R}([S]_{arb}/K_S)^{n_S}}{1+([R]_{arb}/K_R)^{n_R}} \]

where \([G]_{max}\) is the maximum activity observed for the output; \(K_R\), \(K_S\) and \(n_R\), \(n_S\) are the Hill constants and coefficients for HrpR and HrpS, respectively. The steady state levels of the activators were derived from the characterised responses of the two input promoters with the same RBSs as in the characterised AND gate. The parameterisation was implemented by fitting to the characterised response of the engineered AND gate shown in Figure 4.13b together with the fitted transfer functions of the one-dimensional input promoters (Table 4.2). The best fit coefficients with 95% confidence bounds by nonlinear least square optimisation were obtained as shown on the right of Figure 4.14 and the parametrised transfer function was plotted on the left. In order to be easily compared with the experimental data, the fitted transfer function models are plotted with the same concentrations of inducers as used for the characterisation.

\[ K_R = 206.1 \pm 32.5 \]
\[ K_S = 3135 \pm 374 \]
\[ n_R = 2.381 \pm 0.475 \]
\[ n_S = 1.835 \pm 0.286 \]
\[ ([G]_{AND_{max}} = 7858 \text{ au}, R^2 = 0.9781) \]

Figure 4.14 Fitted transfer function of the AND gate. The transfer function model of the engineered AND gate was parameterised by fitting to the experimental data as shown in Figure 4.13b.
4.3.3 The modularity of the AND gate circuit

To verify the modularity of the AND gate, the Plac input was next swapped to the AHL-responsive Plux input. As Plux has been characterised as a strong promoter (Figure 4.8c), the input of Plux/rbs33 pair (maximum response 2000 A.U.) was used to drive the AND gate, to get close to the input of the Plac/rbsH pair. This version of AND gate was first characterised under the standard condition and produced a response as shown on the bottom left of Figure 4.15. It shows that it is similar to logic AND response but with a leaky response on the side of the Plux input. This might not be surprising as there is greater leakage for Plux at low levels of AHL at 30 °C than at 37 °C (Figure 4.11c). By shifting the temperature of characterisation to 37 °C, the same device then produced an improved AND gate characteristic as shown on the bottom right of Figure 4.15. These data show that the context, in which the circuit behaves, has an impact on its behaviour and the construction can be facilitated using the parts that have been characterised in various contexts. In addition, the right performance of the AND gate with this new promoter input verified the input modularity of the device. The observations of this AND gate in the two conditions were consistent with the predictions of the fitted AND gate model as shown in Figure 4.16a, b.
Figure 4.15 The engineering and systematic characterisation of the modular AND gate with alternative inputs. The top is the schematic for the AND gate engineered using $P_{\text{lux}}$ and $P_{\text{BAD}}$ as inputs and the weakest RBS (rbs33) for the interfacing of both. The device was characterised at 30 °C and 37 °C respectively as shown on the bottom left and right. The data shown are for the induction by (left to right) $0$, $2.4 \times 10^{-2}$, $9.8 \times 10^{-2}$, $3.9 \times 10^{-1}$, $1.6$, $6.3$, $25$ and $100$ nM AHL, and by (bottom to top) $0$, $3.3 \times 10^{-4}$, $1.3 \times 10^{-3}$, $5.2 \times 10^{-3}$, $2.1 \times 10^{-2}$, $8.3 \times 10^{-2}$, $0.33$ and $1.3$ mM arabinose. All data are the normalised average of three repeats in *E. coli* MC1061 in M9-glycerol with variations less than 10%.
4.3.4 Fitted model predictions of the AND gate behaviour

To test the prediction capability of the parametrised transfer function, the AND gate behaviour in the two new configurations and conditions as shown in Figure 4.15 were compared with the model predictions by coupling the fitted transfer functions of the characterised input promoters (Table 4.2) and the AND gate (Figure 4.14). The fitted model predictions of the AND gate in these two new configurations and conditions are shown Figure 4.16a, b. It can be seen that the predictions capture well the major experimental characteristics of the engineered AND gate as shown in Figure 4.15, particularly the leakiness of $P_{lux}$ input at the lower temperature 30 °C. In order to be easily compared with the experimental data, the model predictions are plotted with the same concentrations of inducers as used for the experimental characterisations.

![Figure 4.16 Predictions of the AND gate behaviour using fitted transfer functions. a, The predicted characteristics for the AND gate in a new configuration using $P_{lux}$ and $P_{BAD}$ as the two inputs. The prediction is based on the AND gate model by coupling the fitted transfer function of the AND gate in Section 4.3.2 and the fitted transfer functions of promoter $P_{lux}$ and $P_{BAD}$ with the corresponding RBSs at 30 °C in Table 4.2. The experimental validation for this prediction is shown on the bottom left of Figure 4.15. b, The predicted behaviour for the AND gate as in a but in the condition of 37 °C. The prediction is based on the AND gate model by coupling the fitted transfer function of the AND gate in Section 4.3.2 and the fitted transfer functions of characterised input promoters with the corresponding RBSs at 37 °C in Table 4.2. The experimental validation for this prediction is shown on the bottom right of Figure 4.15.](image-url)
4.4 Homogeneity, Metabolic Load and Chassis Compatibility
Assays of the Circuits

4.4.1 Cell homogeneity assay of parts and devices

In this section, cells harbouring the environmental-responsive promoters and the engineered circuits were subjected to analysis by flow cytometry (i.e. the fluorescence of individual cells measured) to examine the GFP expression homogeneity from the parts and modules, and to verify that the circuits function at the single cell level as well as at the ensemble level.

The *E. coli* MC1061 cells harbouring the plasmid construct containing the IPTG-inducible promoter $P_{lac}$ (plasmid pBW105lac-gfp), or the arabinose-inducible $P_{BAD}$ (plasmid pBW203ara-gfp), or the AHL-inducible $P_{lux}$ (plasmid pBW303lux-gfp) were assayed by flow cytometry, with the results shown in Figure 4.17. Referring to Figure 4.17, the cells harbouring promoter $P_{BAD}$ is non-homogenous at intermediate induction levels, i.e. a bimodal distribution of the cell population, while the cells harbouring $P_{lac}$ and $P_{lux}$ promoters have unimodal responses at all graded induction levels, i.e. homogenous. This behaviour is consistent with the findings of a previous study (Siegele and Hu, 1997), which also shows that gene expression from plasmids containing the $P_{BAD}$ promoter represents mixed population at subsaturating inducer concentrations in *E. coli* MC1061. The non-homogeneity of the $P_{BAD}$ promoter in *E. coli* MC1061 is used to illustrate the similar behaviour of the cells containing the engineered AND gate as observed below.

The cells harbouring the functional AND gate circuit were subject to analysis by flow cytometry and the results are shown in Figure 4.18 and Figure 4.19. Referring to Figure 4.18, the AND gate is confirmed to behave with the logic AND function at individual cells, i.e. the output is high when both of the two inputs are fully activated. But the AND gate behaved with bimodal response at intermediate induction levels of $P_{BAD}$ as shown in Figure 4.18b, c. This is because the $P_{BAD}$ promoter response is non-homogeneous in this cell host (*E. coli* MC1061), while $P_{lac}$ promoter is homogeneous.
as shown in Figure 4.18a where the responses are unimodal at all IPTG induction levels under full arabinose induction. The similar behaviour was observed in the assay of another version of the AND gate using \( P_{\text{lux}} \) and \( P_{\text{BAD}} \) as the two inputs as shown in Figure 4.19. Combined with the assays of the promoter responses shown in

Figure 4.17 FACS assays of the promoter \( P_{\text{lac}} \) (a), \( P_{\text{BAD}} \) (b) and \( P_{\text{lux}} \) (c) in \( E. \text{coli} \) MC1061 after 5 hours growth in M9-glycerol at 37 °C. a, The responses of cells harbouring \( P_{\text{lac}} \)-rbsH-gfp construct induced by (bottom to top) 0, 3.9 \( \times \) 10\(^{-4}\), 1.6 \( \times \) 10\(^{-3}\), 6.3 \( \times \) 10\(^{-3}\), 2.5 \( \times \) 10\(^{-2}\), 0.1, 0.4, 1.6, 6.4 and 12.8 mM IPTG. b, Cellular response of \( P_{\text{BAD}} \)-rbs33-gfp induced by (bottom to top) 0, 3.3 \( \times \) 10\(^{-4}\), 1.3 \( \times \) 10\(^{-3}\), 5.2 \( \times \) 10\(^{-3}\), 2.1 \( \times \) 10\(^{-2}\), 8.3 \( \times \) 10\(^{-2}\), 0.33, 1.3, 5.3 and 10.7 mM arabinose. c, Cellular response of \( P_{\text{lux}} \)-rbs33-gfp induced by (bottom to top) 0, 6.1 \( \times \) 10\(^{-3}\), 2.4 \( \times \) 10\(^{-2}\), 9.8 \( \times \) 10\(^{-2}\), 3.9 \( \times \) 10\(^{-1}\), 1.6, 6.3, 25, 100 and 400 nM AHL.
Figure 4.17, it can be concluded that the AND gate itself is homogenous but the homogeneity of the whole functional circuit largely relies on the environment-responsive promoters used as the inputs.

![Image of Figure 4.17](image)

**Figure 4.18** FACS assays of the engineered AND gate using $P_{\text{lac}}$ and $P_{\text{BAD}}$ as the two inputs in *E. coli* MC1061 after 5 hours growth in M9-glycerol at 30 °C. 

a, FACS assay of the AND gate with full induction of the $P_{\text{BAD}}$ input (1.33 mM arabinose) and graded induction of the $P_{\text{lac}}$ input by (bottom to top) 0, $3.9 \times 10^{-4}$, $1.6 \times 10^{-3}$, $6.3 \times 10^{-3}$, $2.5 \times 10^{-2}$, 0.1, 0.4 and 1.6 mM IPTG. 

b, FACS assay of the AND gate with full induction of the $P_{\text{lac}}$ input (1.6 mM IPTG) and graded induction of the $P_{\text{BAD}}$ input by (bottom to top) 0, $3.3 \times 10^{-4}$, $1.3 \times 10^{-3}$, $5.2 \times 10^{-3}$, $2.1 \times 10^{-2}$, $8.3 \times 10^{-2}$, 0.33 and 1.33 mM arabinose. 

c, FACS assay of the AND gate with graded inductions for both of the inputs $P_{\text{lac}}$ and $P_{\text{BAD}}$. 
Figure 4.19 FACS assays of the engineered AND gate using $P_{lux}$ and $P_{BAD}$ as the two inputs in *E. coli* MC1061 after 7 hours growth in M9-glycerol at 37 °C. a, The AND gate assay with induction by 1.33 mM arabinose and graded induction by (bottom to top) 0, $2.4 \times 10^{-2}$, $9.8 \times 10^{-2}$, $3.9 \times 10^{-1}$, 1.6, 6.3, 25 and 100 nM AHL. b, The AND gate assay with induction by 100 nM AHL and graded induction by (bottom to top) 0, $3.3 \times 10^{-4}$, $1.3 \times 10^{-3}$, $5.2 \times 10^{-3}$, $2.1 \times 10^{-2}$, $8.3 \times 10^{-2}$, 0.33 and 1.33 mM arabinose. c, The AND gate assay with graded inductions for both of the inputs $P_{lux}$ and $P_{BAD}$.

### 4.4.2 Metabolic load assay of circuit constructs

Just like characterising the power consumption of an electrical device, it is important to evaluate the metabolic load of an engineered biological device. An ideal biological device is expected to place negligible load on the host metabolism. Otherwise the host fitness will be severely affected by a device having heavy metabolic load. This section considers how the engineered AND gate circuit affects the fitness of the host in which it performs. The experiments were implemented to examine the growth curves of cells containing various circuit constructs to evaluate the metabolic load of the circuit on the host. For the purpose of these experiments, one control (*E. coli* MC1061, wild type) was used; one reference carrying the empty vectors without the circuit constructs, and one carrying the three plasmids with the functional AND gate circuit using $P_{lac}$ and $P_{BAD}$ as the two inputs. The growth curves of the control, the reference, and the cells harbouring the AND gate at the induced (1.33 mM arabinose plus 1.6 mM IPTG) and non-induced (no inductions) conditions are shown in Figure 4.20. The cells diluted from overnight cultures with OD$_{600}$ = 0.05 were grown in a 96 well microplate in the fluorometer with shaking (200 rpm, linear mode) for 20 hours. The
absorbance (OD$_{600}$) of the cell cultures was read every 1h. The data were the average of three repeats from the three bacterial colonies of each strain.

Figure 4.20 Growth curves of *E. coli* MC1061 harbouring various circuit constructs in M9-glycerol at 30 °C. Error bars denote the standard deviations of three repeats.

Referring to Figure 4.20, over a 20 hour period, the greatest growth was achieved by *E. coli* MC1061. The three growth characteristics with the plasmids inserted (ref. - empty vectors, AND - noninduced and AND - induced) have growth characteristics which are very similar. Two key points can be made regarding these results: (i) for all four conditions the growth characteristics show that the cells are viable and that whilst the insertion of the plasmids does affect growth rate, the reduction in growth rate is not catastrophic; (ii) referring again to the figure, the characteristics for the three conditions where the plasmids have been inserted are very similar. This shows that the loss in growth rate is due to the impact of the plasmid carriers in the cell instead of the function of the AND gate circuit. Thus, our engineered AND gate actually placed little metabolic burden on the host. However, the plasmids used for carrying the circuit constructs can reduce the rate of cell growth, which is likely due in a large to the costs of establishing antibiotic resistance for maintaining the plasmids inside the cells.
4.4.3 Chassis compatibility assay of engineered circuits

This section examines the functionality of the AND gate in various *E. coli* cell chassis to evaluate the chassis compatibility of the engineered circuits. Here, the response of the engineered AND gate circuit was tested in seven normally used *E. coli* strains using three inducible promoters, i.e. $P_{lac}$, $P_{BAD}$ and $P_{lux}$, as the driving inputs. The results in the seven cell hosts carrying the AND gate circuit are shown in Figure 4.21 and Figure 4.22.

Referring to Figure 4.21, four input conditions were studied – arabinose plus IPTG; arabinose only; IPTG only; and no inputs, for the seven *E. coli* strains harbouring the AND gate using $P_{lac}$ and $P_{BAD}$ as the two inputs. Cells were grown in M9-glycerol at 30 °C in a 96 well microplate and assayed after 5 hours upon induction. The data shown are the averages of three independent repeats with variations of less than 10% between biological replicates. It can be seen that the device does not work properly in five out of the seven tested cell chassis, i.e. the *E. coli* MC4100, MG1655, DH5α, BW25113 and BL21-DE3. As for most of the non-working chassis, the cells have high output not only with both input inductions (1.3 mM arabinose plus 1.6 mM IPTG) but also with only the induction of the $P_{BAD}$ input (1.3 mM arabinose). This is likely due to the interference between the host genetic backgrounds with the $P_{lac}$ promoter input of the device. The device works well in *E. coli* MC1061 as well as in *E. coli* Top 10, a derivative of *E. coli* MC1061. The levels of the device response upon both input inductions across the seven chassis are also different, which is likely due to the combined effect of the potential interference between $P_{lac}$ input and the host, and the various growth reductions observed for the seven hosts harbouring the same circuit as shown in Table 4.3.
Figure 4.21 Qualitative assays of the functionality of the engineered AND gate using $P_{\text{lac}}$ and $P_{\text{BAD}}$ as the two inputs in seven $E.\ coli$ strains.

Table 4.3 Summary of chassis compatibility assays of the AND gate in Figure 4.21.

<table>
<thead>
<tr>
<th>Chassis</th>
<th>Function</th>
<th>Growth reduction</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E.\ coli$ MC1061</td>
<td>Good</td>
<td>Minor</td>
<td>output only with two input inductions</td>
</tr>
<tr>
<td>$E.\ coli$ MC4100</td>
<td>Poor</td>
<td>Heavy</td>
<td>output with two input inductions - but also with only $P_{\text{BAD}}$ induction</td>
</tr>
<tr>
<td>$E.\ coli$ MG1655</td>
<td>Poor</td>
<td>middle</td>
<td>output with two input inductions - but also with only $P_{\text{BAD}}$ induction</td>
</tr>
<tr>
<td>$E.\ coli$ Top 10</td>
<td>Good</td>
<td>Minor</td>
<td>output only with two input inductions</td>
</tr>
<tr>
<td>$E.\ coli$ DH5a</td>
<td>Poor</td>
<td>Minor</td>
<td>output with two input inductions - but also with only $P_{\text{BAD}}$ induction</td>
</tr>
<tr>
<td>$E.\ coli$ BW25113</td>
<td>Poor</td>
<td>No</td>
<td>output with two input inductions - but also with only $P_{\text{BAD}}$ induction</td>
</tr>
<tr>
<td>$E.\ coli$ BL21-DE3</td>
<td>Poor</td>
<td>Minor</td>
<td>no response for any combination of inputs</td>
</tr>
</tbody>
</table>

Referring to Figure 4.22, four input conditions were studied – arabinose plus AHL; arabinose only; AHL only; and no inputs, for the seven $E.\ coli$ strains harbouring the AND gate using $P_{\text{lux}}$ and $P_{\text{BAD}}$ as the two inputs. Cells were grown in M9-glycerol at 37 °C in a 96 well microplate and assayed after 5 hours upon induction. The data shown are the averages of three independent repeats with
variations less than 10% between biological replicates. It can be seen that the device works well in six out of the seven chassis except in \textit{E. coli} BL21-DE3. The improvement of chassis compatibility is due to the elimination of the potential interference of the input promoter with the host genetic background by using \textit{P\textsubscript{lux}} instead of \textit{P\textsubscript{lac}} for the first driving input of the device now. The \textit{P\textsubscript{lux}} is not endogenous in \textit{E. coli} and is likely orthogonal to the genetic background of this bacteria, while the \textit{P\textsubscript{lac}} is an endogenous \textit{E. coli} promoter. The device has fluorescent outputs in \textit{E. coli} BL21-DE3 for all four input conditions tested. This may be the result of the absence of the Lon protease in this cell strain, which has been shown to degrade the HrpR protein more prominently in \textit{Pseudomonas} bacteria (Bretz et al., 2002). Thus, the background level of HrpR activators may tend to be higher in this cell chassis than in others, which leads to the leakiness on the \textit{P\textsubscript{lux}} input side as observed in Figure 4.22. The variations of the output levels across the seven chassis are largely due to the different growth reductions of the hosts harbouring the same circuit as observed in the experiment (see Table 4.4).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.22.png}
\caption{Figure 4.22 Qualitative assays of the functionality of the engineered AND gate using \textit{P\textsubscript{lux}} and \textit{P\textsubscript{BAD}} as the two inputs in seven \textit{E. coli} strains.}
\end{figure}
Table 4.4 Summary of chassis compatibility assays of the AND gate in Figure 4.22.

<table>
<thead>
<tr>
<th>Chassis</th>
<th>Function</th>
<th>Growth reduction</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> MC1061</td>
<td>Good</td>
<td>Minor</td>
<td>output only with two input inductions</td>
</tr>
<tr>
<td><em>E. coli</em> MC4100</td>
<td>Good</td>
<td>Heavy</td>
<td>output only with two input inductions</td>
</tr>
<tr>
<td><em>E. coli</em> MG1655</td>
<td>Good</td>
<td>Heavy</td>
<td>output only with two input inductions</td>
</tr>
<tr>
<td><em>E. coli</em> Top 10</td>
<td>Good</td>
<td>No</td>
<td>output only with two input inductions</td>
</tr>
<tr>
<td><em>E. coli</em> DH5α</td>
<td>Good</td>
<td>Minor</td>
<td>output only with two input inductions</td>
</tr>
<tr>
<td><em>E. coli</em> BW25113</td>
<td>Good</td>
<td>No</td>
<td>output only with two input inductions</td>
</tr>
<tr>
<td><em>E. coli</em> BL21-DE3</td>
<td>Poor</td>
<td>Minor</td>
<td>output with all four input conditions</td>
</tr>
</tbody>
</table>

By combining the two studies above, it can be concluded that the chassis compatibility of the device is largely related to the particular two driving inputs employed (i.e. the inducible promoters) while the AND gate circuit itself is broadly compatible and has reliable response across many cell chassis tested. This is due to the fact that some of the input promoters used are endogenous in *E. coli* and may interact non-ideally with the host genetic background such as the Plac promoter. Thus the promoter inputs of the AND gate device need to be carefully considered to avoid the potential interference with the host genetic background. Ideally, all the parts used in a device should be orthogonal to the cell host to eliminate or minimise the potential interference between the engineered circuit and the chassis background.
4.5 Discussion

In this chapter, a set of basic genetic elements (i.e. the promoters and RBSs) have been quantitatively characterised in various biophysical and genetic contexts for the systematic forward engineering of synthetic circuits with the modular AND gate function. The initial implementation of the designed genetic AND gate provides us with an evidence to show that rationally designed biological systems may rarely work as might be initially intended. This is partly because the individual parts and modules are not sufficiently well characterised for purpose, and so the effective coupling together of them is not straightforward due to their different characteristics. Moreover, the behaviour of component modules characterised in one context may vary quite differently in another working condition (Klumpp et al., 2009; Tan et al., 2009) and configuration (Salis et al., 2009). By quantitatively characterising a set of promoters, a series of RBSs and the assembled circuits in various contexts, both the biophysical and genetic contexts, in which they behave, were found to have a great impact on their functionalities. Experimental results in this chapter show that the variations of the behaviour of genetic parts and devices in different contexts are due to the many factors affecting gene expression, such as the cell chassis background, medium, temperature, the embedded sequence context. Thus the thorough characterisation of them in various conditions and configurations, particularly in the context of interest, are necessary for any successful wide reuse in the community and more importantly for facilitating the functional assembly of individual parts and modules into customisable large scale systems. While the creation of diverse part and module variants are becoming simplified and at reduced cost (Dougherty and Arnold, 2009; Ellis et al., 2009), high throughput and accurate technology platforms need to be established to accelerate the process of characterisation, like the recently established BIOFAB (biofab, 2010) open technology platform.

The modular genetic AND gate engineered here works as a fundamental module for regulating genetic information transmission in living cells, which can integrate two input signals to generate one output in the digital logic AND manner. Moreover, the core elements of this device are from the specialised *hrp* gene regulatory system.
for Type III secretion in *Pseudomonas syringae*, a plant pathogenic bacterium genetically far away from the standard *E. coli* chassis. Thus, the novel genetic AND gate module is likely orthogonal to the *E. coli* genetic background and can be used simultaneously with existing gene regulatory elements that are widely utilised for regulating gene expression in *E. coli* without any compromise. The orthogonality of the AND gate has been demonstrated by the experimental study in this chapter. The study also provides an example to exploit the diverse natural biological modules and to engineer them for creating novel orthogonal parts and modules to expand the limited toolbox of synthetic biology at the current stage. With the increase of the number orthogonal parts and modules in the toolbox, we can engineer more complicated systems which can contain many parts to enable the cells with high level functions.

From the fitted transfer function of the AND gate (Figure 4.15), it was noticed that the fit Hill constants and coefficients for the two activators HrpR and HrpS are quite different. This indicates that the two enhancer binding proteins likely have unequal roles in the binding and activation of the $\sigma^{54}$-dependent *hrpL* promoter although they are originally expressed from the same operon and have many sequence homologies. On this aspect, the work corroborates that the study of simple synthetic biological circuits can contribute to the uncovering of the design principles of their natural counterparts (Mukherji and van Oudenaarden, 2009). In addition, the characterisation scope for synthetic parts and devices was extended. The AND gate circuit was subjected to homogeneity, metabolic load and chassis compatibility assays respectively beyond the normal population-level phenotype assays. The thorough characterisation of parts and devices will enable their wide reuse in the community and facilitate the reliable prediction of their behaviour when integrated into other larger systems.

The work in this chapter also shows that the functional assembly of component parts into devices and systems can be executed more predictably and reliably using engineered, in-context quantitatively characterised parts and sub-modules aided by modelling. This approach minimises the unexpected or high-order effects which could
occur during circuit construction by characterising components in the same biophysical and genetic context as anticipated in their target system. This was illustrated by converting an assembled non-functional logic AND circuit to be a functional one. The modular AND gate can be reconnected to different sensor inputs to detect and integrate various environmental signals (Anderson et al., 2007; Kobayashi et al., 2004; Voigt, 2006) or easily incorporated into a large system as a fundamental building block to regulate the cell signalling in a desired logic manner. The behaviour of the constructed modules are well captured by the parameterised mathematical models for their transfer functions, which are reusable for modelling the behaviour of larger integrated modules (Guido et al., 2006). The functional assembly approach described here provides an effective strategy and a guide for the engineering of synthetic gene circuits with predictable functions across different contexts for their application in many areas such as biotechnology, biocomputing and biosensors.
Chapter 5

Engineering a Set of Modular NOT and NAND Gates
In the previous chapter, a set of parts and sub-modules have been quantitatively characterised in various contexts and it has shown that the functional assembly of BioParts into a modular genetic AND gate can be facilitated using these ‘in-context’ characterised components. In this chapter, to further verify this functional assembly approach at the device level, it proceeded to engineer a set of modular logic NOT gates and the combinatorial logic NAND gates, which comprise component modules of the AND gate and NOT gate.

5.1 Engineering a Set of Modular NOT Gates

5.1.1 The experimental design and construction of the NOT gates

Here, three types of modular logic NOT gate were rationally designed on the basis of the three classic biological repressor modules – lacI/P\textsubscript{Llac}, tetR/P\textsubscript{tet} and cI/P\textsubscript{lam}. As Figure 5.1 shows, the designed NOT gates function in the way that the expressed repressor proteins bind their cognate operator sites on the downstream promoter, which results in the inhibition of the transcription of the corresponding promoter. The design is modular as both the inputs and output of the designed NOT gates are promoters. Thus, various environment-responsive input promoters can be connected to the NOT gate and the output can be used to drive different cellular responses. The activity level of the NOT gate output is in inverse proportion to that of the input promoter.

![Diagram of NOT gates](image)

**Figure 5.1 Design of the modular genetic NOT gates.** a, The lacI/P\textsubscript{Llac} repressor module based NOT gate. The P\textsubscript{Llac} promoter is a hybrid regulatory region consisting of the promoter P\textsubscript{L} of phage lambda with the CI binding sites replaced with lacO1 operator sites to allow strong expression (Lutz and Bujard, 1997). b, The tetR/P\textsubscript{tet} based NOT gate. c, The cI/P\textsubscript{lam} based NOT gate. P\textsubscript{lam} is based on the P\textsubscript{R} promoter of phage lambda.
To construct the three types of NOT gate, the Inverters that have been constructed with BioBrick standard parts in the Part Registry (Registry of Standard Biological Parts, 2010d) were used as templates here instead of cloning their natural elements by PCR from scratch. Each repressor gene of the inverters (lacI/P_{L_{lac}}, tetR/P_{tet} and cI/P_{lam}) from the registry has been modified with a LVA tail. The LVA tail (a peptide tag of AADENYA_{LVA}) is attached to the C-terminal of the target protein. The resulting protein is then recognised and rapidly degraded by intracellular tail-specific proteases in *E. coli* (Andersen et al., 1998; Keiler et al., 1996). Consequently, the LVA tail can lead to rapid degradation of the repressor proteins in the inverters which enables dynamic control of the output gene expression and brings down the repressor background level to allow larger dynamic range of the output. The inverters in the registry are designed only in limited number of RBS versions and are not well characterised. Thus, various RBS sequences for each repressor gene construct were introduced with primers containing the corresponding RBS and appropriate restriction sites, which can modify the behaviour of the NOT gates to different extent and provide more choice for the construction of subsequent composite devices. The constructs of the NOT gates were next cloned into the expression vectors under appropriate inducible promoters as shown in Figure 5.2. The arabinose inducible P_{BAD} promoter in pBAD18-cm vector (Guzman et al., 1995) and IPTG inducible P_{lac}

![Figure 5.2 Plasmid maps showing circuit constructs used for the characterisation of the NOT gates.](image-url)

All maps shown are NOT gates of the rbs34 version. a, The lacI/P_{L_{lac}} NOT gate was cloned under an arabinose inducible P_{BAD} promoter in pBAD18-cm vector. b, The tetR/P_{tet} NOT gate was cloned under the IPTG inducible P_{lac} promoter in pAPT110 vector. b, The cI/P_{lam} NOT gate was cloned under the IPTG inducible P_{lac} promoter in pAPT110 vector.
promoter in pAPT110 vector (Polard and Chandler, 1995) were used here as they have been quantitatively characterised in Chapter 4 and their transfer functions are already known.

5.1.2 The characterisation of the NOT gates

Here the aim is to engineer genetic logic NOT gates that can closely mimic the behaviour of a digital logic NOT gate, e.g. rapid output state switching across a narrow transition region of the input and a relatively large dynamic range of the output. To increase the diversity of the device behaviour and thus the flexibility for

Figure 5.3 The characterisation of the lacI/P_{Lac} based NOT gate. The NOT gate module was characterised under the arabinose inducible P_{BAD} promoter using 3 versions of RBS (rbs31, rbs33, rbs34) of various strengths. The bottom is the dose responses (fluorescence/OD600) of the engineered NOT gate induced by (left to right) 0, 3.3 × 10^{-4}, 1.3 × 10^{-3}, 5.2 × 10^{-3}, 2.1 × 10^{-2}, 8.3 × 10^{-2}, 0.33, 1.3, 5.3 and 10.7 mM arabinose. The data were fitted to the transfer function model of the NOT gate module. Error bars, s.d. (n=3).
selection, a device library containing eleven versions of the NOT gates using various RBSs were finally constructed and characterised under appropriate inducible promoters in the standard culturing condition (i.e. *E. coli* MC1061, M9-glycerol, 30 °C). The characterised cellular responses of the three types of the NOT gates using various induction levels are shown in Figure 5.3, Figure 5.4 and Figure 5.5 respectively.

![Diagram of NOT gate](image)

**Figure 5.4 The characterisation of the tetR/P_{tet} based NOT gate.** The NOT gate module was characterised under the IPTG inducible P_{lac} promoter using 3 versions of RBS (rbs31, rbs33, rbs34). The bottom is the dose responses (fluorescence/OD600) of the NOT gate by (left to right) 0, 3.9 × 10^{-4}, 1.6 × 10^{-3}, 6.3 × 10^{-3}, 2.5 × 10^{-2}, 0.1, 0.4, 1.6, 6.4 and 12.8 mM IPTG. The data were fitted to the transfer function model of the NOT gate module. Error bars, s.d. (n=3).

Referring to Figure 5.3, it can be seen that all the RBS versions of the lacI/P_{Llac} NOT gate are roughly similar and show basic transition characteristics. However, the
transition is over a much wider inducer concentration range and there is significant divergence between the characteristics. Referring to Figure 5.4, the characteristics for two of the three RBS versions (rbs31, rbs34) of the tetR/P tet NOT gate are similar and there is a clear transition from one state to the other. However, the transition is over a wide inducer concentration range and there is divergence between the transition characteristics as well. The non responsive (flat) characteristic of the rbs33 version is probably due to the almost no response of the P lac/rbs33 pair as shown in Chapter 4.

Figure 5.5 The characterisation of the cI/P lam based NOT gate. The NOT gate module was characterised under the IPTG inducible P lac promoter using 5 versions of RBS (rbs31, rbs32, rbs33, rbs34, rbsH). The bottom is the dose responses (fluorescence/OD600) of the NOT gate by (left to right) 0, 3.9 × 10^−4, 1.6 × 10^−3, 6.3 × 10^−3, 2.5 × 10^−2, 0.1, 0.4, 1.6, 6.4 and 12.8 mM IPTG. The data were fitted to the transfer function model of the NOT gate module. Error bars, s.d. (n=3).
Referring to Figure 5.5, it can be seen that the cl/P_lam based NOT gate exhibited the best characteristics overall. Four of the five versions of this NOT gate each produced similar responses with rapid state transition over a narrow range of inducer concentration, a large dynamic range and a low level of output at the “off” state when comparing with the lacI/P_Llac and tetR/P_tet based NOT gates.

Table 5.1 The model fits of the three types of NOT gates using various RBSs in the standard condition with 95% confidence bounds otherwise fixed at bound. cl/P_lam and tetR/P_tet based NOT gates were characterised under the IPTG-inducible P_lac promoter in pAPT110 vector. The lacI/P_Llac based NOT gate was characterised under the arabinose-inducible P_BAD promoter in pBAD18-cm vector.

<table>
<thead>
<tr>
<th>NOT gate</th>
<th>k_3 (au)</th>
<th>a_3</th>
<th>n_3</th>
<th>K_3 (au)</th>
<th>R^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>rbs31-cl/P_lam</td>
<td>7.538e4 ± 0.234e4</td>
<td>0.0527 ± 0.0144</td>
<td>7.647 ± 1.379</td>
<td>111 ± 4.7</td>
<td>0.9997</td>
</tr>
<tr>
<td>rbs32-cl/P_lam</td>
<td>7.191e4 ± 0.385e4</td>
<td>0.0631 ± 0.0141</td>
<td>5.09 ± 0.714</td>
<td>47.47 ± 5.42</td>
<td>0.9997</td>
</tr>
<tr>
<td>rbs33-cl/P_lam</td>
<td>7.5e4 ± 3.55e11</td>
<td>0.5225 ± 5.10e6</td>
<td>1.005 ± 5.3e6</td>
<td>40.01 ± 2.2e8</td>
<td>2.2e-9</td>
</tr>
<tr>
<td>rbs34-cl/P_lam</td>
<td>7.46e4 ± 0.11e4</td>
<td>0.0509 ± 0.0108</td>
<td>2.905 ± 1.028</td>
<td>367.8 ± 29.2</td>
<td>0.9988</td>
</tr>
<tr>
<td>rbsH-cl/P_lam</td>
<td>7.392e4 ± 0.097e4</td>
<td>0.0635 ± 0.0986</td>
<td>2.967 ± 0.324</td>
<td>272.3 ± 15.3</td>
<td>0.9999</td>
</tr>
<tr>
<td>rbs31-tetR/P_tet</td>
<td>4.484e4 ± 0.662e4</td>
<td>0.0181 ± 0.0873</td>
<td>2.299 ± 0.702</td>
<td>205.7 ± 18.7</td>
<td>0.9993</td>
</tr>
<tr>
<td>rbs33-tetR/P_tet</td>
<td>4.003e4 ± 1.34e11</td>
<td>0.5276 ± 1.32e7</td>
<td>1.156 ± 8.5e6</td>
<td>41.02 ± 2.2e8</td>
<td>1.6e-9</td>
</tr>
<tr>
<td>rbs34-tetR/P_tet</td>
<td>3.879e4 ± 0.11e4</td>
<td>0.0595 ± 0.0229</td>
<td>2.175 ± 0.303</td>
<td>690.6 ± 83.3</td>
<td>0.9994</td>
</tr>
<tr>
<td>rbs31-lacI/P_Llac</td>
<td>3.332e4 ± 0.225e4</td>
<td>5.65e-9</td>
<td>3.787 ± 1.768</td>
<td>4.32e4 ± 0.75e4</td>
<td>0.9805</td>
</tr>
<tr>
<td>rbs33-lacI/P_Llac</td>
<td>4.781e4 ± 0.261e4</td>
<td>1.45e-9</td>
<td>2.849 ± 1.018</td>
<td>5943 ± 1018</td>
<td>0.986</td>
</tr>
<tr>
<td>rbs34-lacI/P_Llac</td>
<td>2.121e4 ± 0.088e4</td>
<td>3.06e-10</td>
<td>1.866 ± 0.455</td>
<td>2.28e4 ± 0.42e4</td>
<td>0.9954</td>
</tr>
</tbody>
</table>

The good characteristics of the rapid state transition over a narrow input range, a large dynamic range and a low output level at the “off” state for the NOT gate is important because it is intended to be combined with the AND gate to compose a composite NAND gate and a correct choice of the NOT gate with an appropriate RBS is necessary to achieve a good performance. The characterisation data of the three types of NOT gate using various RBSs were next fitted to the transfer function model of the NOT gate in the form \( f([R_3]_{ss}) = k_3 (\alpha_3 + K_3 n_3 / (K_3 n_3 + [R_3]_{ss} n_3)) \) as derived in...
Chapter 3, where \([R_3]_{ss}\) is the concentration of the repressor at steady state, \(K_3\) and \(n_3\) are the Hill constant and coefficient respectively, \(k_3\) is the maximum expression level due to induction and \(\alpha_3\) is a constant relating to the basal level of the regulated promoter. The steady state levels of the repressors were derived from the fitted transfer function of the characterised inducible promoters with the same RBS as used for the NOT gate characterisation. The results for the best fit coefficients with 95% confidence bounds using the nonlinear least square curve fitting in Matlab are listed in Table 5.1, where shows the favouring characteristics of the cI/P_{lam} based NOT gate as well.

### 5.2 Engineering the Modular Composite NAND Gates

#### 5.2.1 The experimental design and construction of the NAND gates

On the basis of the characterised component gate modules and parts, the combinatorial NAND gate designed in Chapter 3 was subsequently constructed. As shown in Figure 5.6 and Figure 5.7, the NAND gate was assembled from the characterised AND gate and NOT gate modules and the pertinent characterised parts. As the individual component parts and device modules have been well characterised in their corresponding contexts, the behaviour of the composite NAND gate can be simulated directly using the parameterised models of the characterised components. The transfer function of the NAND gate can be derived by coupling the derived transfer functions of the individual modules, i.e. the NOT gate, AND gate and environment-responsive promoters, in the system. The output of a forward module acts as the input of the next module in the system cascade. As the model derivation in Chapter 3 shows, the NAND gate transfer function is give by

\[
[G]_{\text{NAND}} = f([R_3]) = k_3 \left(\alpha_3 + K_3^{n_3} / (K_3^{n_3} + [R_3]^{n_3})\right)
\]

(5.1)

where \([R_3]_{ss} = f([R]_{ss}, [S]_{ss}) = [G]_{\text{ANDmax}} \frac{([R]_{ss} / K_R)^{n_3} ([S]_{ss} / K_S)^{n_3}}{1 + ([R]_{ss} / K_R)^{n_3})(1 + ([S]_{ss} / K_S)^{n_3})} \).
where \([R] \text{ or } [S] = f([I_1]) = k_1 I_1 \frac{[I_1]}{(K_1 + [I_1])^n} \), \(I_1\) is the inducer of the inducible promoter for the regulation of \(hrpR\) or \(hrpS\) in the AND gate. All parameters in this transfer function has the same meanings as in their individual transfer functions derived previously for the environment-responsive promoters, the AND gate and the NOT gate. The only difference is the fit value of \([G]_{\text{ANDmax}}\), which needs to be adjusted according to the corresponding RBS used in the NOT gate because the previous fit is based on the rbs30-\text{gfp} reporter used for the AND gate characterisation.

Figure 5.6a and Figure 5.7a are the computational predictions of two versions of the modular composite NAND gate, which shows the desired NAND gate characteristics. The first NAND gate comprises the characterised AND gate using \(P_{\text{lac}}\) and \(P_{\text{BAD}}\) as the inputs and the rbs34-cI/\(P_{\text{lam}}\) based NOT gate, and the prediction of the device behaviour is based on their fitted transfer functions in the standard condition. The second NAND gate comprises the characterised AND gate using \(P_{\text{lux}}\) and \(P_{\text{BAD}}\) as the inputs and the rbs32-cI/\(P_{\text{lam}}\) based NOT gate, and the prediction of the device behaviour is based on their fitted transfer functions at 37 °C in the otherwise standard condition for the inducible input promoters. For the second device, it was assumed that the characteristic of cI/\(P_{\text{lam}}\) based NOT gate does not change much at the two temperatures 30 °C and 37 °C. The cI/\(P_{\text{lam}}\) based NOT gate module is used because this type of NOT gate exhibits the best characteristic among all characterised NOT gates as shown in the previous section.

Based on the model predictions, the NAND gates were constructed by directly assembling the appropriate modules and parts. The two environment-responsive inputs and the output are carried on three compatible plasmids respectively as shown in Figure 5.6b and Figure 5.7b. The \(hrpL\) followed by a NOT gate module with \text{gfp} reporter was carried on pSB4A3 (pSC101 ori, Amp\(^r\)) (Shetty et al., 2008) as the output plasmid. The two input plasmids for driving \(hrpR\) and \(hrpS\) are the same as used in the characterised AND gates. The use of separate plasmids for the inputs and output allows for the easy exchange of the individual environment-responsive inputs and the output gene, and facilitates the assembly of them into one functional device.
Figure 5.6 Two-input NAND gate using P$_{lac}$ and P$_{BAD}$ inputs, and the cI/P$_{lam}$ based NOT gate. **a**, The composite modular NAND gate comprising the characterised AND gate using inducible promoter P$_{lac}$ and P$_{BAD}$ as the inputs and the rbs34-cI/P$_{lam}$ based NOT gate, rbs30-gfp as the output reporter. The bottom is the model prediction of the NAND gate response based on the fitted models of the component gate modules at 30 °C. **b**, Plasmid maps showing circuit constructs for this NAND gate device.
Figure 5.7 Two-input NAND gate using P\textsubscript{lux} and P\textsubscript{BAD} inputs, and the cI/P\textsubscript{lam} based NOT gate. a, The composite modular NAND gate comprising the characterised AND gate using inducible promoter P\textsubscript{lux} and P\textsubscript{BAD} as the inputs and the rbs32-cI/P\textsubscript{lam} based NOT gate, rbs30-gfp as the output reporter. The bottom is the model prediction of the NAND gate response based on the fitted models of the component gate modules at 37 °C. b, Plasmid maps showing circuit constructs for this NAND gate device.
5.2.2 The characterisation of the NAND gates

The two assembled devices were next characterised in their corresponding conditions. The devices were first fully characterised by fluorometric assays in a fluorometer as the results shown in Figure 5.8a and Figure 5.9a. It can be seen that the responses of the two versions of the NAND gate mimic the digital logic NAND function well, where the outputs of gates are in the “off” state (Point A) only when both inputs are high. Points B, C and D show high fluorescence and point A low fluorescence, which are the outputs corresponding to the four logic input conditions. The experimental results are close to the model predictions apart from the slight difference at the corner with low IPTG and high arabinose induction (Point B, Figure 5.8a) on the two dimensional map of full characterisation and, similarly, in Figure 5.9a, the corner with low AHL and high arabinose induction. The slight inhibition at Point B might be due to the small leaky expression of HrpR from the corresponding input inducible promoters.

The NAND gate circuits were then subject to flow cytometry assay under four logic input conditions and borne out to work at individual cellular level as well (Figure 5.8b and Figure 5.9b), where the data were acquired 16 hours after induction from the cultures inoculated from a single colony containing the circuit. Referring to Figure 5.8b and Figure 5.9b, the entire cell population is turned off only when both inputs are highly induced (low fluorescent lane A corresponding to the Point A in the fluorometric assay). When either input is not induced, the entire population is switched on (high fluorescent lanes B, C and D). The high fluorescence of a small portion of the cells at lane A of the second NAND gate (Figure 5.9b) might be due to a not fully induced cell population under the induction conditions employed (i.e. 1.3 mM arabinose plus 100 nM AHL). Particularly, this might be due to the non-homogenous input promoter P_{BAD} which displays all or none (bimodal) activity at subsaturating arabinose conditions in E. coli MC1061 as shown in Section 4.4.1. Overall, the results are consistent with the full characterisation results at the population level by fluorometric assay as shown in Figure 5.8a and Figure 5.9a.
Figure 5.8 The systematic characterisation of the first composite NAND gate. **a**, The NAND gate responses by fluorometric assay for 64 combinations of input inductions by (left to right) 0, $3.9 \times 10^{-4}, 1.6 \times 10^{-3}, 6.3 \times 10^{-3}, 2.5 \times 10^{-2}, 0.1, 0.4$ and $1.6$ mM IPTG, and induction by (bottom to top) 0, $3.3 \times 10^{-4}, 1.3 \times 10^{-3}, 5.2 \times 10^{-3}, 2.1 \times 10^{-2}, 8.3 \times 10^{-2}, 0.33$ and $1.3$ mM arabinose. Data are the normalised average of three repeats in *E. coli* MC1061 in M9-glycerol at 30 °C with variations less than 10% between biological repeats. **b**, FACS assay of the NAND gate under four logic combinations of input inductions: A - 1.3 mM arabinose plus 1.6mM IPTG; B - 1.3 mM arabinose; C - 1.6mM IPTG; D - none.
Figure 5.9 The systematic characterisation of the second composite NAND gate.

a, The NAND gate responses by fluorometric assay for 64 combinations of input inductions by (left to right) 0, 2.4 × 10^{-2}, 9.8 × 10^{-2}, 3.9 × 10^{-1}, 1.6, 6.3, 25 and 100 nM AHL, and by (bottom to top) 0, 3.3 × 10^{-4}, 1.3 × 10^{-3}, 5.2 × 10^{-3}, 2.1 × 10^{-2}, 8.3 × 10^{-2}, 0.33 and 1.3 mM arabinose. Data are the normalised average of three repeats in E. coli MC1061 in M9-glycerol at 37 °C with variations less than 10% between biological repeats. b, FACS assay of the NAND gate under four input inductions: A - 1.3 mM arabinose plus 100 nM AHL; B - 1.3 mM arabinose; C - 100 nM AHL; D - none.
The long growth time of the cell cultures for this FACS assay is used because the NAND gate default output is high without input inductions, which can lead to high fluorescence background after short time growth upon induction, due to inoculating cells carrying over high GFP levels. As shown in Figure 5.10a-c, the cells harbouring the first composite NAND gate were subject to the flow cytometry assay after 5 hours growth with various input inductions. It can be seen that the output in terms of fluorescence is reduced when both input inducers are present, but the lowest output is still high even when both inputs are fully induced. By extending the growth time to 7 hours, the lowest fluorescent output of the device is reduced further under the full induction of both inputs, as Figure 5.10d shows. This is due to the cells with initial high fluorescence having undergone a longer period of growth with more cell divisions, which resulted in more dilutions of the background fluorescent proteins in each cell. Thus, the assay cell cultures were finally inoculated from a single bacterial colony for long time growth (16 h) to completely dilute out the background fluorescent proteins. The complete dilution of the initial background GFPs is corroborated by the FACS assay results of these cultures (Figure 5.8b and Figure 5.9b).
Figure 5.10 FACS assays of the engineered composite NAND gate using $P_{lac}$ and $P_{BAD}$ as the two inputs in *E. coli* MC1061 after 5 hours (a-c) or 7 hours (d) growth in M9-glycerol at 30 °C. **a**, FACS assay of the NAND gate with full induction of the $P_{BAD}$ input (1.33 mM arabinose) and graded induction of the $P_{lac}$ input by (bottom to top) 0, 3.9 × 10^{-4}, 1.6 × 10^{-3}, 6.3 × 10^{-3}, 2.5 × 10^{-2}, 0.1, 0.4 and 1.6 mM IPTG. **b**, FACS assay of the NAND gate with full induction of the $P_{lac}$ input (1.6 mM IPTG) and graded induction of the $P_{BAD}$ input by (bottom to top) 0, 3.3 × 10^{-4}, 1.3 × 10^{-3}, 5.2 × 10^{-3}, 2.1 × 10^{-2}, 8.3 × 10^{-2}, 0.33 and 1.33 mM arabinose. **c**, FACS assay of the NAND gate with graded inductions for both of the inputs $P_{lac}$ and $P_{BAD}$. **d**, FACS assay of the NAND gate after 7 hours growth under four logic combinations of input inductions.
5.3 Discussion

This chapter describes the engineering and quantitative characterisation of a set of modular genetic logic NOT gates as well as the combinatorial modular NAND gates. Particularly, the composite NAND gates are directly assembled from the characterised component gate modules, i.e. the AND and NOT gates, whose behaviour can be accurately predicted from the fitted transfer functions of the individual modules in the corresponding contexts. This further demonstrated that the assembly of BioParts into functional devices and systems can be executed predictably and reliably using previously ‘in-context’ quantitatively characterised parts and submodules guided by modelling. Combining the results in Chapters 4 and 5, they show that the functional assembly approach proposed in this thesis not only works at the basic part level, but also at the device level.

Strictly speaking, the component modules should be characterised in the exact same biophysical and genetic context as in their final target system. This will minimise the variations arising from the difference of contexts. However, the same genetic context is more difficult to achieve than the same biophysical context. For instance, the 5' UTR sequence in the context for characterising the module might be different from the one in the target system. Thus, it will be significantly beneficial to standardise the 5' UTR sequence before the RBS site used for the characterisation. Unfortunately, there is no work on this aspect currently in the synthetic biology community. As shown in Chapter 4, the difference of genetic context such as the 5' UTR can lead to the large variations of the absolute activities of parts and modules. It will be worthwhile to systematically design several standard 5' UTRs of various lengths and secondary structures to investigate their impact on characterisation brought by these variations. In addition, a set of standard inducible characterisation systems need to be developed to standardise the characterisation of synthetic parts and modules, i.e. in the same genetic context and biophysical (abiotic) context. This will enable the reuse and exchange of the characterisation data between different labs. The PBAD, P lac and P tet inducible promoters are the good candidates for such systems.
Ideally, these inducible promoters need to be placed in a standard robust vector to minimise the variations of plasmid copy numbers.

The engineered modular NOT and NAND gates are reusable and can be reconnected to different sensor inputs to detect and integrate various environmental signals or easily incorporated into a large system as fundamental building blocks to regulate the genetic information processing in a specified logic manner. The NOT and NAND gates are the basic modules for engineering genetic circuits with desired logic functions just as their counterparts in electronic circuit engineering. To expand the toolbox of synthetic biology, more orthogonal modules with logic functions need to be engineered from diverse natural biological modules. As a result, multiple modules with the same logic can be used in one composite system, and large systems with increasingly complicated logic functions can be developed just like the combinatorial logic circuits seen in electronic engineering.
Chapter 6

Engineering a Modular Cell Density-dependent Microbe-based Biosensor
In this chapter, on the basis of the gene circuits engineered in the previous chapters, a microbe-based biosensing application platform was then built. As mentioned before, the constructed AND, NOT and NAND gate devices are all modular and promoter-input based. Thus, various promoters that are responsive to specific environmental signals can be connected to these modular devices, and the \textit{E. coli} hosts harbouring the device can respond with a corresponding output when the signal in the environment appears or a specific logic combination of multiple signals is detected. In this sense, the engineered microbe-based biosensor encompasses the three equivalent modules of a traditional sensor, i.e. the sensory, regulatory and the actuation modules, which are the parts corresponding to the environment-responsive input promoters, the engineered gene circuits and the output genes in the biosensor. The biosensor conforms to a modular structure because each module is within limits independent and exchangeable in the chosen framework.

On the other hand, the natural world has shown that more complex tasks can be performed precisely through the cooperation of otherwise separated individuals, such as the developmental pattern formation in some eukaryotes and the quorum sensing mechanism found in some specific bacteria (Bassler, 2002; Surette and Bassler, 1998; Waters and Bassler, 2005). Usually, the coordinated jobs are completed through the communication of biochemical signals between them (Taga and Bassler, 2003), which are regulated by the underlying gene circuits in each host. Thus, to program the cells with high level functions, we need to exploit these natural cell-cell communication circuits and to engineer a function that is dependent on the behaviour of the entire population instead of only the individuals. Here in Section 6.1, a module was engineered in \textit{E. coli}, which regulates gene expression as a function of the density of the whole population. The cell density-dependent module is inspired from the quorum sensing mechanism in \textit{Vibrio fischeri}, a marine symbiotic bacterium. Based on this cell density-dependent module and other engineered parts and modules in the previous chapters, an \textit{E. coli} based cell density-dependent biosensor was then constructed in Section 6.2, which integrates its own cell density signal through a cell-cell communication module and a second environmental signal in a logic AND manner, with green fluorescent protein level as the output readout.
6.1 Quorum Sensing and the Engineering of a Cell Density-dependent Module

6.1.1 The quorum sensing mechanism in Vibrio fischeri

It is now known that the ability to communicate is essential for the survival and interaction of many different species of bacteria in their natural habitats. One cell-cell communication mechanism, named quorum sensing, was first discovered and described in the marine symbiotic bacterium Vibrio fischeri about 35 years ago (Fuqua and Greenberg, 2002; Miller and Bassler, 2001). V. fischeri usually exists in two states. When in the free-living state, it swims freely in seawater at a low cell density (approximately 100 cells or less per ml) and emits almost no light. When in its symbiotic light organ of some eukaryotic hosts, like the Hawaiian squid Euprymna scolopes, the bacteria can grow up to a concentration of $10^{10}$ to $10^{11}$ cells per ml and emit light. This autoinduction of light emission is intertwined with cell-cell communication and cell density. Due to its correlation to the population density, the mechanism underlying this behaviour is referred to as quorum sensing. The host and the bacteria are in a symbiotic relationship. The bacteria can get nutrients continuously from the host while the host uses the bacterial bioluminescence for various specific purposes like the squid avoiding its potential predators by counter illumination.

The quorum sensing system was first identified by the direct transfer of a 9 kb fragment from V. fischeri to E. coli (Engebrecht et al., 1983), and the transformed cells surprisingly showed bioluminescence. Following this work, the quorum sensing system has been extensively studied and the basic regulatory structure is found conserved in a variety of other bacterial species as well. Figure 6.1 shows the disclosed gene regulatory circuit underlying the quorum sensing mechanism in V. fischeri. The mechanism is controlled by the lux regulatory system, which comprises the left luxR operon and right luxI operon. The right operon transcribes a set of six genes (luxICDABE) encoding the luciferase enzymes required for light production, in which the LuxI proteins expressed from luxI are the synthase for the signalling molecule acyl-homoserine lactone (AHL). The signalling molecule AHL can freely
Chapter 6 - Engineering a Modular Cell Density-dependent Microbe-based Biosensor

Figure 6.1 The quorum sensing circuit in *Vibrio fischeri* and the regulatory region of *lux* operon. a, The quorum sensing circuit in *V. fischeri* comprises two *lux* operons transcribed in two directions respectively. The five luciferase structural genes (*luxCDABE*) and two regulatory genes (*luxR* and *luxI*) are required for the cell density-dependent bioluminescence in *V. fischeri* through a positive feedback circuit. b, The 218 bp intergenic region between the left and right *lux* operons represents the control region for the bioluminescent system. The regulatory region contains the two promoters for the two operons, and the *crp* box and *lux* box which are the binding sites of CRP-cAMP and LuxR-AHL complexes respectively. The left promoter has the full -10 and -35 binding sites for $\sigma^{70}$ whilst the right promoter only contains the -10 site.

diffuse across the cell membrane. The left operon encodes the LuxR proteins under a separate constitutive promoter. The constitutively expressed LuxR binds its cognate AHL molecule to form the LuxR-AHL complex, which can bind a site, named *lux* box, on the regulatory region of the *lux* operon and activate the transcription of the right operon. Thus, at low cell densities, the right *luxI* operon is transcribed at a low basal level and a low level of the autoinducer AHL is produced. Since the genes encoding
luciferase are located downstream of the luxI gene, only a low level of light is produced. When V. fischeri culture grows to a certain density, the AHL molecules accumulate to a threshold in the culture and AHL can then bind its cognate LuxR proteins to activate the transcription of the rightward operon. Due to the positive feedback, it results in an exponential increase in both the AHL production and bioluminescence. The LuxR-AHL complex also negatively regulates the left promoter for LuxR expression as a compensatory mechanism which decreases the expression of the genes in the rightward operon in response to the positive feedback circuit (Miller and Bassler, 2001).

The detailed regulatory region for the lux system is shown in Figure 6.1b. The regulatory region includes the two regulatory promoters and the two conserved binding sites, i.e. the crp box and lux box, for the CRP-cAMP complex and LuxR-AHL complex respectively. The crp box allows the catabolite repression of the left luxR operon. The lux box is a 20 bp palindromic sequence centred approximately 40 bp upstream from the transcription start site of the right operon (Stevens and Greenberg, 1997), which is a conserved binding site for the LuxR-like family proteins. The specificity in response to a signalling molecule is encoded in the N-terminal regulatory domain of LuxR-like proteins, and the signaling molecule used in V. fischeri is N-3-oxohexanoyl-L-homoserine lactone.

### 6.1.2 The engineering of a cell density-dependent module

In this section, a genetic module for cell-cell communication in a cell density-dependent manner was designed and constructed on the basis of the lux quorum sensing regulatory system in V. fischeri as introduced in Section 6.1.1.

Figure 6.2 shows the design of the genetic module. The module was created on the basis of the synthetic P_{lux} promoter that has been characterised in Chapter 4 and the luxR and luxI regulatory genes from the lux quorum sensing system. Although the natural lux regulatory system has been isolated and demonstrated to have the population density-dependent behaviour, a modular design approach was taken here in
building the module using standardised BioBrick parts based on the intrinsic mechanisms of quorum sensing (see Section 6.1.1). The module is different from the natural one, but, essentially, employs the same positive feedback mechanism. The $P_{\text{lux}}$ promoter contains the $\text{lux}$ box and the right $P_{\text{luxI}}$ promoter from the native $\text{lux}$ system. The $\text{luxR}$ gene is constitutively expressed under the $P_{\text{tet}}$ promoter. The $\text{luxI}$ and $\text{gfp}$ are expressed as an operon under the regulation of the $P_{\text{lux}}$ promoter. Thus, when at low cell densities, LuxI and GFP proteins are expressed at low levels and the AHL molecules synthesised from LuxI is also at a low level and can not sufficiently bind LuxR to activate the transcription from $P_{\text{lux}}$. When the cells continue to grow, the AHL accumulates in the environment and reaches a threshold at a certain cell density, at which it sufficiently binds LuxR to activate the positive feedback loop as shown in the graph and then more LuxI and GFP proteins are expressed. As a result, the output in terms of fluorescence increases with cell density, and the module has the population density-dependent behaviour.

To conform to the BioBrick standard, all the module components (promoters and genes) were obtained from the Registry of Standard Biological Parts and were assembled on the BioBrick standard vector pSB3K3 as shown on the right of Figure 6.2a. Since the $\text{luxI}$ and $\text{gfp}$ are expressed as an operon in the module, the $gfp$ expression level is likely different from the one that is first expressed downstream from the $P_{\text{lux}}$ promoter even with the same RBS used. Polarity is known to occur in bacterial systems, and later transcribed genes in an operon are often not as highly expressed as those initially transcribed. Such effects can be due to changes in the amounts of message where the message is degraded from its 3' end, competing RNA structures that disfavour translation, and so on. This view is borne out by the initial experimental results which showed that, for the constructs with $\text{rbs33-gfp}$ or $\text{rbsH-gfp}$ downstream the $\text{luxI}$ gene, there were nearly no fluorescent outputs of the cells harbouring the module for all 6 RBSs ($\text{rbs30 - rbsH}$) used for $\text{luxI}$ (data not shown). Thus, the strong RBS – $\text{rbs30}$ was finally chosen for obtaining measurable levels of $gfp$ expression, which gave a clear fluorescent output for cells harbouring the module. The weak RBS – $\text{rbs33}$ for $\text{luxI}$ in the final construct was chosen from the screenings which showed that, for all 6 RBSs ($\text{rbs30 – rbsH}$) used for $\text{luxI}$ gene ($\text{luxI-rbs30-gfp}$),
Figure 6.2 The engineering of a cell density-dependent module. a, The regulatory module is designed on the basis of the synthetic AHL-responsive \( P_{\text{lux}} \) promoter and the \( \text{luxI} \) gene, which forms a positive feedback loop in response to the density of the cell population. On the left is the plasmid construct of the module. The \( \text{luxI} \) gene is modified with a LVA degradation tag. b, The characterised results of the module in \( E. \ coli \) MC1061 after 5 hours growth in M9-glycerol at 37 °C. The sample cultures are initially inoculated with various numbers of cells of OD\(_{600}\) at (bottom to top) \( 1 \times 10^{-4}, 3 \times 10^{-4}, 5 \times 10^{-4}, 1 \times 10^{-3}, 2 \times 10^{-3}, 5 \times 10^{-3}, 1 \times 10^{-2}, 2 \times 10^{-2} \) and \( 5 \times 10^{-2} \). The negative control is the culture of cells carrying the \( \text{gfp} \)-free construct, and the positive control is the culture of cells carrying the functional construct induced with 100 nM AHL. The inoculation density for both the negative and positive controls is of OD\(_{600}\) at 0.02. On the right is the plotted curve for the module output (mean fluorescence from FL1 filter) as a function of cell density for the FACS assays.

where there was only significant output difference between the non-induced culture and fully induced culture (by 100 nM AHL) for cells harbouring the module with \( \text{rbs33-luxI-rbs30-gfp} \). Therefore, the rbs33 and rbs30 were finally selected for \( \text{luxI} \) and
The constructed module (Figure 6.2a) was next characterised in *E. coli* MC1061 by FACS assay and the results are shown in Figure 6.2b. To prepare the day cultures, the overnight cultures were pelleted by centrifugation, and re-suspended in pre-warmed M9 media and left for 10 mins before being diluted into the day cultures at various cell densities. This washing step is important for removing the AHL molecules expected to have built up in the media because the overnight culture was likely saturated with high level of AHL molecules after a long time of growth. The diluted day cultures were next grown at 37 °C for 5 hours to various cell densities (measured by spectrophotometer at 600 nm) before being harvested for analysis by flow cytometry. Referring to Figure 6.2b, it shows that the cells containing the engineered module displayed the desired cell density-dependent property. The fluorescent outputs of the sample cultures with various initial cell densities varied from each other after the same length of time of growth, and the output increases with the density of the cell culture. Extrapolating back towards a fully “off” low density cell culture, it seems that the positive feedback of the module was already partly switched on at a lower cell density than the lowest density shown in the graph. However, due to the detection limit of the spectrophotometer used, the lowest OD$_{600}$ that can be accurately measured for the cell culture is 0.01. Although the completely “off” culture was not directly observed during the experiment, the cell cultures analysed indeed exhibited the density-dependent behaviour. It is known that AHL (Acyl-homoserine lactones) molecules are quite stable in mildly acidic or neutral pH environments (Leadbetter and Greenberg, 2000). The quorum-sensing signalling molecules AHL are found to degrade significantly only in some specific bacterial species like *Bacillus* bacteria which can express the AiiA enzyme to degrade the signalling molecules (Dong et al., 2000) or in *Variovorax paradoxus* which can metabolise the signalling molecules (Leadbetter and Greenberg, 2000). Thus, the effect of AHL degradation can be negligible in the *E. coli* chassis used here. However, if used in an *E. coli* chassis expressing the AiiA degrading enzyme, it might slow down the accumulation of AHL in the media and lead to a delayed density switching response, i.e. switching at a higher cell density.
6.2 Engineering a Modular Cell Density-dependent Microbial Biosensor

In this section, a microbe-based biosensing application platform is described which uses the biological parts and modules that have been engineered in this project. Specifically, an *E. coli*-based biosensor has been constructed, which can sense in the environment (i) its own cell density and (ii) the arabinose concentration. The observed GFP level was the output readout for the device. The design and construction of the biosensing system is introduced in Section 6.2.1 and the characterisation results are described in Section 6.2.2.

6.2.1 System design and construction

Figure 6.3a shows the detailed structure of the designed microbe-based biosensing system, which consists of three independent and exchangeable modules (i.e. the sensory module, the internal regulatory module and the reporter module). The input sensors are inducible promoters which can respond to specific molecules either inside the cell or outside the cell via a signalling pathway. The regulatory module is an engineered genetic circuit that can transmit and integrate the changes from the input promoters to generate a single output signal in a pre-determined way. The reporter module is an output gene that is expressed in proportion to the signal level from the regulatory module. The reporter protein, e.g. GFP, usually can be detected by a traditional biophysical or biochemical assay. Thus, the microbial biosensor is modular and can be easily customised to detect various environmental signals and generate a quantifiable output in a desired logic manner according to different applications. Here, a cell density-dependent biosensor was constructed which uses the cell density-dependent module engineered in the previous section and the arabinose-responsive PBAD promoter as the two sensor inputs, the logic AND gate engineered in Chapter 4 as the internal regulatory circuit, and the *gfp* as the output reporter.

As shown in the graph (Figure 6.3a), the first sensor input of the device is the LuxI/LuxR quorum sensing module, which can synthesise and detect the AHL signalling molecules. The positive feedback in the module enables the response level
Cell density-dependent biosensor using P_{BAD} as the 2\textsuperscript{nd} input

**Figure 6.3** The engineering of a cell density-dependent microbe-based biosensor.  
\textbf{a}, The biosensor is designed to comprise three exchangeable modules, i.e. the promoter-based input sensors, the internal circuit for regulating genetic information transmission, and the output gene reporter. Here, the two sensor inputs used are for detecting the cell density and the arabinose (P_{BAD}) respectively.  
\textbf{b}, Plasmid maps showing the circuit constructs used for the characterisation of the biosensor.

The second sensor input, the inducible P_{BAD} promoter, detects and responds to arabinose in the environment. The AND gate circuit integrates the levels of the two sensor inputs in the specific logic AND gate manner as characterised before, and drives the expression of the \textit{gfp} reporter. As a result, the biosensor will generate a clear output (fluorescence) when both the cell population density and the arabinose concentration are sufficiently high and reach the detection
thresholds used by the device. This tight control property is useful for some applications in which we would like the cells to produce a product in large amounts and in a desired environment while the product is toxic to the cells. The desired environmental condition could be defined by the second sensor input of the biosensor, such as using an engineered pH sensitive promoter to specify the acidic condition. Therefore, the cells will not produce the product until they reach a certain high density to avoid the growth inhibition effect induced by the product during the early growth phase.

Figure 6.3b shows the plasmid constructs of the cell density-dependent biosensor capable of arabinose sensing. The system was assembled directly from the individual modules that have been characterised (see Chapters 4-6). The three plasmids represent the cell density-dependent input (Input 1, plasmid pBW630luxI-hrpR driving \( hrpR \)), the arabinose input (Input 2, plasmid pBW213ara-hrpS driving \( hrpS \)) and the output (plasmid pBW400hrpL-gfp driving \( gfp \)) respectively. The plasmid separation of the two inputs and output allows modular and quick exchange of the sensor inputs and the output responsive gene.

6.2.2 Characterisation: results and analysis

The plasmid constructs of the biosensor were next transformed in \( E. \ coli \) MC1061, which was subject to flow cytometry assay. Figure 6.4 shows the characterisation results, where the sample cultures were initially inoculated from the same overnight culture, but with various numbers of cells and were grown to different cell densities after the same length of time. As explained in the previous section, the overnight cultures were washed with M9 media before being diluted into the day sample cultures. The biosensor was first assayed under the condition of full induction by 1.3 mM arabinose (Figure 6.4a, c). It shows that the output (in terms of fluorescence) of the biosensor sample cultures varied after the same length of time of growth, and the output increases with the cell density of the culture. The biosensor was next assayed under the condition of no arabinose induction (Figure 6.4b, c). The results show that there is nearly no output response without the presence of arabinose in the culture.
Figure 6.4 Characterisation of the cell density-dependent biosensor. The cells (E. coli MC1061) harbouring the engineered biosensor were assayed by flow cytometry after 5 hours growth in M9-glycerol at 37 °C. The negative control is the culture of cells containing the gfp-free constructs with inoculation cell density of OD$_{600}$ = 0.05.

a, FACS assay of sample cultures with 1.3 mM arabinose. The sample cultures were initially inoculated with various numbers of cells from the same overnight culture at OD$_{600}$ of (bottom to top) 5 × 10$^{-4}$, 1 × 10$^{-3}$, 2 × 10$^{-3}$, 1 × 10$^{-2}$, 2 × 10$^{-2}$, 3 × 10$^{-2}$, 5 × 10$^{-2}$, 1 × 10$^{-1}$ and 2 × 10$^{-1}$. b, FACS assay of sample cultures without arabinose. The sample cultures were initially inoculated with various numbers of the cells from the same overnight culture at OD$_{600}$ of (bottom to top) 5 × 10$^{-4}$, 1 × 10$^{-3}$, 1 × 10$^{-2}$, 5 × 10$^{-2}$, and 2 × 10$^{-1}$. c, The plotted curves of the biosensor output (mean fluorescence from FL1 filter) as a function of cell density for the assays of a and b.
whatever the cell density of the culture is. Combing the two assay results leads to conclusion that the engineered biosensor exhibits the cell density-dependent sensory property as well as the arabinose sensing capability. This means that the engineered biosensor will not output with high fluorescence until both the cell density and the arabinose concentration are sufficiently high and each reach a certain threshold.

As seen in the characterisation of the cell density-dependent module, the completely “off” culture of the biosensor was not observed under the condition of high arabinose concentration (Figure 6.4a). This might be due to the cells harbouring the biosensor switch still in an “on” state even at a lower cell density than the lowest cell density shown in the graph. Working below this already low cell density is below the detection limit of the spectrophotometer used, and so seeking a fully “off” cell population is technically not easy unless cell density was checked by e.g. manually counting cells directly in a microscope or through colony viable counting. The partly “on” population of cells studied here is reflected by the two peaks of GFP levels seen in the sample culture of OD$_{600}$ of 0.010, one showing low fluorescence (“off” state cells) while the other higher fluorescence (“on” state cells). The two subpopulations were also observed on the scatter graphs of the FACS assays. Figure 6.5 shows the side-scattering and forward-scattering graphs for the first four sample cultures (Figure 6.4a) with OD$_{600}$ of 0.010, 0.020, 0.034 and 0.088. The same gate was applied to each graph for calculating the fluorescent outputs of the samples. The first graph clearly shows that there are two populations corresponding to the high and low fluorescence respectively. With the increase of the cell culture density, the low fluorescent subpopulation gradually disappears and the high fluorescent subpopulation begins to dominate as shown in Figure 6.5 b-d.
Figure 6.5 FACS side-scattering and forward-scattering graphs of the cells (*E. coli* MC1061) harbouring the engineered biosensor. The sample cultures are from the assays in Figure 6.4a, which were initially inoculated with various numbers of cells and grown to various densities after 5 hours growth in M9-glycerol at 37 °C. a, FACS scatter graph for the sample culture of OD$_{600}$ = 0.010. b, FACS scatter graph for the sample culture of OD$_{600}$ = 0.020. c, FACS scatter graph for the sample culture of OD$_{600}$ = 0.034. d, FACS scatter graph for the sample culture of OD$_{600}$ = 0.088.
6.3 Discussion

In this chapter, a cell-cell communication module was successfully engineered with the cell density-dependent sensory property, which is based on the *lux* quorum sensing system in *Vibrio fischeri*. By integrating the biological parts and modules that have been engineered in the project, a scalable application platform for microbe-based biosensors was built. Due to the modular structure of the platform, these promoter-based microbial biosensors can be engineered to detect and integrate various environmental conditions precisely using various input-specific promoters, and initiate different cellular responses on demand. Specifically, an *Escherichia coli* based biosensor were constructed, which can sense its own cell density signal and a second environmental signal (arabinose concentration) and integrate them in a logic AND gate manner, with GFP level as the output readout. The work here further validates the modular functional assembly approach proposed in this thesis. It also represents the capability that we can design an engineered version of the naturally discovered biological circuits using standardised biological components. In addition, we could program emergent, high level functions into the living cells through the communication and coordination of the individuals instead of just the standalone cellular behaviour. Up to now, various advanced functions have been implemented by synthetic biologists using multiple cell populations or species that have been engineered with different dedicated circuits and thus can communicate with each other in a pre-determined manner, such as programmed pattern formation (Basu et al., 2005), artificial ecosystems (Balagadde et al., 2008; Brenner et al., 2007; Song et al., 2009; Weber et al., 2007) and synchronised oscillation (Danino et al., 2010).

However, the experimental results show that the transition region of the engineered cell density-dependent module is not completely within the readily observable range of cell densities. It seems that the positive feedback in the module switches on at a lower threshold of density than the lowest one shown in the experiment. This is largely due to that the current architecture of the module is not optimised for the experimentally observable region of cell density (OD$_{600}$ between 0.01 and 1.00). Thus as an alternative to studying cells at very low cell density, three
modified architectures of the cell density-dependent module are proposed here for future improvement as Figure 6.6 shows. In the first architecture (Figure 6.6a), by using a weak RBS for luxR instead of the currently strong RBS (rbs34), the density of the induction threshold of the module will increase and thus likely falls in the observable range of cell density. The effect of altering luxR RBS strength has been shown in a study of shuffling the architecture of the natural lux operon (Haseltine and Arnold, 2008). In the second architecture (Figure 6.6b), luxI and gfp (output gene) are expressed separately under the same P_{lux} promoter instead of as an operon. This will improve the modularity and predictability of the module because the balanced expression of multiple genes in an synthetic operon are difficult to achieve (Pfleger et al., 2006). As observed in the experiment, the second gene (gfp) expression in the operon of the engineered module has varied a lot compared to the single gene (gfp) expression directly under the inducible promoter P_{lux}. In the third architecture (Figure 6.6c), the luxR is regulated by the P_{lux} promoter instead of the P_{tet} constitutive promoter. Thus, luxR expression is under the positive feedback control as well as the expression of luxI and gfp. This structure will increase the transition sharpness of the module as shown in a relevant study (Haseltine and Arnold, 2008), but might lead to a bistable response due to the increased strength of the positive feedback.

It will be of great value to investigate the behaviour of the three proposed architectures above. The study will assist the engineering of a genetic module with customised cell density-dependent property for its wide application in many areas, for example in biotechnology including fermentation.
Figure 6.6 Potential improved architectures for the cell density-dependent module.
Chapter 7 - Overview and Future Work

Chapter 7

Overview and Future Work
7.1 Overview

Synthetic biology represents a new, interdisciplinary field at the interface between engineering and biology. The initial proof-of-principle examples in this area are encouraging and exhibit the great potential of designing de novo genetic circuits from modular biological elements, under the guidance of some engineering principles. However, the materials that are to be engineered in gene circuits are fundamentally different from their engineering counterparts. The living biological materials are inhabited in a complex, yet orderly, aqueous environment, self-replicable and subject to continuous background fluctuations. Thus many engineering principles may not be directly applicable to the entire process of synthetic gene circuit engineering. In addition, the complexity of natural biological systems can be overwhelming and we still do not know enough details of their design principles and underlying components. This is reflected in the relatively slow progress made towards engineering large complex biological circuits, with desired properties, during the past decade. The complexities of currently bottom-up constructed systems appear to enter into a platform with no more than six regulatory elements in one system (Purnick and Weiss, 2009). To a large extent, engineering of a simple biological system, even with just a few components, is still an ad hoc research process with a number of trial-and-error steps and some retrospective empirical tinkering (Kwok, 2010; Lu et al., 2009). Several major challenges remain to be solved before the engineering of biology has the same predictability and reliability as seen in other mature engineering disciplines. Therefore, the study in this thesis has been centred on overcoming some current bottlenecks in gene circuit design and construction.

The thesis introduces the design methodology for engineering modular and orthogonal biological devices to tackle the modularity and orthogonality problems. A novel genetic logic AND gate was designed and implemented in *Escherichia coli* by redesigning the natural HrpR/HrpS hetero-regulation motif in the specific *hrp* system of *Pseudomonas syringae pv tomato* DC3000. The essential gene components of the AND gate are exogenous and thus orthogonal to the *E. coli* genetic background. Both inputs and output of the AND gate are promoters and thus the device is designed to be
modular and reusable. The inputs can be easily rewired for various input promoters and the output can be connected to different genes to drive various cellular responses. Furthermore, the AND gate applies an alternative regulation mode, i.e. the $\sigma^{54}$-dependent gene transcription, to allow tight control of the underlying transcriptional activation and closely mimic digital logic AND behaviour. The successful engineering of this device sets an example for harnessing the diverse natural biological networks to design modular and orthogonal devices to expand the currently limited toolbox of synthetic biology.

The thesis determines the effect of context on part and device behaviour through the systematic quantitative characterisation of a series of biological parts and devices, e.g. the three inducible promoters ($P_{lac}$, $P_{BAD}$ and $P_{lux}$) and six ribosome binding sites ($\text{rbs30-rbsH}$), in various biophysical and genetic contexts. It shows that both the biophysical (e.g. media, temperature, carbon source) and genetic (e.g. chassis background, embedded genetic sequence context) contexts in which the parts and devices behave have a large impact on their behaviour. This suggests that we should take into account the context effect when designing and characterising synthetic parts and devices. This is corroborated by the forward engineering of the designed logic AND gate. The current characterisation scope for parts and devices were extended by characterising the homogeneity, metabolic load and chassis compatibility of the engineered AND gate beyond the normal population-averaged behaviour. The results show that that the engineered AND gate is homogenous, though the whole device relies on the homogeneities of the input promoters. It imposes a negligible metabolic load on the tested host chassis ($E. coli$ MC1061) and functions compatibly across the majority of the seven tested $E. coli$ strains. Thorough characterisation of parts and devices is also of great importance for their wide reuse within the synthetic biology community.

The thesis presents a new effective strategy for assembling individual synthetic parts and devices into functional biological systems using engineered ‘in-context’ quantitatively characterised modules. By characterising the candidate component parts and modules in the same biophysical and genetic context as anticipated for their final target system, it eliminates or reduces the behavioural variations arising from the
difference of working contexts and thus increases the predictability of circuit construction. As a result, this approach minimises the unexpected and other high-order effects which are likely to occur during circuit construction. Following this functional assembly approach, the modular logic AND gate was successfully engineered with behaviour that can be precisely predicted from the components parts (i.e. promoters and RBSs) characterised in various contexts. The subsequent fit of the abstract AND gate model to the characterisation data showed that the two enhancer binding activators HrpR and HrpS are likely to play unequal roles in the cooperative activation of the $\sigma^{54}$-dependent hrpL promoter. A set of modular logic NOT gates (i.e. the lacI/P$_{lac}$, tetR/P$_{tet}$ and cI/P$_{lam}$ based inverters) were also constructed and characterised with different RBS versions. Under the guidance of modelling, a set of composite combinatorial logic NAND gates were then assembled from the AND and NOT gate modules with behaviour that can be quantitatively predicted from the individual transfer function models of the component, ‘in-context’ characterised parts and submodules. Hence, the proposed functional assembly method was demonstrated to work not only at the basic part level, but also at the device level.

Lastly, the thesis presents the design and construction of a cell density-dependent module using standard modular biological parts on the basis of the Vibrio fischeri lux operon. The module allows the cell population to co-ordinate local cellular behaviour through cell-cell communication, although the density transition region of the module remains to be improved. By assembling the multiple parts and modules that have been engineered in this study, a cell density-dependent microbe-based biosensor was constructed. This performs as a logic AND gate by integrating its own cell density signal, through a synthetic cell communication module, and a second input signal, through an environment-responsive promoter. The successful construction of this in vivo biosensing platform signifies that more advanced biological devices can be engineered through cellular cooperation and the integration of many modular biological parts and devices with practical applications.

Thus the work in this thesis describes the steps I have taken towards the goal of making routine the engineering of synthetic gene circuits. There is a major focus on the design and engineering of modular and orthogonal biological parts and devices, as
well as their predictable functional assembly. Though many challenges remain to be overcome, the task of making biology easier to engineer will continue since the benefits of programming cells by customised gene circuits are enormous ranging from practical applications in industry and medicine, to uncovering the design principles of natural biological circuits.

7.2 Future Work

This section discusses future research directions relating to the work of the thesis and other potential research projects arising from the current study.

1. Part/device characterisation and improvement of modularity

As shown in the thesis, the context in which the biological parts and devices behave has a large impact on their behaviour. The majority of currently constructed synthetic biological parts and devices often only work for their own specific purposes in a particular context, and lack sufficient modularity and reusability. Thus the future direction of part (device) standardisation and characterisation should be directed towards the decoupling of context dependency, in particular the dependency on the embedded sequence context, to improve their modularity. For instance, it would be advantageous to design several standard 5' UTR sequences ahead of the RBS site in order to lower the interference with adjacent parts, i.e. the promoters and RBSs. In addition, the functional modularity of the parts and modules can be improved by deploying some insulator sequences between their physical connections. These insulator sequences should be designed to have the least interaction with their adjacent parts. This would isolate the function of individual parts and modules sufficiently to eliminate or reduce the sequence context-dependency effect. As a result, the predictability of assembling genetic parts and modules into functional biological systems would be greatly improved without the need of re-characterisation in a new sequence context. This is illustrated in a recent example (Davis et al., 2010), in which a set of insulated constitutive promoters was designed to act more predictably in different 5'- or 3'-sequence contexts.
Another component pertaining to characterisation could be automation. The thorough characterisation of many parts and modules in various contexts is a large amount of work if done manually, and manual characterisation is not totally reproducible. Therefore, automated, high throughput platform (e.g. liquid handing machines) should be established to speed up the characterisation process and to generate detailed data sheets for each part or module.

2. Optimising the quorum sensing module of the microbial biosensor
The transition region of the engineered cell density-dependent module in this study is not completely within the readily observable range of cell densities. Therefore the next step is to optimise the module to have customisable cell density-dependent characteristics. As discussed at the end of Chapter 6, the current architecture of the module is not optimised. The three architectures described in Chapter 6 could be explored to improve the switching characteristic, as well as the modularity, of this quorum sensing module. By doing this, the module which is an important component of the cell density-dependent microbial biosensor will enable the biosensor to have customised quorum sensing thresholds. As a demonstration for its application in potential biotechnological or environmental areas, an engineered pH-responsive promoter such as the P170 and P1 promoters in *Lactococcus lactis* (Madsen et al., 2005) can be coupled to the second input of the biosensor to detect or specify the acidic condition of the environment.

3. Device diversification with orthogonality
As mentioned at the beginning of Chapter 3, there is a pressing need to enrich and diversify the current device library of synthetic biology to produce many orthogonal modules for constructing larger and more complex biological circuitry. It has been shown in the thesis that the engineering of novel modular and orthogonal devices can be implemented by exploiting new regulation motifs in other bacterial species (e.g. the HrpR/HrpS hetero-regulation motif in *P. syringae*) and utilising alternative regulation modes (e.g. the $\sigma^{54}$-dependent gene transcription). However, an alternative approach could potentially be to use additional distinct orthologs for device diversification. Taking the engineered AND gate as an example, it might be
diversified using orthologous motifs appearing in other *Pseudomonas* species. The *hrp* regulatory system has evolved and diverged among a number of species of the plant *P. syringae* pathovars and other related bacteria such as *P. mendocina* and *P. viridiflava*, but the majority conserves the hetero-regulated HrpR/HrpS motif (Jovanovic et al., 2011). Thus the various orthologs of this motif among these bacterial species can be extracted out to test their functionality and orthogonality as logic AND gates in *E. coli*. As a result, a device library containing many orthogonal logic AND gates might be generated. This method has advantages over random evolution to generate functional orthogonal modules because the evolution space for random mutations of a genetic module is generally too large and thereby limits the likelihood to produce desired mutants.

4. Engineering a distributed bacterial optimiser

One potential research project arising from the current study is to engineer a distributed bacterial optimiser as an upgrade of the cell density-dependent biosensor. This cellular optimiser aims to use living cells as the computing agents to demonstrate a high level biocomputing function – optimising (Amos et al., 2007). Figure 7.1 shows the design principle of this optimiser. Imagine the optimisation task is to seek the optimal solution \((x, y)\) that maximises the function \(f(x, y)\), which is the distribution of the concentration of a specific ligand in a two dimensional space (defined by \(x\) and \(y\)) as shown in Figure 7.1a. Using engineered gene circuits, it may be possible to harness the parallel and distributed power of living cells to accomplish the search process and locate the position with the highest ligand concentration.

The bacterial cells are engineered by integrating multiple functional modules, i.e. the genetic logic circuit module, sensory module, cell-cell communication module and the chemotactic navigation module, to push the limit of gene circuit engineering. As Figure 7.1b shows, the bacterial optimiser will use the logic AND gate to integrate the input signal from the sensory module and the cell density signal from the communication module to drive the chemotaxis module such as *cheZ* output in a *cheZ* knockout *E. coli* strain (Tan et al., 2007; Topp and Gallivan, 2007). Consequently, the cells will be programmed to sense the ligand concentration gradient and migrate.
towards the position with higher ligand concentration and higher cell density. As a result, the engineered cells will aggregate around the point with highest concentration of the target ligand to complete the optimisation task. Because the cells work in parallel, the efficiency of this optimisation process is much higher than serial searching. In addition, three of the four modules in this system have been engineered as part of the work described in this thesis. Therefore, the focus will be the coupling of the chemotactic module to the output of the cell density-dependent biosensor and the fine tuning of these individual modules.

**Figure 7.1 The design of the distributed bacterial optimiser.** The cellular optimiser is engineered to migrate towards the location with the highest ligand concentration, i.e. the optimal solution. **a**, The optimiser’s objective is to locate the position of highest ligand concentration on a surface. **b**, The designed architecture of the optimiser.
# Appendix A: Plasmids

## Table A.1 The plasmid DNAs used in this study.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Notes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAPT110</td>
<td>IPTG inducible P$_{lac}$ promoter expression vector p15A ori, Kan'</td>
<td>Polard et al., 1995</td>
</tr>
<tr>
<td>pBAD18-cm</td>
<td>arabinose inducible P$_{BAD}$ promoter expression vector pBR322 ori, Cm'</td>
<td>Guzman et al., 1995</td>
</tr>
<tr>
<td>pSB4A3</td>
<td>BioBrick vector, pSC101 ori, Amp'</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pSB3K3</td>
<td>BioBrick vector, p15A ori, Kan'</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pSB1A2</td>
<td>BioBrick vector, pMB1 ori, Kan'</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pSB2K3</td>
<td>BioBrick vector, F' ori, P1 lytic ori under P$_{lac}$, Kan'</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pGA15hrpR</td>
<td>GENEART cloning vector harbouring synthesised hrpR, pUC ori, Kan'</td>
<td>From GENEART</td>
</tr>
<tr>
<td>pGA14hrpS</td>
<td>GENEART cloning vector harbouring synthesised hrpS, Col E1 ori, Amp'</td>
<td>From GENEART</td>
</tr>
<tr>
<td>pGA18hrpL</td>
<td>GENEART cloning vector harbouring synthesised hrpL, Col E1 ori, Amp'</td>
<td>From GENEART</td>
</tr>
<tr>
<td>pBW1A2-30G</td>
<td>pSB1A2 harbouring rbs30-gfp-ter</td>
<td>A gift from K. Jensen</td>
</tr>
<tr>
<td>pBW100lac-gfp</td>
<td>pAPT110 encoding rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW101lac-gfp</td>
<td>pAPT110 encoding rbs31-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW102lac-gfp</td>
<td>pAPT110 encoding rbs32-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW103lac-gfp</td>
<td>pAPT110 encoding rbs33-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW104lac-gfp</td>
<td>pAPT110 encoding rbs34-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW105lac-gfp</td>
<td>pAPT110 encoding rbsH-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW200ara-gfp</td>
<td>pBAD18-cm encoding rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW201ara-gfp</td>
<td>pBAD18-cm encoding rbs31-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW202ara-gfp</td>
<td>pBAD18-cm encoding rbs32-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW203ara-gfp</td>
<td>pBAD18-cm encoding rbs33-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW204ara-gfp</td>
<td>pBAD18-cm encoding rbs34-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW205ara-gfp</td>
<td>pBAD18-cm encoding rbsH-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW120260</td>
<td>pSB3K3 harbouring BioBrick Part I20260 (rbs32-gfp-ter under the constitutive BBa_J23101 promoter)</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pBWBBF2620</td>
<td>pSB1A2 harbouring BioBrick Part BBa_F2620 (synthetic AHL inducible P$_{lux}$ promoter)</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pBWF2620</td>
<td>pSB3K3 harbouring BioBrick Part BBa_F2620 BioBrick Part</td>
<td>This study</td>
</tr>
<tr>
<td>pBW300lux-gfp</td>
<td>pBWF2620 encoding rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW301lux-gfp</td>
<td>pBWF2620 encoding rbs31-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW302lux-gfp</td>
<td>pBWF2620 encoding rbs32-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW303lux-gfp</td>
<td>pBWF2620 encoding rbs33-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW304lux-gfp</td>
<td>pBWF2620 encoding rbs34-gfp-ter</td>
<td>This study</td>
</tr>
</tbody>
</table>
### Appendix A: Plasmids

<table>
<thead>
<tr>
<th>Plasmid Name</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBW305lux-gfp</td>
<td>pBWF2620 encoding rbsH-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBWBBP1010</td>
<td>pSB4A3 harbouring BioBrick BBa_P1010 (ccdB gene)</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pBW1A2-30I</td>
<td>pSB1A2 harbouring BioBrick BBa_K081008 (rbs30-luxI-LVA)</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pBW1A2-34I</td>
<td>pSB1A2 harbouring BioBrick BBa_C0261 (rbs34-luxI-LVA)</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pBWBBQ04510</td>
<td>pSB2K3 harbouring BioBrick BBa_Q04510 (rbs34-cl-transfer-P&lt;sub&gt;lux&lt;/sub&gt;)</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pBWBBQ01510</td>
<td>pSB2K3 harbouring BioBrick BBa_Q01510 (rbs31-cl-transfer-P&lt;sub&gt;lux&lt;/sub&gt;)</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pBWBBQ04400</td>
<td>pSB2K3 harbouring BioBrick BBa_Q04400 (rbs34-tetR-transfer-P&lt;sub&gt;lux&lt;/sub&gt;)</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pBWBBQ03400</td>
<td>pSB2K3 harbouring BioBrick BBa_Q03400 (rbs33-tetR-transfer-P&lt;sub&gt;lux&lt;/sub&gt;)</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pBWBBQ01400</td>
<td>pSB2K3 harbouring BioBrick BBa_Q01400 (rbs31-tetR-transfer-P&lt;sub&gt;lux&lt;/sub&gt;)</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pBWBBQ04121</td>
<td>pSB2K3 harbouring BioBrick BBa_Q04121 (rbs34-lacI-transfer-P&lt;sub&gt;lux&lt;/sub&gt;)</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pBWBBQ03121</td>
<td>pSB2K3 harbouring BioBrick BBa_Q03121 (rbs33-lacI-transfer-P&lt;sub&gt;lux&lt;/sub&gt;)</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pBWBBQ01121</td>
<td>pSB2K3 harbouring BioBrick BBa_Q01121 (rbs31-lacI-transfer-P&lt;sub&gt;lux&lt;/sub&gt;)</td>
<td>BioBrick Registry</td>
</tr>
<tr>
<td>pBW110lac-hrpR</td>
<td>pAPT110 encoding rbs30-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW111lac-hrpR</td>
<td>pAPT110 encoding rbs31-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW112lac-hrpR</td>
<td>pAPT110 encoding rbs32-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW113lac-hrpR</td>
<td>pAPT110 encoding rbs33-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW114lac-hrpR</td>
<td>pAPT110 encoding rbs34-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW115lac-hrpR</td>
<td>pAPT110 encoding rbsH-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW310lux-hrpR</td>
<td>pBWF2620 encoding rbs30-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW311lux-hrpR</td>
<td>pBWF2620 encoding rbs31-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW312lux-hrpR</td>
<td>pBWF2620 encoding rbs32-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW313lux-hrpR</td>
<td>pBWF2620 encoding rbs33-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW314lux-hrpR</td>
<td>pBWF2620 encoding rbs34-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW315lux-hrpR</td>
<td>pBWF2620 encoding rbsH-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW210ara-hrpS</td>
<td>pBAD18-cm encoding rbs30-hrpS-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW211ara-hrpS</td>
<td>pBAD18-cm encoding rbs31-hrpS-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW212ara-hrpS</td>
<td>pBAD18-cm encoding rbs32-hrpS-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW213ara-hrpS</td>
<td>pBAD18-cm encoding rbs33-hrpS-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW214ara-hrpS</td>
<td>pBAD18-cm encoding rbs34-hrpS-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW215ara-hrpS</td>
<td>pBAD18-cm encoding rbsH-hrpS-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW145lac-hrpS</td>
<td>pAPT110 encoding rbsH-hrpS-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW343lux-hrpS</td>
<td>pBWF2620 encoding rbs33-hrpS-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW243ara-hrpR</td>
<td>pBAD18-cm encoding rbs33-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW4A3hrpL</td>
<td>pSB4A3 harbouring synthetised hrpL</td>
<td>This study</td>
</tr>
<tr>
<td>pBW400hrpL-gfp</td>
<td>pSB4A3 encoding hrpL-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>Plasmid Code</td>
<td>Description</td>
<td>Remarks</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>pBWLAC32cI</td>
<td>pAPTT110 harboring rbs32-cl-ter-P lamin (XbaI/KpnI)</td>
<td>This study</td>
</tr>
<tr>
<td>pBWLAC33cI</td>
<td>pAPTT110 harboring rbs33-cl-ter-P lamin (XbaI/KpnI)</td>
<td>This study</td>
</tr>
<tr>
<td>pBWLACHcI</td>
<td>pAPTT110 harboring rbsH-cl-ter-P lamin (XbaI/KpnI)</td>
<td>This study</td>
</tr>
<tr>
<td>pBW3K3-32cI</td>
<td>pSB3K3 harboring rbs32-cl-ter-P lamin (XbaI/PstI)</td>
<td>This study</td>
</tr>
<tr>
<td>pBW3K3-33cI</td>
<td>pSB3K3 harboring rbs33-cl-ter-P lamin (XbaI/PstI)</td>
<td>This study</td>
</tr>
<tr>
<td>pBW3K3-HcI</td>
<td>pSB3K3 harboring rbsH-cl-ter-P lamin (XbaI/PstI)</td>
<td>This study</td>
</tr>
<tr>
<td>pBWLAC31cIG</td>
<td>pAPTT110 (EcoRI/KpnI) harboring rbs31-cl-ter-P lamin-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBWLAC33cIG</td>
<td>pAPTT110 (EcoRI/KpnI) harboring rbs33-cl-ter-P lamin-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBWLAC34cIG</td>
<td>pAPTT110 (EcoRI/KpnI) harboring rbs34-cl-ter-P lamin-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBWLACHcIG</td>
<td>pAPTT110 (EcoRI/KpnI) harboring rbsH-cl-ter-P lamin-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW121lac-clgfp</td>
<td>pAPTT10 (XbaI/KpnI) encoding rbs31-cl-ter-P lamin-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW122lac-clgfp</td>
<td>pAPTT10 (XbaI/KpnI) encoding rbs32-cl-ter-P lamin-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW123lac-clgfp</td>
<td>pAPTT10 (XbaI/KpnI) encoding rbs33-cl-ter-P lamin-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW124lac-clgfp</td>
<td>pAPTT10 (XbaI/KpnI) encoding rbs34-cl-ter-P lamin-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW125lac-clgfp</td>
<td>pAPTT10 (XbaI/KpnI) encoding rbsH-cl-ter-P lamin-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW412hrpL-clgfp</td>
<td>pSB4A3 encoding hrpL-rbs32-cl-ter-P lamin-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW414hrpL-clgfp</td>
<td>pSB4A3 encoding hrpL-rbs34-cl-ter-P lamin-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW121lac-tetRgfp</td>
<td>pAPTT10 (EcoRI/KpnI) harboring rbs31-tetR-ter-P tet-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW123lac-tetRgfp</td>
<td>pAPTT10 (EcoRI/KpnI) harboring rbs33-tetR-ter-P tet-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW124lac-tetRgfp</td>
<td>pAPTT10 (EcoRI/KpnI) harboring rbs34-tetR-ter-P tet-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW125lac-tetRgfp</td>
<td>pAPTT10 (EcoRI/KpnI) harboring rbsH-tetR-ter-P tet-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW131lac-lacIgfp</td>
<td>pSB1A2 harbouring rbs31-lacI-ter-P lux-LVA</td>
<td>This study</td>
</tr>
</tbody>
</table>
### Appendix A: Plasmids

<table>
<thead>
<tr>
<th>Plasmid Code</th>
<th>Description</th>
<th>Study Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBW1A2-32I</td>
<td>pSB1A2 harbouring rbs32-luxI-LVA</td>
<td>This study</td>
</tr>
<tr>
<td>pBW1A2-33I</td>
<td>pSB1A2 harbouring rbs33-luxI-LVA</td>
<td>This study</td>
</tr>
<tr>
<td>pBW1A2-HI</td>
<td>pSB1A2 harbouring rbsH-luxI-LVA</td>
<td>This study</td>
</tr>
<tr>
<td>pBW1A2-30I33G</td>
<td>pSB1A2 harbouring rbs30-luxI-LVA-rbs33-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW1A2-32I33G</td>
<td>pSB1A2 harbouring rbs32-luxI-LVA-rbs33-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW1A2-33I33G</td>
<td>pSB1A2 harbouring rbs33-luxI-LVA-rbs33-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW1A2-34I33G</td>
<td>pSB1A2 harbouring rbs34-luxI-LVA-rbs33-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW1A2-HI33G</td>
<td>pSB1A2 harbouring rbsH-luxI-LVA-rbs33-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW1A2-30I30G</td>
<td>pSB1A2 harbouring rbs301-luxI-LVA-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW1A2-31I30G</td>
<td>pSB1A2 harbouring rbs31-luxI-LVA-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW1A2-32I30G</td>
<td>pSB1A2 harbouring rbs32-luxI-LVA-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW1A2-33I30G</td>
<td>pSB1A2 harbouring rbs33-luxI-LVA-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW1A2-34I30G</td>
<td>pSB1A2 harbouring rbs34-luxI-LVA-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW1A2-HI30G</td>
<td>pSB1A2 harbouring rbsH-luxI-LVA-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW503luxI-gfp</td>
<td>pBWF2620 encoding rbs30-luxI-LVA-rbs33-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW513luxI-gfp</td>
<td>pBWF2620 encoding rbs31-luxI-LVA-rbs33-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW523luxI-gfp</td>
<td>pBWF2620 encoding rbs32-luxI-LVA-rbs33-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW533luxI-gfp</td>
<td>pBWF2620 encoding rbs33-luxI-LVA-rbs33-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW543luxI-gfp</td>
<td>pBWF2620 encoding rbs34-luxI-LVA-rbs33-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW553luxI-gfp</td>
<td>pBWF2620 encoding rbsH-luxI-LVA-rbs33-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW500luxI-gfp</td>
<td>pBWF2620 encoding rbs30-luxI-LVA-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW510luxI-gfp</td>
<td>pBWF2620 encoding rbs31-luxI-LVA-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW520luxI-gfp</td>
<td>pBWF2620 encoding rbs32-luxI-LVA-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW530luxI-gfp</td>
<td>pBWF2620 encoding rbs33-luxI-LVA-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW540luxI-gfp</td>
<td>pBWF2620 encoding rbs34-luxI-LVA-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW550luxI-gfp</td>
<td>pBWF2620 encoding rbsH-luxI-LVA-rbs30-gfp-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW630luxI-hrpR</td>
<td>pBWF2620 encoding rbs33-luxI-LVA-rbs30-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW631luxI-hrpR</td>
<td>pBWF2620 encoding rbs33-luxI-LVA-rbs31-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW632luxI-hrpR</td>
<td>pBWF2620 encoding rbs33-luxI-LVA-rbs32-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW633luxI-hrpR</td>
<td>pBWF2620 encoding rbs33-luxI-LVA-rbs33-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW634luxI-hrpR</td>
<td>pBWF2620 encoding rbs33-luxI-LVA-rbs34-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW635luxI-hrpR</td>
<td>pBWF2620 encoding rbs33-luxI-LVA-rbsH-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW620luxI-hrpR</td>
<td>pBWF2620 encoding rbs32-luxI-LVA-rbs30-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW621luxI-hrpR</td>
<td>pBWF2620 encoding rbs32-luxI-LVA-rbs31-hrpR-ter</td>
<td>This study</td>
</tr>
<tr>
<td>pBW622luxI-hrpR</td>
<td>pBWF2620 encoding rbs32-luxI-LVA-rbs32-hrpR-ter</td>
<td>This study</td>
</tr>
</tbody>
</table>
### Appendix B: Primers

#### Table B.1 The oligo DNAs (primers) used in this study.

<table>
<thead>
<tr>
<th>Primer (Set)</th>
<th>Sequence (5’ – 3’)</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>gfp-C F</td>
<td>CGGAATTCGGCGCCGCTCTAGAG</td>
<td>PCR gfp to be cloned into the three expression vectors</td>
</tr>
<tr>
<td>gfp-C R</td>
<td>GGTTACCTGCGAGCGCTACAGAG</td>
<td></td>
</tr>
<tr>
<td>gfp-S F</td>
<td>CAAGAGTGGCCATGCAGGAAG</td>
<td>Middle sequencing primers for gfp</td>
</tr>
<tr>
<td>gfp-S R</td>
<td>AATGGTTGCTGGTAAAGGG</td>
<td></td>
</tr>
<tr>
<td>pBAD18-cm F</td>
<td>ATGCCCATAGCAATTTTTATCC</td>
<td>pBAD18-cm vector sequencing primer set for gfp</td>
</tr>
<tr>
<td>pBAD18-cm R</td>
<td>GATTTAATCTGGATACAGG</td>
<td></td>
</tr>
<tr>
<td>BioBrick F</td>
<td>TGGCCACCTGACGGCTCTAGGA</td>
<td>Sequencing or analyzing parts in BioBrick vectors for gfp</td>
</tr>
<tr>
<td>BioBrick R</td>
<td>ATTACCGCTTTGAGTGAGGC</td>
<td></td>
</tr>
<tr>
<td>pAPT110 F</td>
<td>GGCTTTACACTTATATCTTC</td>
<td>pAPT110 vector sequencing primer set</td>
</tr>
<tr>
<td>pAPT110 R</td>
<td>GAAATCCTTGGGATGGATACAGG</td>
<td></td>
</tr>
<tr>
<td>hrpR F</td>
<td>GTGGATCTGGGCTGATTCAGG</td>
<td>Middle sequencing primers for hrpR</td>
</tr>
<tr>
<td>hrpR R</td>
<td>CTTCAATTTCGCGATGGTGAGC</td>
<td></td>
</tr>
<tr>
<td>hrpS F</td>
<td>CGACGATCAAGCTGGATATAC</td>
<td>Middle sequencing primers for hrpS</td>
</tr>
<tr>
<td>hrpS R</td>
<td>CAACACTCGGGACGGGAAAC</td>
<td></td>
</tr>
<tr>
<td>RBS30_gfp F</td>
<td>CGTGCTAGATTAAGGAAGAGGAATACTAG</td>
<td>PCR gfp with RBS30 and relevant restriction sites</td>
</tr>
<tr>
<td>RBS31_gfp F</td>
<td>CGTCTAGATGCAACAGGAAAGTACTAGATG</td>
<td>PCR gfp with RBS31 and relevant restriction sites</td>
</tr>
<tr>
<td>RBS32_gfp F</td>
<td>CGTCTAGATGCAACAGGAAACCTACTAGATG</td>
<td>PCR gfp with RBS32 and relevant restriction sites</td>
</tr>
<tr>
<td>RBS33_gfp F</td>
<td>CGTCTAGATGCAACAGGACTACTAGATG</td>
<td>PCR gfp with RBS33 and relevant restriction sites</td>
</tr>
<tr>
<td>RBS34_gfp F</td>
<td>CGTCTAGAGAAGAGAGAAATACTAGATG</td>
<td>PCR gfp with RBS34 and relevant restriction sites</td>
</tr>
<tr>
<td>RBSH_gfp F</td>
<td>CCGAATCAGTGGAGATAGCATG</td>
<td>PCR gfp with RBSH and relevant restriction sites</td>
</tr>
<tr>
<td>RBS30-H_GRS R</td>
<td>GGGGTTACCTGCGAGCGCTACTAGATATATAAAC</td>
<td>Reverse primer for gfp, hrpR, hrpS with the 6 RBSs of various strengths</td>
</tr>
<tr>
<td>RBS30_hrpR F</td>
<td>CGTCTAGAGATTAAGGAAGAGGAATACTAG</td>
<td>PCR hrpR with RBS30 and relevant restriction sites</td>
</tr>
<tr>
<td>RBS31_hrpR F</td>
<td>CGTCTAGAGATGCAACAGGAAACCTACTAG</td>
<td>PCR hrpR with RBS31 and relevant restriction sites</td>
</tr>
<tr>
<td>RBS32_hrpR F</td>
<td>CGTCTAGAGATGCAACAGGAAAGTACTAG</td>
<td>PCR hrpR with RBS32 and relevant restriction sites</td>
</tr>
<tr>
<td>RBS33_hrpR F</td>
<td>CGTCTAGAGATGCAACAGGACTACTAG</td>
<td>PCR hrpR with RBS33 and relevant restriction sites</td>
</tr>
</tbody>
</table>
Appendix B: Primers

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequences</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBS34_hrpR F</td>
<td>CGTCTAGAGAAAGGAGGAAATCTAG</td>
<td>PCR hrpR with RBS34 and relevant restriction sites</td>
</tr>
<tr>
<td></td>
<td>ATGAGTACAGCCATCGATAGGG</td>
<td></td>
</tr>
<tr>
<td>RBSH_hrpR  F</td>
<td>CGTCTAGAGGAGGATATACC</td>
<td>PCR hrpR with RBSH and relevant restriction sites</td>
</tr>
<tr>
<td></td>
<td>ATGAGTACAGCCATCGATAGGTC</td>
<td></td>
</tr>
<tr>
<td>RBS30_hrpS  F</td>
<td>CGTCTAGAGATTTAAGAGGAGAATCTAG</td>
<td>PCR hrpS with RBS30 and relevant restriction sites</td>
</tr>
<tr>
<td></td>
<td>ATGAGTCTAGTAGAAAGGTTTGG</td>
<td></td>
</tr>
<tr>
<td>RBS31_hrpS  F</td>
<td>CGTCTAGAGTACACAGAGAATCTAG</td>
<td>PCR hrpS with RBS31 and relevant restriction sites</td>
</tr>
<tr>
<td></td>
<td>ATGAGTCTAGTAGAAAGGTTTGG</td>
<td></td>
</tr>
<tr>
<td>RBS32_hrpS  F</td>
<td>CGTCTAGAGTACACAGAGAATCTAG</td>
<td>PCR hrpS with RBS32 and relevant restriction sites</td>
</tr>
<tr>
<td></td>
<td>ATGAGTCTAGTAGAAAGGTTTGG</td>
<td></td>
</tr>
<tr>
<td>RBS33_hrpS  F</td>
<td>CGTCTAGAGTACACAGAGAATCTAG</td>
<td>PCR hrpS with RBS33 and relevant restriction sites</td>
</tr>
<tr>
<td></td>
<td>ATGAGTCTAGTAGAAAGGTTTGG</td>
<td></td>
</tr>
<tr>
<td>RBS34_hrpS  F</td>
<td>CGTCTAGAGGAGGAGGAAATCTAG</td>
<td>PCR hrpS with RBS34 and relevant restriction sites</td>
</tr>
<tr>
<td></td>
<td>ATGAGTCTAGTAGAAAGGTTTGG</td>
<td></td>
</tr>
<tr>
<td>RBSH_hrpS  F</td>
<td>CGTCTAGAGGAGGATATACC</td>
<td>PCR hrpS with RBSH and relevant restriction sites</td>
</tr>
<tr>
<td></td>
<td>ATGAGTCTAGTAGAAAGGTTTGG</td>
<td></td>
</tr>
<tr>
<td>RBS31_luxI  F</td>
<td>CGTCTAGAGTACACAGAGAATCTAG</td>
<td>PCR luxI with RBS31 and relevant restriction sites</td>
</tr>
<tr>
<td></td>
<td>ATGACTATAATGATAAAAAAATCG</td>
<td></td>
</tr>
<tr>
<td>RBS32_luxI  F</td>
<td>CGTCTAGAGTACACAGAGAATCTAG</td>
<td>PCR luxI with RBS32 and relevant restriction sites</td>
</tr>
<tr>
<td></td>
<td>ATGACTATAATGATAAAAAAATCG</td>
<td></td>
</tr>
<tr>
<td>RBS33_luxI  F</td>
<td>CGTCTAGAGTACACAGAGAATCTAG</td>
<td>PCR luxI with RBS33 and relevant restriction sites</td>
</tr>
<tr>
<td></td>
<td>ATGACTATAATGATAAAAAAATCG</td>
<td></td>
</tr>
<tr>
<td>RBSH_luxI  F</td>
<td>CGTCTAGAGGAGGATATACC</td>
<td>PCR luxI with RBSH and relevant restriction sites</td>
</tr>
<tr>
<td></td>
<td>ATGACTATAATGATAAAAAAATCG</td>
<td></td>
</tr>
<tr>
<td>RBS31-H_luxI  R</td>
<td>GGTCACCTGCAGCGCGCGCTACTAGTA</td>
<td>Reverse primer for PCR luxI with the 4 RBSs (RBS31-H)</td>
</tr>
<tr>
<td></td>
<td>TTATTAAGCTACTAAAAAGCGTAG</td>
<td></td>
</tr>
<tr>
<td>RBS32_cI  F</td>
<td>CGTCTAGAGTACACAGAGAATCTAG</td>
<td>PCR cI with RBS32 and relevant restriction sites</td>
</tr>
<tr>
<td></td>
<td>ATGAGCACAACAAAAAACAAACC</td>
<td></td>
</tr>
<tr>
<td>RBS33_cI  F</td>
<td>CGTCTAGAGTACACAGAGAATCTAG</td>
<td>PCR cI with RBS33 and relevant restriction sites</td>
</tr>
<tr>
<td></td>
<td>ATGAGCACAACAAAAAACAAACC</td>
<td></td>
</tr>
<tr>
<td>RBSH_cI  F</td>
<td>CGTCTAGAGGAGGATATACC</td>
<td>PCR cI with RBSH and relevant restriction sites</td>
</tr>
<tr>
<td></td>
<td>ATGAGCACAACAAAAAACAAACC</td>
<td></td>
</tr>
<tr>
<td>RBS31-H_cI  R</td>
<td>GGTACCCTGCAGCGCGCGCTACTAGTA</td>
<td>Reverse primer for PCR cI with the 3 various RBSs</td>
</tr>
<tr>
<td></td>
<td>GCAACCATTATCACCACCACC</td>
<td></td>
</tr>
</tbody>
</table>
Appendix C: Essential DNA Sequences

The sequences of the essential gene elements used in this study are listed below.

hrpL promoter sequence:
>Part-only sequence, 208 bp
GCCGGATATATGTCCCGCTAGGTGGTCAGTGCACTCCATCGGAGGATATATGGGCCTAGGAGATTCTTTT
GTGCCAAACTTTGTGTGCGAAAAATTAAAGGGTATTGGCTATCAACTTCGTCAGTTTTA
AAATATCTTTTATAATAATCATCAGTTATTTCTTATATTGTATGTTATAAGCTTGGAATTTAATCGCTATAGGGCTTGAGC

hrpR gene in BioBrick format:
>Part-only sequence, 948 bp
ATGAGTACAGCGATCAGATGAGGTACCGTCCGAGAGTGTTGGGGCGTAACTGCATTATCAGCGGGTCATCAA
ATTGCAATGCATACGCACTCCATCGGACGAGTGGTCAGCTGCTATCGGGTGAGACCGGCACCGGCAAAGACGT
CGTGGAGTCTCTGCACAGTTTTTCTGCCGCTACAGTGGGCTACCTGGCGCTTGACCGGCAAGTACCG
TTATTCGCGCTGCGGACGATGTTGATGTTGAAATCGAGAGGTCGACGAGATTTGGCAGATCGCTGGAGGCTG

hrpS gene in BioBrick format:
>Part-only sequence, 912 bp
ATGAGTCTTGATGAAAGGTTTGAGGATGATCTGGACGAGGAGCGGGTTCCGAATCTGGGGATAGTTGCC
GAAAGTTTCTCGCACTTGACAGATGTCAGGCCGCAAAGGGCGCCTGGTGGCGATGAATTGCGCGGCCATTCCG
CTGAGTCCCTCGCCGAGAGCGAGTTATTCGGCGTGGTCAGCGGTGCCTACACCGGCGCTGATCGCTCCAGA
GTCGGTTATAGGGCTGGCCAGGCGGCGGACGTCGAGTGGTCAGTGCATACGATACGCTGCCG

gfpmut3b gene in BioBrick format:
>BBa_E0040 Part-only sequence, 720 bp
ATCGCTATGAGGAGATTGTTTGGAGATGATCTGGGAGGACGCGGTCCGGAATCTGGGGATAGTTGCC
GCAAAGTTTCTCGCACTTGACAGATGTCAGGCCGCAAAGGGCGCCTGGTGGCGATGAATTGCGCGGCCATTCCG
CTGAGTCCCTCGCCGAGAGCGAGTTATTCGGCGTGGTCAGCGGTGCCTACACCGGCGCTGATCGCTCCAGA
GTCGGTTATAGGGCTGGCCAGGCGGCGGACGTCGAGTGGTCAGTGCATACGATACGCTGCCG
Appendix C: Essential DNA Sequences

**tetR/Ptet Inverter:**
>Part-only sequence, 884 bp, *tetR-LVA*

**lacI/PLac Inverter:**
>Part-only sequence, 1352 bp, *lacI-LVA*

**cI/Plam Inverter:**
>Part-only sequence, 969 bp, *cI-LVA*
Appendix C: Essential DNA Sequences

The AHL-responsive synthetic P_lux promoter (BBa_F2620):

>`BBa_F2620 Part-only sequence, 1061 bp
TCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACTACTAGAGAAAGAGG
AGAAATACTAGATGAAAAACATAAATAATGCGGACAGGACTACAGAATAATTAATAAAATTATGATGTT
GAAAGGCAATATTTATAGAAGCTCTACTTTGATTTTATATGAATTTTTATACACTGTTGGTTGTTTGTC
GTAATCTGCTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTCT
GGTTTATATACTAGAGACCTGTAGGATCGTACAGGTTTACGCAAGAAAAGTTTGTTATAGTCGAATAAA

luxI gene – the synthase of AHL molecule:

>`BBa_C0061 Part-only sequence, 618 bp, luxI-LVA
ATGACTATAATGATAAAAAAATCGGATTTTTTGGCAATTCCATCGGAGGAGTATAAAGGTATTCTAAGT
CTTCGTTATCAAGTGTTTAAGCAAAGACTTGAGTGGGACTTAGTTGTAGAAAATAACCTTGAATCAGAT
GAGTATGATAACTCAAATGCAGAATATATTGCTTGTGATGATACTGAAAATGTAAGTGGATGCTGG
CGTTTATTACCTACAACAGGTGATTATATGCTGAAAAGTGTTTTTCCTGAATTGCTTGGTCAACAGAGT
GCTCCCAAAGATCCTAATATAGTCGAATTAAGTCGTTTTGCTGTAGGTAAAAATAGCTCAAAGATAAAT
AACCTCTGCTAGTGAAATTACAATGAAACTATTTGAAGCTATATATAAACACGCTGTTAGTCAAGGTATT
ACAGATAATGTAACAGTAAACTCAACAGCAACAGAGGCTAATTATAGGCTGATGGCAGTTTACGCAAAAC
GCTTCTTAAATAACGACTGTTTATAGTGCATTTATAAT

The BioBrick double terminator BBa_B0015:

>`BBa_B0015 Part-only sequence, 129 bp
CCAGGCATCAAATAAAAACGCAAAGGCTCGATCGTAAATATTACGATATTTTTATCTGTTGTTTTATGCT
GGTGACGCTCTACTAGAGTCACACTGCGCTACCTCTGGTTGTTTATA
References


References


