Towards Light Based Dynamic Control of Synthetic Biological Systems

Marios Tomazou

Imperial College London
Department of Bioengineering

Submission for the Degree of Doctor of Philosophy (PhD)

September 2014
Declarations

'I hereby confirm, that all the material in this document are the results of my work, with the exception of data, quotes or figures from other published or unpublished sources that are clearly labelled as such and referenced accordingly, both within the text and in the figure captions.'

Marios Tomazou
September, 2014

'The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.'
Acknowledgments

I would like to thank my supervisors Dr Karen Polizzi, Dr Guy-Bart Stan and Prof Mauricio Barahona for giving me the opportunity to work with you and complete this work. Your help and guidance, was invaluable and your dedication on securing funding for me was most certainly a core driving force behind this work. I really have no words to express how appreciated this is. Karen, I couldn't have done any meaningful work, if it wasn't for your endless patience with me, all the discussions and brainstorming throughout the project's duration and the fact that you turned me from an undergraduate student, with only basic lab practical experience, to become a self sufficient experimentalist. Guy, in my opinion, you are by far the most passionate and dedicated academic for bringing together biology and engineering, something that kept me interested and willing to learn, as much as possible from you, about modelling and other theoretical approaches for biology. Your ideas (often wild), energy and the fact that you were always there when needed, kept me motivated and productive on all aspects of this project. Prof Barahona, your ability to see through the trivial details and all the problems that occurred, at every step of the project, only to identify what was really important and relevant, is something that kept me on track. I don't know if your claim that you 'don't know biology' is true, but your catalytic input in both parts of this work, was truly inspiring.

A very special thanks goes to my amazing friends and fellow researches Dr James Arpino, Dr James Chappell, Dr Jennifer Samson, Dr Jordan Ang, Dr Christ Hirst, Dr Arturo Casini and Dr Diego Oyarzun that helped me survive and overcome everything from real life problems to the stubborn cloning that just didn't want to work. You are undeniably the best group of people that I had the luck to work (and have a pint) with and your presence was the one thing that kept getting me through the jungle *i.e.* London tube, every single day. Your contribution on an every day basis regarding the lab and theoretical aspects of the work is just too much for mentioning here in detail.

I would also like to thank specifically, Dr Travis Bayer for the idea of period and amplitude decoupling, Dr Eric Davidson for the help with the light receptors, Dr Tom Ellis, Dr Ben Blount and Dr John Heap for all the useful discussions on various topics, Dr Jensen Kirtsen for ensuring the smooth operation of the lab, Prof Paul Freemont for helping me on the funding issue and everyone that I had the luck to work with in CSynBi.

Last but most important, is a warm thank you to my parents, my grandparents and my brother. I have lost count of how many sacrifices you have done for me, in order to be here and complete this work. For providing for me the best possible education and funding for my studies, supporting me at every step of the way and the faith you showed in me, I couldn't be more grateful. I feel that you deserve a sincere apology from me for not seeing you or calling you, as much as I wanted and owed you through these years, because I was often putting less important things as a priority. My thesis is dedicated to you,
Abstract

For the field of synthetic biology, the adaptation of principles, from the well established traditional engineering disciplines, like mechanical and electrical engineering, in order to realise complex synthetic biological circuits, is an intriguing prospect. These principles can enable a forward engineering, rational design and implementation approach, where a system's properties can be predicted or designed in silico followed by the manufacturing of the in vivo system, that can be tested, used or redesigned in the most efficient possible way. Achieving control over these circuits, is one of the important topics of the field, for these applications to become robust and render useful functions applicable to energy, medicine, pharmaceuticals and agriculture industries. In this work, I attempt to explore light, as a promising control 'dial' for synthetic circuitry. Light is fast, economic compared to chemicals, it can be interfaced with electronics, it is reversible in its effect and can be applied at a fine spatio-temporal resolution. These characteristics, are absent from the classically used chemical inducers, meaning that light, can open new possibilities for the user to control synthetic systems, or even facilitate the cell to cell communication, within population based networks.

This work, is a contribution towards harnessing the advantages of light, for achieving control over synthetic circuits. More specifically, I start with the detailed theoretical and experimental study of the Cph8 two component system, a synthetic chimeric receptor which is responsive to red light. This is done, in order to develop a sufficient theoretical understanding of it, through detailed mechanistic modelling, in order to connect the specific system with the toggle switch and the dual feedback oscillator, in an optimal way and achieve control of these devices through light. The developed model, was able to highlight the main aspects and mechanisms inherent to its structure, describe most of the observations from the experimental system, to also make quantitative predictions. The second part of this work, was the development of novel promoters, that can be regulated by a commonly used transcription factor, such as LacI, but also, light responsive regulators like OmpR and CcaR. This yielded a direct way to integrate light and chemical inputs, into a single output, while the dual regulation, allowed to connect and modulate the toggle switch without the need of additional transcription factors. The latter, a light tuneable toggle switch, showed indications that it can function as a memory controller that can be reset by light. Finally, I show the design and modelling of a light tuneable dual feedback oscillator, where light of one wavelength can be used to tune the amplitude, while another wavelength can tune the period. The developed models and synthetic circuits are expected to contribute towards implementing finely tuned and controlled synthetic circuits through light.
# Table of Contents

Abstract 4  
Table of Figures 8  
List of Tables 9  

Chapter 1. Introduction 11  

1.1. Synthetic Biology 11  
1.1.1. Definition and a Brief History 11  
1.1.2. Adapting Engineering Principles for Synthetic Biology 13  
1.1.3. The Engineering Process for Synthetic Biology 15  
1.2. Biological 'Knobs', Control and Induction of Synthetic Genetic Networks 22  
1.2.1. Regulation Biological 'Dials' for Bacterial Systems 22  
1.2.2. Confounding Factors and Control Aspects 25  
1.2.3. User Input - Induction Approaches 28  
1.3. Optogenetic Control of Gene Expression in bacteria 31  
1.3.1. Two Component System Light Receptors 31  
1.3.2. Alternative Light Based Gene Regulation Mechanisms 32  
1.4. Motivation and Objectives 34  

Chapter 2. Materials and Methods 39  

2.1. Experimental Methods 39  
2.1.1. Cloning and DNA Assembly 39  
2.1.2. Experimental Assays 44  
2.2. Theoretical and Computational Methods 50  
2.2.1. Computational Modelling 50  
2.2.2. Data Analysis and Illustrative Software 53  
2.3. Materials 54  
2.3.1. Light Induction Source 54  
2.3.2. Strains, Genetic Material and Solutions 54  

Chapter 3. The Light Receptor Characterisation 62  

3.1. The Cph8 Two Component System Model 62  
3.1.1. System Overview 62  
3.1.2. Biochemical Model 64  
3.1.3. Model Reduction 70
3.1.4. Model Simulations

3.2. The Experimental System
   3.2.1. The Light Source
   3.2.2. The Light Receptor Construct
   3.2.3. Cph8 Experimental Characterisation Results
   3.2.4. Model Fitting - Parameter Estimation
   3.2.5. The Multichromatic System Characterisation - CcaS / Cph8

Chapter 4. Light and Chemical Input Promoters
   4.1. Background
   4.2. The Light-Chemical Induced Promoter Design and Model
      4.2.1. Design
      4.2.2. The Compact TCS Model and Extension for LacI Binding
   4.3. Experimental Implementation
      4.3.1. Dual Input Promoter Experimental Assembly
      4.3.2. Experimental Characterization
      4.3.3. Model Refinement and Data Fitting

Chapter 5. A Light Tuneable Toggle Switch
   5.1. Background
   5.2. Light Tuneable Toggle Switch Model
      5.2.1. Two-State Phenomenological Model of the Modified Light Toggle Switch (LTS)
      5.2.2. Mechanistic Model Derivation
   5.3. Experimental Implementation
      5.3.1. Assembly of the Light Tuneable Genetic Toggle Switch
      5.3.2. Characterisation of the pRG021-C8 dual Reporter Toggle Switch
      5.3.3. Characterisation of the pIKEPlo Single Reporter Light Toggle Switch

Chapter 6. Theoretical Design, A Light Tuneable Oscillator
   6.1. Background
      6.1.1. Model Implementation and Preliminary System Design
      6.1.2. Genetic Construct and Plasmid Maps Design
   6.2. Model Refinement and Modifications
      6.2.1. DNA Loop Formation by LacI
      6.2.2. Degradation Mechanism and GFP SsrA Tag Feedback Effect
   6.3. System Sensitivity Assessment, re-Design and Variability Profile of the Model
      6.3.1. Amplitude and Period Sensitivity on Light
6.3.2. System re-Design for Independent Amplitude and Period Tuning 194
6.3.3. On the Robustness of Oscillations - Stochastic Model 197

Chapter 7. Conclusions and Discussion 211

7.1. Conclusion and Discussion on the Scope of the Project 211
7.2. Results Discussion for Each Module 212
   7.2.1. Module 1: Light Receptor Characterisation 212
   7.2.2. Module 3: Light Tuneable Toggle Switch 215
   7.2.3. Module 3: Light Tuneable Toggle Switch 218
   7.2.4. Module 4: Light Tuneable Oscillator 221
7.3. Implications and Future Prospects 224
   7.3.1. Module 1: Light Receptor Characterisation 224
   7.3.2. Module 3: Light Tuneable Toggle Switch 226
   7.3.3. Module 4: Light Tuneable Oscillator 227
   7.3.4. Other Future Prospects 228

8. Bibliography 229

9. Supplementary Information 243
   8.1. Appendix A: Light Receptors 243
      8.1.1. Cph8 TCS Reduced Computational Model 243
      8.1.2. Parameter Fitting Supplementary Information 247
   8.2. Appendix B: Dual Input Model and LTS 248
      8.2.1. Dual Input Promoter Model 248
      8.2.2. LTS Model 252
   8.2. Appendix C: Light Tuneable Oscillator 254
      8.2.1. LTO Model 254
## Table of Figures

<table>
<thead>
<tr>
<th>Figure Number</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1.1</td>
<td>The abstraction hierarchy</td>
<td>13</td>
</tr>
<tr>
<td>Figure 1.1.2</td>
<td>The engineering cycle</td>
<td>17</td>
</tr>
<tr>
<td>Figure 1.1.1</td>
<td>Biological 'Dials', effectors and regulation of gene expression in bacteria.</td>
<td>25</td>
</tr>
<tr>
<td>Figure 1.1.2</td>
<td>The TCS light receptors</td>
<td>32</td>
</tr>
<tr>
<td>Figure 1.1.2</td>
<td>Other optogenetic regulation strategies</td>
<td>34</td>
</tr>
<tr>
<td>Figure 2.1.1</td>
<td>Microplate layouts and setup used for the light toggle switch assays discussed in chapter 5.</td>
<td>50</td>
</tr>
<tr>
<td>Figure 2.1.2</td>
<td>Circuit diagram of the LED devices</td>
<td>54</td>
</tr>
<tr>
<td>Figure 2.1.1</td>
<td>Incubator setup for fixed and variable light intensity</td>
<td>55</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Diagrammatic overview of the galloping model</td>
<td>62</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>The biosynthesis pathway for Phycocyanobilin in <em>Synechocystis</em> sp.</td>
<td>62</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>The network diagram of Cph8 receptor model</td>
<td>64</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Simulations for the mass-action, reduced and stochastic model</td>
<td>73</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Time Course run and Transfer function for varying the Gene Copy number of gPCB</td>
<td>74</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Time Course run for OmpR-p and its transfer function</td>
<td>79</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Transfer function for L versus light</td>
<td>82</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Retroactive effect of the PompC promoter copy number on the reporter output</td>
<td>85</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Diagram of the reduced model</td>
<td>89</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Time course simulation with induction at t=700</td>
<td>89</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>The deviation in the transition dynamics between the reduced and mass-action model</td>
<td>90</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Sensitivity analysis results</td>
<td>91</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Time course simulation of the full Multichromatic control system</td>
<td>94</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Transfer function of red and green light at steady state</td>
<td>94</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>LED the power output</td>
<td>96</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Plasmid maps of the pCph8 chimeric light receptor</td>
<td>98</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Phenotypic characterisation of Cph8</td>
<td>99</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Steady state characterisation for LacZ expression</td>
<td>100</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Time course run for pL-Cph8FD</td>
<td>104</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Determination of the range of Heme levels</td>
<td>106</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Relative fluorescent levels of Oxalic acid treated samples</td>
<td>106</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Diagrammatic overview of the parameter fitting</td>
<td>108</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Data fits, estimated value boxplots and residuals plot</td>
<td>109</td>
</tr>
<tr>
<td>Figure 3.1.1</td>
<td>Plasmid maps for the multichromatic control system</td>
<td>111</td>
</tr>
<tr>
<td>Figure 4.1.1</td>
<td>Repression mechanism of LacI</td>
<td>116</td>
</tr>
<tr>
<td>Figure 4.1.1</td>
<td>Examples of Signal transduction strategies</td>
<td>116</td>
</tr>
<tr>
<td>Figure 4.1.1</td>
<td>Chemical and light inducible promoter design</td>
<td>119</td>
</tr>
<tr>
<td>Figure 4.1.1</td>
<td>Re-fitted compact model</td>
<td>122</td>
</tr>
<tr>
<td>Figure 4.1.1</td>
<td>Construction of the pl-PloPCB plasmid</td>
<td>125</td>
</tr>
<tr>
<td>Figure 4.1.1</td>
<td>Construction of the pl-PlcPCB plasmid</td>
<td>126</td>
</tr>
<tr>
<td>Figure 4.1.1</td>
<td>Plac/ompR IPTG induction</td>
<td>128</td>
</tr>
<tr>
<td>Figure 4.1.1</td>
<td>Model and experimental data comparison</td>
<td>131</td>
</tr>
<tr>
<td>Figure 4.1.1</td>
<td>Experimental Time course runs for Plac/ompR</td>
<td>132</td>
</tr>
<tr>
<td>Figure 4.1.1</td>
<td>Fitted model on the time course data for both promoters</td>
<td>134</td>
</tr>
<tr>
<td>Figure 5.2.1</td>
<td>The genetic toggle switch model</td>
<td>138</td>
</tr>
<tr>
<td>Figure 5.2.1</td>
<td>The red light tuneable toggle switch</td>
<td>139</td>
</tr>
<tr>
<td>Figure 5.2.2</td>
<td>Model simulations for the LTS</td>
<td>143</td>
</tr>
<tr>
<td>Figure 5.2.1</td>
<td>Network diagram of the single reporter Light tuneable Toggle Switch (LTS)</td>
<td>146</td>
</tr>
<tr>
<td>Figure 5.2.2</td>
<td>Construction of the dual reporter version of the light inducible toggle</td>
<td>147</td>
</tr>
<tr>
<td>Figure 5.2.3</td>
<td>Construction of the single reporter version of the light inducible toggle</td>
<td>148</td>
</tr>
<tr>
<td>Figure 5.2.3</td>
<td>Preliminary screening</td>
<td>150</td>
</tr>
<tr>
<td>Figure 5.2.5</td>
<td>Induction plate assay</td>
<td>152</td>
</tr>
</tbody>
</table>
Figure 5.3.6. Growth profile of the strains transformed with the dual reporter LTS
Figure 5.3.7. Growth assays for the single reporter LTS
Figure 5.3.8. LTS response on chemical inducers under light or dark conditions.
Figure 5.3.9. Surface plot of two dimensional induction by light and IPTG
Figure 5.3.10. Light pulse assay
Figure 5.3.11. Time course and stability assay
Figure 6.1.1. Types of synthetic oscillator designs
Figure 6.1.2. Numerical simulation of the dual feedback oscillator
Figure 6.1.3. Diagram of the expanded model
Figure 6.1.4. Modelled effect of green and red light on the amplitude and period
Figure 6.1.5. Modelled effect of Arabinose
Figure 6.1.6. The proposed plasmid maps
Figure 6.2.1. Crystal structure of one LacI tetramer
Figure 6.2.2. Numerical simulations demonstrating the effect of red and green light
Figure 6.2.3. Examples of the time courses simulations
Figure 6.2.4. Numerical simulations showing the relationship of period on green light
Figure 6.2.5. Numerical simulations for tuneable region shifting
Figure 6.2.6. Numerical simulations of the light tuneable oscillator with the dual proteosome pathway
Figure 6.2.7. Stochastic simulations
Figure 6.2.8. Period distributions
Figure 6.2.9. Amplitude distributions
Figure 6.2.10. Stochastic simulations for the light tuneable oscillator model
Figure 6.2.11. Period distributions
Figure 6.2.12. Amplitude distributions
Figure 6.2.13. Boxplots

List of Tables

Table 1.1.1. Databases and Repositories of Biological Parts and Computational Tools
Table 1.2.1. Induction Methods and Relative Comparison of their Properties
Table 2.1.1. Restriction digest reaction mix
Table 2.1.2. Ligation reaction mix
Table 2.1.3. Ligation reaction mix
Table 2.1.4. Thermocycler program for PCR reactions
Table 2.1.5. In-Fusion® Cloning reaction mix
Table 2.1.6. Excitation, emission and cutoff wavelengths
Table 2.3.1. List of Electronic components for building the LED devices
Table 2.3.2. Strains used in this work
Table 2.3.3. BioBricks™ used in this work. (MIT Registry of Biological Parts)
Table 2.3.4. Plasmids used in this work
Table 2.3.5. List of oligonucleotides used in this work
Table 2.3.6. Solutions and buffers
Table 3.1.1. The three identified modules of the Cph8 system and their components*
Table 3.2.1. Initial conditions ( at time t=0 ) used for the numerical simulations
Table 3.2.2. Qualitative activity for the multichromatic sensor for different light wavelengths
Table 4.1.1. Most commonly used regulators in bacterial synthetic genetic circuits
Table 5.3.1. Observed fluorescence and interpretation of the qualitative behaviour of the LTS
Table 6.1.1. Modular Plasmid combinations for different types of LTO
Table 8.2.1. Oscillatory Phase Table
Chapter 1

Introduction

1.1. Synthetic Biology

1.1.1. Definition and a Brief History
1.1.2. Adapting Engineering Principles
1.1.3. The Engineering Process

1.2 Biological 'Knobs', Control and Induction of Synthetic Genetic Networks

1.2.1 Regulation Biological 'Dials' for Bacterial Systems
1.2.2 Confounding factors and control aspects
1.2.3 User Input - Induction Approaches

1.3. Light sensing systems in prokaryotes

1.3.1. Two Component System Light Receptors
1.3.2. Alternative Light Based Gene Regulation Mechanisms

1.4. Motivation and Objectives

In the introductory chapter I discuss the state of the field of Synthetic Biology and how terms such as "parts" and "modular design", borrowed from other disciplines, are put into a biological context. I present the widely proposed, but yet to be implemented, forward engineering approach for designing, modelling, assembling and characterising synthetic genetic applications. For each step of the engineering cycle I discuss the up to date enabling technologies, but also the current limitations. A summary of the regulatory mechanisms that can be used as control 'dials' for synthetic circuits is given along with a short discussion about current induction methods as a 'user' input. From the latter I explain the advantages that light holds as an inducer and review the light sensing systems and strategies that have been previously published. Finally I discuss the project's motivation and objectives along with how I have structured the work under a forward engineering context.
1.1. Synthetic Biology

1.1.1. Definition and a Brief History

Synthetic Biology is the newly emerged multi-disciplinary field that aims to rationalise and develop the framework for the engineering or re-engineering of anything from basic biological pathways and functions up to entire synthetic life forms. It brings together biologists, engineers, physicists, mathematicians, computer scientists and chemists not only to contribute their expertise but also build the necessary communication bridges between disciplines as the only way to realise the field’s premise. Harnessing the accumulated knowledge and current understanding of living organisms, the field addresses the question of how new non-naturally occurring functions can be implemented and used for our benefit. As an engineering driven field, the potential applications are focused on introducing fundamental research principles and foundational technologies up to the implementation of real world synthetic biological applications in medicine, energy, agriculture, pharmaceutical, remediation and chemicals production industries. The field, employs and adapts the principles that enabled the other traditional engineering fields to deliver the applications that shaped the modern world. These principles include the abstraction, decoupling, modularity and interchangeability of parts and circuitry, sufficient and reliable theoretical framework to guide the engineering process and finally a clearly defined set of standards in terminology, part composition and characterisation data. Although there is not an agreed and clearly defined definition for synthetic biology, all the above summarise the main aspects of the field’s profile, that have been proposed in the literature during the last decade (Endy, 2005; Benner & Sismour, 2005; Andrianantoandro et al, 2006; Heinemann & Panke, 2006; Serrano, 2007; Purnick & Weiss, 2009; Cameron et al, 2014).

Luis Campos (2009) places the first use of the term ‘synthetic biology’ back in 1912 by S. Leduc ‘La biologie synthétique’, before even the DNA structure was solved ( Watson & Crick, 1953 ) and reappears after the discovery or recombinant DNA and restriction endonucleases (Szybalski & Skalka 1978). The modern era of the field is actually placed more recently, around 2000 (Campos, 2009; Purnick & Weiss, 2009; Benner & Sismour, 2005, Cameron et al, 2014). This is when the ”first wave of synthetic biology” (Purnick & Weiss, 2009) begins by the introduction of small genetic blocks such as promoters, ribosome binding sites (RBS), and regulatory components and used to build relatively simple devices like the “repressilator” (Elowitz & Leibler, 2000) and genetic toggle switch (Gardner et al. 2000). Ever since numerous such devices have been proposed such as genetic switches (Atkinson et al, 2003; Dueber et al 2003; Kramer et al 2004 & 2005; Ham et al, 2006...
Chapter 1. Introduction


The "second wave of synthetic biology" (Purnick and Weiss, 2009) started during the last ten years, when the modules and circuits were coupled in order to render higher order functional systems. Some examples were programmed pattern formation through quorum sensing (Basu, 2005), quorum sensing mediated synchronisation of oscillators (Danino et al, 2010, Prindle et al, 2014), a multichromatic optogetic control system (Tabor et al, 2011), an edge detection system (Tabor et al, 2009), a logic iRNA based evaluator in mammalian cells (Rinaudo et al, 2007), an anti-malaria drug production pathway (Ro et al, 2006) even tumor invading bacteria (Anderson et al, 2006). A "third wave" was more recently discussed by Amos (2014), where the author suggests that the use of consortia of engineered cells can overcome many of the challenges of building large and complex networks within the same cell.

Compared to electrical, chemical and mechanical engineering, synthetic biology emerged much later. It has delivered a notable amount of applications (Khalil & Collins, 2010), yet is far from realising its full potential i.e. Industrialised applications (Kwok 2010, Kitney & Freemont, 2012). Partially, this is because the field is still undergoing an incubation period, where fundamental technologies, methodologies and repositories of physical genetic parts, model databases and standards are still under development. The latter means that most of the synthetic genetic circuits to date were built on an ad hoc basis, making the reusability and incorporation of these circuits into larger systems a time and resource demanding task. Kwok, (2010) highlights the "five hard truths for synthetic biology" and how some hypes of the field were almost naively based on oversimplified notions. Notions like "parts can work like lego" collapse quickly into the problems of 'parts' being incompatible or not characterised, or characterised in an arbitrary ad hoc way. The crosstalk of synthetic circuits with other components make the system unpredictable and some times too complex to understand. Finally the most important and fundamental difference of living organisms compared to mechanical or electrical systems is the "variability that crashes the system" (Kwok, 2010). McAdams & Arkin (1997) and Arkin et al (1998) discussed the stochastic mechanisms of gene expression, while Elowitz et al (2002) showed experimentally how the stochastic nature in gene expression is not only due to the mechanism of a gene's expression (i.e. Intrinsic noise) but also the variability from cell to cell in terms of other cellular substances (i.e. extrinsic noise). The contributions of extrinsic and intrinsic noise was also discussed by the same group (Swain et al, 2002) while in the same year Ozbudak et al (2002) showed the contribution of translational and transcriptional efficiency to phenotypic noise. The above, highlight the consequences of the fact that a cell is not a well stirred uniform environment and that biological components cannot be approached, in most cases, as a well-defined and
Adapting Engineering Principles for Synthetic Biology

Addressing the above challenges is of course at the centre of current research, by developing methods and technologies to overcome many of them (Heinemann & Panke 2006). Kahl & Endy (2013) presented a survey on the enabling technologies for synthetic biology that have aided and are continuing the field’s progression. Amongst others, next generation sequencing, affordable gene synthesis, fast DNA assembly techniques, microfluidic platforms, computational tools able to cope with both deterministic and stochastic systems, DNA repositories and model databases are only some of the main contributors to the future of the field. Some of the above are discussed more extensively in the next sections.

1.1.2. Adapting Engineering Principles for Synthetic Biology

Applying engineering to biology inevitably requires living organisms to be approached as a form of machinery, consisting of various parts, sensors, processing modules, actuators etc. A first such parallelism was drawn by Andrianantoandro et al (2006) (shown in Figure 1.1.1(a) ) that aligned the hierarchy on which a computer is built i.e. from the physical layer up to computer networks, with the layers that a cell operates i.e. from the molecular level up to cell populations. An analogy to this hierarchy was drawn by Endy (2005) where each layer exchanges minimal information with its higher or lower order layers as a way to abstract information and simplify the engineering of larger systems (Figure 1.1.1 (b)).

Both hierarchies give rise to the concept of ‘abstraction’ (Endy, 2005) where a very complex system can be broken down to layers that exchange minimum information between each other. For example the protein (‘part’) binding to DNA does not hold any information on the molecular mechanics of the gene (‘DNA’ level) expression that produces it. Similarly, ‘a device’, in this example a circuit that has two inputs and one output can be implemented by a number of combinations of transcription factors and the only information required from the previous layer is what regulation the part performs. ‘Abstraction’ helps both the intuitive understanding of the system but of course simplifies the computational and mathematical representation of a given system.

A complex problem can be solved by splitting it into a number of smaller tasks. This principle called ‘decoupling’ (Endy, 2005) can be applied at different levels. First, the abstraction hierarchy provides a fair separation where one individual can work on the system level by modelling the connection of various devices, while a different individual project can be focused on cloning and expressing the genes needed for operating the specific devices. Apart from an application focused decoupling, a workload can be split between the researchers who develop a cloning method, or engineer a robust cell strain with researchers that analyse a system using computational tools.

For both the abstraction and decoupling to work effectively, a set of standard rules has to be introduced. Standardisation is perhaps one of the most important principles in engineering, as it allows a fast, precise and economic workflow. Standards are needed in all aspects of the field, for example the way that DNA parts can be assembled, like the BioBrick™ assembly protocols where each ‘standardised’ genetic part bears...
Figure 1.1.1. The abstraction hierarchy. **(a)** Figure adapted from Andrianantoandro *et al.* (2006). The author shows how a computer network’s complexity can be broken down to layers and the analogies with a cell’s population organisation. In this case the bottom layer is the physical electronic parts like transistors, resistors and capacitors that correspond to the molecular level of a cell that includes proteins, DNA and other molecules. The parts of the bottom layer put together can render simple functions like chemical reactions parallel to electronic gates. A collection of reactions and electronic components eventually can consist a module that performs a complicated yet self defined function, like sugar metabolism or data storage in computing. Eventually the cell or computer is defined by the functions of its constituent modules collectively. Multiple cells are organised into populations and multiple computers connected through a network. It is worth noting that each layer’s properties are emergent properties of the combined components of the lower ordered layer. **(b)** Figure adapted from Endy (2006), In this case the author proposes a slightly different organisation where the genetic information encoded in DNA is a layer by itself while proteins are placed in the ‘parts’ layer. Again small ‘devices’ are emerging from the combination of various parts while interconnected ‘devices’ constitute a ‘system’.
specific restriction sites that allow for an large number of parts to be assembled together (Knight, 2003). A standard way of characterising parts is needed to enable the development of a computer aided design (CAD) tool to help the design of complex networks having the required performance characteristics in a compatible way. Such standardised characterisation techniques were proposed by Canton et al (2008), where the data produced for a quorum sensing part (F2620) was used to populate a ‘specifications sheet’, similar to the ones used in engineering, to quickly give a rough idea about the part’s properties and performance. The transient and static performance data were presented in physical units (GFP synthesis rate per cell, per second), something that allows for quantitative models to be developed. On the other hand, Kelly et al (2009) demonstrated how standard biological parts can be characterised with respect to a reference part. In this case they showed how relative fluorescence units (RFU) can describe the strength of a library of promoters compared to a reference promoter (J23101). This approach, on one hand solves the problem that different labs using different machinery will measure fluorescence in non comparable arbitrary units (A.U) or with deviations from one another, but on the other hand this method is not quantitative in physical units, making the implementation of predictive models of more complex systems, a much more challenging task. Finally, enzyme nomenclature, model databases, parts databases, enzyme activity units are also highlighted by Endy (2005) as important targets for standardisation.

The idea of standardised parts and circuits is also emphasised by Andrianantoandro et al (2006), as the only way to interface different functional blocks within a given host with a predictable result. ‘Modularity’ and ‘Interchangeability’ of these blocks will allow one researcher to reuse part or a whole previously characterised synthetic construct to build a more complex system without the need of revisiting the structure of the constituent parts. However, it is understood that coupling genetic parts is not such a straight forward process, as cross talk will always need to be taken into account as well as burden and load effects i.e. ‘retroactivity’ (Del Vecchio et al, 2008; Sontag & Del Vecchio, 2009) can alter or ‘break’ genetic circuits when connected together.

1.1.3. The Engineering Process for Synthetic Biology

Apart from the principles adapted from other traditional engineering fields, it is only natural that an established workflow is also needed to guide the implementation of a given synthetic biological application. The engineering process in biology is normally categorised into:

a) An iterative forward engineering process, where a demand is translated into an initial design, which is modelled and built as a prototype that is assessed and refined with each iteration of the cycle. This engineering cycle is further discussed below. Such an example was the dual feedback oscillator (Stricker et al, 2008), which was based on a thorough modelling framework that was published earlier by Hasty et al (2002).

b) A mutagenesis or ‘evolution’ process where numerous variants of the same component are generated following a screening process for selecting the ones that exhibit the desired behaviour. The mutagenesis process can be random (e.g. the evolution of antibiotic resistance in cells), or directed where a defined target
or small part of the device of interest is selected for mutagenesis (Arnold, 1998). Unlike the entirely random process, directed mutagenesis methods require some prior knowledge of the system’s function or possible targets for mutation. e.g. to create a promoter library with different affinities to a transcription factor (TF) the targeted sequence is usually the DNA recognition site for that TF (Ellis et al, 2009; Blount et al, 2012). This method is also used in directed protein engineering with some examples from Frances Arnold’s group where entirely new reactions could be performed by engineered enzymes (Cirino & Arnold, 2002; Brustad & Arnold, 2011).

c) A combination of both a) and b). In fact most of the libraries of biological parts available to date is the product of a mutagenic process, natural or artificial. e.g. the Anderson constitutive promoter library (Registry of Biological Parts) or the SsrA degradation tag library (Andersen et al, 1998) were products of mutagenesis. However, they are used in rational designed, model-guided engineered systems. A clear example was the time-delay switches developed by Ellis et al (2009) where a generated library of different strength variants of one promoter was used into a model-guided construction process. In fact a more recent review by Arpino et al (2013) places 'evolution' as part of the engineering cycle.

**Step 1-Specifications:** This engineering cycle, as presented by Arpino et al (2013) and adapted in Figure 1.1.2 which is followed for a rational forward engineering approach, normally starts by defining the problem, research interest or industrial demand that needs to be addressed. Based on the that, the researcher at step one sets the design objectives, specifications and performance characteristics that the system needs to exhibit. An example is the system presented in Chapter 6 where the design specification and objective was 'A genetic light tuneable oscillator operating in *E. coli* where its frequency and amplitude can be tuned independently'.

**Step 2-Design:** Designing an *in silico* representation of the system is the second step where the designer will use prior knowledge on basic network motifs and design principles for biological circuits (Tyson et al, 2003; Alon, 2007; MacDonald et al, 2011; Baldwin et al, 2012), e.g. feedforward or feedback loops and in what way these can be used to generate the desired behaviour and meet robustness and stability requirements. For example stability analysis of the genetic toggle switch (Gardner et al, 2000) comprising two mutually repressed TFs (shown in Chapter 5, section 1) shows the two stable steady state equilibrium points and one unstable. It is also known that engineering delay in controlled systems, via negative feedback loops can result in oscillations (Goodwin, 1963). This knowledge allows for an initial intuitive design which has to be followed by a more specific detailed design. For this, specific components have to be selected from the literature or preferably libraries e.g. Registry of Biological Parts, Addgene, *E. coli* Genetic Stock Centre (Table 1.1.1). Based on the characteristics and properties of these parts a set of composition rules (Andrianantoandro et al, 2006) must be met for a realistic design. For example, if two small circuits share the same components, the resultant crosstalk can have unpredictable impact on the overall function of the system. Choosing a suitable strain with the required gene knock-outs is equally important in order to eliminate crosstalk as much as possible. Plasmid incompatibility, cytotoxic effects, gene orientation, possible burden and retroactivity (discussed below) are
also important factors that have to be taken into account when complex systems are designed. It is worth noting that unfortunately having all the required information regarding the available biological parts is rarely the case. Hence the designer has to compensate for the large uncertainty of the system with mechanisms that allow a relatively simple substitution of parts that are more likely to require fine tuning. Another strategy can be the engineering of redundancy in the network, when this is possible, so that if one component fails another one can compensate for the loss of function. A recent example by Dr James Arpino (personal communication - unpublished results) is a plasmid that harbours multiple identified terminator sequences. These can increase the probability of recombination. However, the plasmid was designed in a way that if any parts of it is lost by recombination the cells cannot survive, by loosing either the origin of replication or one of the two antibiotic resistance cassettes. A plethora of computational tools is available that can assist the design process. Examples include the Systems Biology Markup Language (SBML) based model editors (Hucka et al, 2003), Tinker Cell (Chandran et al, 2009), CellDesigner (Funahashi et al, 2003), Systems Biology ToolBox and Simbiology® in MATLAB. These tools include built in functions that are normally used to describe biochemical systems, like mass-action, Michaelis-Menten kinetics (Johnson & Goody, 2011: Michaelis, Menten, 1913) and hill equations. They also provide access to model databases like BioModels (Le Novère, 2006) and normally have a graphical user interface (GUI) that allows for users with relatively basic computational and mathematical background to assemble biological networks. However, none of the available software has proven or reached the status of the corresponding CAD-like tools in the other engineering disciplines, not necessarily because of programming limitations but mainly due to limitations in the available parts, characterisation data, standards and even the lack of general understanding of biological processes (Kwok, 2010). A non-exhaustive list of biological parts, databases and modelling tools is given in Table 1.1.1.

**Step 3-Modelling:** The design process leads to the third step where a mathematical model is derived and informed with specific parameter values or range of conditions that the experimental system will operate. Again, often not all parameter values are available in the literature or the databases of biological parts. Instead, a combination of literature mined values, statistical observations, experimentally determined parameters, fitted parameters or even initial guesses within biologically relevant limits are employed. The most common types of models used in synthetic biology are nonlinear (NL) systems of ordinary differential equations (ODE model). These can be derived using first principles, by converting a set of biochemical reactions, also known as the kinetic scheme, that represents the system into their corresponding ODEs, according to the law of mass-action. These can be a sufficient representation of a system that is assumed to be evolving in a homogenous environment (McDonald et al, 2011), and its species are changing with time as the only independent variable. For including a second independent variable like space, a partial differential equation (PDE) system can be used instead. Of course the nonlinearities in ODE models do not allow for an exact analytical solution of the time evolution of the system. Therefore, these are normally solved by numerical solvers (e.g. ODE15s, Sundials, ODE45 ) for a given time range under specific starting conditions. These solvers are implementations of the Euler’s and Runge-Kutta algorithms (Stan, 2014) and will return a deterministic solution of the system. Deterministic simulations will always return the same result and allow an
Chapter 1. Introduction

Figure 1.1.2. The engineering cycle. The figure was adapted from Arpino et al., (2013). The figure shows the steps that are normally followed for a forward engineering process. The dashed lined boxes show the current enabling technologies and methods for each step of the cycle. The cycle begins with the definition and the specifications of the system in order to meet a specific demand or problem to be solved. The second step starts with a conceptual design, based on standard network motifs and network topology, that can give the desired behaviour and finishes with the more detailed design that includes specific biological ‘parts’ to be used and connected with respect to known composition rules. The third step is the derivation of the mathematical model that can be interrogated in order to describe, provide some further understanding or even predict the properties of the physical system. At this point if the model indicates that the proposed design needs further improvement the cycle between steps two and three can be reiterated. Next, during the fourth step the system is constructed from parts found in repositories and other sources of biological material, or simply synthesised chemically, followed by sequencing verification. The main cycle ends with the testing and characterisation of the system. The outcome from this step can determine whether the application meets the set specifications and objectives, in which case the cycle can stop, or if it needs refinement to reiterate. The data obtained at this point can by used to inform the model at step three in order to increase its predictability. Alternatively, an evolutionary method can be applied in order to obtain variants of the system with improved performance.
estimation of how the system will evolve on average, depending on the initial conditions, parameter values, external stimulus/inputs, for an infinite number of samples or repeats. Deterministic solutions, although fast and useful in some aspects, are often not a good representation of a living cell, simply because the basic assumption of a homogeneous environment is not the case within the cytoplasm. Microenvironments, protein localisation, random fluctuations (Elowitz et al, 2002; Ozbudak et al, 2002) become significant especially when operating with a small number of molecules and give rise to stochastic effects. To approximate these effects in silico, probabilistic models are used. These include stochastic differential equation models (SDE) and Chemical Master Equations (CME). SDE models like ODE, are continuous models but they include a noise function which will return a different value for every run of the solution. Chemical Master Equations describe the joint probabilities of all time variables to have specific values at specific time points. Their solutions are extremely demanding in computation which is done through a Stochastic Simulation Algorithm (SSA). SSAs, implementations of the Gillispie algorithm (Gillespie, 1976 &1 977 ), return discrete solutions and simulate exact particle interactions by assigning a propensity function that describe the probability of a given reaction to 'fire' at a specific time step. Stochastic models are better approximations of the physical system and can reveal states and behaviours not observable in the deterministic regime, yet the quality and predictability of the model is still subject to how well structured and informed the model is in terms of parameters.

Populating the parameter set of a given model is one of the most challenging tasks. As mentioned above, if a parameter cannot be mined from the literature, targeted experiments or databases, then it has to be inferred from data obtained from a prototype of the system (Step 5 of the cycle). Numerous methods have been proposed in the literature that are applicable for biological model parameter estimation and optimisation (Mendes & Kell, 1998; Moles et al, 2003; Lillacci & Khammash, 2010; Beguerisse-Diaz et al, 2012). A large number of parameter sets that explain the data can be reached due to multiple local error minima that exist in the hyper-dimensional space that these algorithms search. Hence posterior selection and manual constraints need to be applied in order to reach a parameter set that falls within biologically relevant values. Nevertheless, the NL ODE models normally used to describe biological systems are just not identifiable (Ljung, 1987; McDonald et al, 2011) since the "minimum information needed to identify the model" is almost impossible to obtain. Thus any estimated values, even if they satisfy a given dataset, cannot be assumed to be the real physical values. An alternative way to tackle the complexity of large dimensional non linear ODE models is model reduction. This can reduce both the number of time variables and number of parameters by non-dimensionalisation or lumping them together. Model reduction is based on approximations (e.g. quasi-steady state approximation) that can reduce the complexity of the system and still reproduce the rough behaviour of the system.

Other types of modelling include the derivation of inductive models (Stan, 2014) or phenomenological models. These models require prior knowledge and data of the system’s behaviour, where instead of constructing a model from a principal reaction scheme, the modeller uses a combination of intuition and pre-determined biologically relevant expressions to derive a minimum size model that can explain the system’s behaviour like in the case of the toggle switch, the ‘repressillator’ (Gardner et al, 2000; Elowitz et al, 2000) and more recently...
a light sensing system (Olson et al, 2014). Finally the literature is rich in 'reverse engineering' approaches for selecting or inferring models and network topologies, provided that data is already available (Gardner et al, 2003; Bansal et al, 2007; Sontag, 2008; Gonçalves & Warnick, 2008; Yuan et al, 2011).

Regardless of the type, biochemical models can vary in complexity and detail depending on what is necessary for the model to include, measurable and what are the possible tuning 'dials' of the physical system (Baldwin et al, 2012; Arpino et al, 2013; Stan, 2014). They can be used to estimate the behaviour of the system under a range of conditions and inputs and they can be tested for robustness by sensitivity analysis, or used to predict behaviours and emergent properties of more complex systems. Depending on their predictive abilities, models can guide a rational design approach by indicating which parts to change and how they can be connected, something that eliminates the need for building large libraries of a circuit and screening for the required function.

**Table 1.1.1. Databases and Repositories of Biological Parts and Computational Tools**

<table>
<thead>
<tr>
<th>A. Databases for Models, Parameter Values and Bioinformatics Tools*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bionumbers</strong></td>
</tr>
<tr>
<td><strong>BioModels Database</strong></td>
</tr>
<tr>
<td><strong>CellML Model Repository</strong></td>
</tr>
<tr>
<td><strong>The Cell Collective</strong></td>
</tr>
<tr>
<td><strong>KEGG</strong></td>
</tr>
<tr>
<td><strong>BioCyc</strong></td>
</tr>
<tr>
<td><strong>ExPASy</strong></td>
</tr>
<tr>
<td><strong>EcoCyc</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Main Repositories of Biological Parts*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli Genetic Stock Center</strong></td>
</tr>
<tr>
<td><strong>AddGene</strong></td>
</tr>
<tr>
<td><strong>Registry of Biological Parts</strong></td>
</tr>
<tr>
<td><strong>ATCC</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C. Examples of Model Editors and Tools**</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SimBiology 3</strong></td>
</tr>
<tr>
<td><strong>Systems Biology Toolbox</strong></td>
</tr>
<tr>
<td><strong>Systems Biology Workbench</strong></td>
</tr>
<tr>
<td><strong>TinkerCell</strong></td>
</tr>
<tr>
<td><strong>CellDesigner</strong></td>
</tr>
<tr>
<td><strong>Copasi</strong></td>
</tr>
<tr>
<td><strong>MatCont</strong></td>
</tr>
</tbody>
</table>

* A More extensive list of Databases relevant to synthetic biology is given by Kahl & Endy (2013)

** A more complete list of model editors and simulation tools is given by Marchisio & Stelling (2009) and McDonald et al (2011)
Step 4-Implementation, Assembly: Following the *in silico* part the engineering cycle proceeds to the *in vivo* or *in vitro* implementation of the system. During step four the system can be constructed by cloning together the necessary open reading frames (ORFs) and regulatory sequences *e.g.* promoters, Ribosome Binding Sites (RBS) and terminators, based on the proposed design. A big part of the enabling technologies for this step is focused on standardising and making the part assembly procedure, simpler, cheaper and faster. The spectrum of available methods range from traditional cloning methods like enzymatic restriction digestion and ligation to modern methods like Goldengate (Engler *et al.*, 2008), CPEC (Quan & Tian, 2014), USER (Nour-Eldin *et al.*, 2010), inFusion™, SLIC (Li & Elledge, 2007) Gibson assembly (Gibson *et al.*, 2008) and CRISPR (Burgess, 2013). The most modern techniques allow for the combinatorial assembly of multiple DNA fragments in one cycle in only a few hours. In addition, the price per base pair that gene synthesis companies offer has radically decreased over the past decade. Hence, it is now a viable option for ordering relatively large constructs instead of assembling everything from scratch. An example is the first organism operating on a chemically synthesised genome (Gibson *et al.*, 2010) and the refactoring of the nitrogen fixation pathway (Temme *et al.*, 2012), where both projects included the chemical synthesis of genetic constructs and parallel assembly of them in an unprecedented scale. Apart from synthesis, or individual groups exchanging constructs, parts can be obtained from open access repositories. Some of the existing libraries are shown in Table 1.1.1 (B) like the MIT's Registry of Biological Parts, that facilitates thousands of BioBricks™ generated from undergraduate teams taking part in the international genetically engineered machine (iGEM) competition. However the data and the annotation that accompanies these parts is often incomplete, poor in quality or even wrong, thus this library although large, lacks reliability. AddGene, and Biofab are also open access repositories where individual research groups can deposit their genetic parts and constructs for distributing around the synthetic biology community. These libraries are still small but operate with much higher standards in terms of verifying the sequences and annotating their DNA. In either case, the necessary specification and performance data is still a grey area, because as mentioned above the standards are not clearly defined and applied to high throughput technologies for characterising these parts. Finally an alternative source for diversity and generation of new parts as mentioned earlier is the use of random or directed mutagenesis. The latter along with most of the modern assembly methods are currently being scaled up by the use of robotics and other forms of automation (Nottebaum *et al.*, 2008; Wang *et al.*, 2009; Wang & Church, 2011; Lux *et al.*, 2012).

Step 5-Characterisation: The last step of the cycle is the testing and characterisation of the system. Current methods aiming to characterise the activity of a network involve the use of a fluorescent reporter or an enzyme (GFPs and variants, LacZ, Luciferase). However characterisation data have been mostly presented in relative arbitrary units (Kelly *et al.*, 2006) with fewer examples of absolute quantification (Canton *et al.*, 2008). For one the relative units (based on a reference) help to ease the inevitable differences between different types of machinery in various labs around the world, however quantitative measurements can be more informative for mechanistic models. Depending on the application the characterisation can be determined at a population level by simple means, like a fluorometer, plate readers, flow cytometry or at a single cell level by fluorescent or even confocal microscopy. This has enabled ultra high resolution imaging
down to single molecule chase assays (Hammar et al, 2014) while advances in microfluidics have enabled the observation of single cells under chemostatic conditions (Stricker et al, 2008) and scaled up to high throughput characterisation platforms (Maerkl, 2009). An engineered genetic network can also be expressed in a cell-free environment as a way to eliminate most of the confounding factors that accompany an in vivo assay. The TX-TL cell free system is gaining ground amongst biologists for rapid prototyping and assessment of synthetic biological systems (Gerber et al, 2009; Niederholtmeyer et al, 2013; Sun et al, 2013 & 2014; Takahashi et al, 2014). Other ways of assessing the system include measurements at a transcriptional level by Real Time quantitative PCR (Heid, 1996) and more recently scaled up to high throughput RNAseq technology (Marguerat & Bähler, 2010).

The data collected at the end of the final step indicate whether the engineered system meets the specifications and objectives set at the first step. However, even in other established engineering fields, one cycle is not sufficient to generate a fine tuned product. Instead the available data can be fed back to the model, in order to increase its predictive ability. This will allow for a re-design proposal aiming at increasing the systems performance and robustness. This re-design will ideally be directed towards the fine tuning of one or few components (e.g. strength of an RBS) rather than changing the structure of the network. These 'control dials' are discussed further in the next section.

1.2. Biological 'Knobs', Control and Induction of Synthetic Genetic Networks

1.2.1. Regulation Biological ‘Dials’ for Bacterial Systems

A synthetic genetic network can generate an output at different magnitudes, at a specific temporal or spatial pattern, all in response to a user input (induction methods discussed in the next section). In bacteria, the transformation of a 'signal' to the final output goes through a genetic regulatory network comprised of transcriptional regulation of promoters by specific transcription factors (TFs), translational regulation through riboswitches, tuning through RBS strength or post-transcriptional regulation of the mRNA and finally post-translational modification and degradation mechanisms. Throughout these regulatory stages a system can be tuned in numerous ways. The most common and practical approaches in terms of time and resources include plasmid copy number, RBS strength, modifications on the promoter strength, mRNA secondary structure and the use of degradation tags. These tuning points can be referred to as 'biological knobs' or 'dials'. The discussion below aims to highlight some of the commonly used strategies, in bacteria, to fine tune a system’s performance or behaviour within the forward engineering framework discussed in section 1.1.3. The most common dials are summarised in Figure 1.2.1 and discussed further below.

On a transcriptional level the most commonly used knob is the strength of the promoter. It is worth noting that ‘promoter strength’ is an arbitrary term. More recently the RNA polymerases per second (PoPS) that pass through a point of the DNA strand is used as a metric (Endy, 2005; Varadarajan & Del Vecchio, 2009).
The recruitment rate can vary depending on the type of sigma factors ($\sigma^{70}$, $\sigma^{54}$, $\sigma^{32}$) and their affinity to the -35 promoter sequence along with the distance and the sequence composition between the -10 and -35 (Arpino et al, 2013). TF binding sites also block or promote the recruitment of polymerase. Examples include different version of LacI regulated promoters like the Ptac, Ptoc and Ptrc families (Lutz & Bujard, 1997) which are all regulated by the same TF but have different strengths. Such libraries were shown by Ellis et al, (2009) to be relatively simple to derive by applying a random mutagenesis method following sorting and selection of the desired strength. Other examples include promoters that recruit the T7 RNA polymerase found in T7 phage (Tabor & Richardson, 1985). Lutz & Bujard (1997) demonstrated how hybrid promoters of different strength can be constructed like the Plac/ara-1. Small RNA (sRNA) has also been used to interfere with transcription by stabilising or destabilising the secondary mRNA structure of the 5' Untranslated Region (UTR) between the transcription start site and the start codon of the ORF and or the 3' UTR between the stop codon and the 3' end of the mRNA (Chappell et al, 2013). The secondary structures that form naturally depending on the sequence composition can terminate transcription by the rho-mediated (Richardson, 2002) or rho-independent mechanism (Farnham & Platt, 1981). A similar approach includes regulation by riboswitches, which again are 5' UTRs that form terminating secondary structures but can be switched on by a small molecule (inducer), promoting or repressing transcription. An example of the latter was recently demonstrated by (Chinnappan et al, 2013) where both the transcription efficiency and regulatory activity of a riboswitch and its analogs were measured with high time resolution. Finally, CRISPRi was demonstrated as an alternative regulatory mechanism where an inactive Cas9 mutant (dCas9) was targeted to DNA through a small guide RNA (sgRNA) and terminated transcription prematurely (Qi et al, 2013).

On a translational level, the ribosome recruitment rate on the mRNA can be controlled by modifying the Ribosome Binding Site (RBS) sequence in a fairly predictable manner (Arpino et al, 2013; Reeve et al, 2014). A number of computational tools have been proposed (Na & Lee, 2010; Seo et al, 2013) that can be used to design RBS sequences of varying strengths, but the most widely used one is the RBS calculator from Salis et al (2009). The latter is able to predict and design individual or libraries of RBSs with a strength spanning from a range of 1 to 100000 AU, using a free energy model that takes into account long distance interactions, sequence composition and hybridisation of the Shine-Dalgarno (SD) sequence with the 16S rRNA, secondary structure unfolding, ribosome spacing on a given mRNA and organism (Salis et al, 2009; Salis et al, 2011). This tool even with high degree of uncertainty has been used for forward engineering design and refinement (Chen et al, 2012; Temme et al, 2012; Pothoulakis et al, 2014). Zelcbuch et al, (2013) showed how combinatorial assembly of a set of RBSs can be used to tune protein production over several orders of magnitude, while Egbert & Klavins (2012) used repeated sequences to vary the 5' UTR sequence lengths to modulate translation rates and fine tune a toggle switch circuit. Similar to transcription the use of small RNA and riboswitches have been used to occlude the RBS site and block its access to the ribosome, or affect the rate at which the mRNA is degraded by stabilising its structure and the RNase's access to it (Deana & Belasco, 2005, Liand et al, 2009; Chappell et al, 2013; Mackie, 2013). Finally, codon optimisation is employed for improving translation efficiency given the differences in the abundance and incorporation rates of tRNA between different organisms.
during the translation process (Hershberg & Petrov, 2008; Angov, 2011).

At a protein level, the most straightforward way of fine tuning a synthetic circuit is by changing the rate at which a protein is degraded. Proteins, either regulators, enzymes, reporters or other effectors are very stable in commonly used bacterial organisms like *E. coli*. Their natural half life extends from several hours to days. Hence during a typical time course run these proteins are only ‘removed’ from the system due to dilution during cell growth and division. On the other hand targeting a protein to a proteosome allows for faster degradation rates. Several protein degradation pathways have been identified in bacteria that are vital components of the cell physiology and cell division (Jayasekera *et al*, 2000; Camberg *et al*, 2011). They involve a protease or proteosome structure, like Lon (Tsilibaris *et al*, 2006), Tsp (Silber *et al*, 1992), FtsH (Narberhaus *et al*, 2009) and ClpX, ClpS, ClpA and ClpP complexes (Ortega *et al*, 2004; De Donatis *et al*, 2010), that recognise N-terminus and C-terminus oligopeptide signals (Flynn *et al*, 2003) called degradation tags, on which the protease binds, unfolds and cleaves the targeted proteins. Andersen *et al* (1998) has characterised a library of SsrA tags recognised by the ClpXP proteosome and showed how a GFP's half life can vary between 60 to 120 min depending on the amino acid residue composition of the signal peptide. These have been shown reliable in controlling protein decay rate in well known synthetic circuits like the ‘repressilator’ (Elowitz *et al*, 2000) and dual feedback oscillator (Stricker *et al*, 2008). Another example is the expression and use of the tobacco etch virus (TEV) protease (Kapust & Waugh, 2000) to cleave fusion proteins. Expression of non native proteases in bacteria like the latter, can potentially be a powerful tool in order to eliminate the overloading and cross feedback effects of the native proteases that can severely affect the normal cell operation as discussed by Camberg *et al* (2011) and Cookson *et al* (2011) and observed in Chapters 5 and 6.

Alternative to degradation mechanisms, numerous protein engineering approaches (with their detailed discussion extending beyond the scope of this project) can be applied to tune the 'effectiveness' of a protein *i.e.* catalytic rate for enzymes (Böttcher & Bornscheuer, 2010), binding affinity to ligands or inducers (Betz, 1986) and DNA binding site for TFs. More interesting though are cases where entirely new functions arise from protein engineering. Examples include a light sensing chimeric receptor between the EnvZ osmolarity receptor in *E. coli* and the Cph1 light sensing domain from *Synechocystis* sp. (Levskaya *et al*, 2005), enzymes engineered to catalyse new reactions considered 'difficult' by chemistry means (McIntosh *et al*, 2014) and engineered TetR inducible to the non-natural inducer, cmt3 (Scholz *et al*, 2003). Dueber, *et al* (2009) showed how engineered protein scaffolds can be used to achieve control over metabolic pathways. Although powerful and with wide scope of applications, protein engineering approaches are normally slower and more resource demanding compared to the methods discussed above, thus they are not the primary choice for refining a synthetic circuit.

Various other factors that contribute to the regulation and fine tuning of systems exist. Gene copy number is a very common biological dial and can range from a single copy integrated into the genome or cloned into a bacterial artificial chromosome (BAC) (Wild *et al*, 2002) up to hundreds of copies on a high copy plasmid backbone. Plasmid backbones commonly used in bacterial systems normally fall under the three
incompatibility groups of origins of replication. Replicons within the same group were thought not to be stable within the same host however more recent results contradict this (Velappan et al, 2007). A CoIE1 based high (CoIE1 and pMB1R12 at 50 copies) and very high (pMB1 at 300 copies) copy number origin, a p15A medium copy number (15-20) and a low copy number origin pSC101 (3-5 copies) have been used in synthetic biological circuits. Finally, inducible copy number plasmids have been reporter that can be induced from 50 up to 300 copies per cell when induced (Bachvarov et al, 1990). Arpino et al (2013) reports a more extensive list, however the authors point to the fact that the copy number associated with each replicon was estimated by using bare backbones and these numbers can change significantly depending on factors dictated by the insert constructs. One of these is the genetic construct architecture, that was shown to affect the optimal expression of genes expressed from the same vector, depending whether they are placed in a divergent, convergent or tandem orientation while the spacer region between genes is also a contributing factor (Yeung et al, 2014).

Last but not least, is the 'chassis' in which the network is expressed (Arpino et al, 2013). Expressing the same synthetic genetic network in different organisms or even strains of the same species can be significantly different, due to variability regarding cellular resource supplies (e.g. amino acids, tRNA etc.), availability of ribosomes and RNA polymerase, growth rates (thus different dilution rates), tolerance to antibiotics and stress from environmental conditions like temperature, osmolarity, pH, nutrient supply and aeration. e.g. A 2-fold difference in GFP expression between E. coli TOP10 and E. coli W3110 (Kelly et al, 2009 in Chappell, 2013) strain was previously reported. These factors can be tuned be either choosing (from biological repositories shown in Table 1.1.1) or engineering the right genotype or changing if possible the environmental conditions that the host operates until the synthetic circuit's performance is optimised.

### 1.2.2. Confounding Factors and Control Aspects

Expressing even very simple synthetic genetic devices e.g. expression of GFP from a constitutive promoter will almost inevitably have some impact on the host's physiology. This is due to the fact that synthetic genetic circuitry is operating on the same shared resource pool as the 'housekeeping' genes of the host. As the synthetic systems become larger and more complicated these effects and new causes of 'failure' arise that become significant to the point that the system might behave in an unpredictable manner. The main confounding factors that can potentially disrupt and cause this unpredictability were previously discussed in an excellent review by Cardinale & Arkin (2012) and the main points are summarised below.

Burden is the most commonly observed complication when large or heavily overproducing circuits are used. Burden occurs when the synthetic circuitry recruits cellular resources e.g. dNTPs, amino acids, ribosomes and RNA Polymerases at such high rates that the host's normal cell cycle is affected, resulting in slower growth rates and slower overall protein production, (Vind et al, 1993; Dong et al, 1995; Algar et al, 2013). Under these conditions it is apparent that other metabolic processes are slowed down and 'emergency' responses are triggered e.g. the stringent response (Harcum, 2002; Ferullo & Lovett, 2008), something that inevitably
Chapter 1. Introduction

Biological 'Dials' and Regulation

**Figure 1.2.1.** Biological 'Dials', effectors and regulation of gene expression in bacteria.
feeds back and alters the circuits behaviour in terms of overall protein production but also changes in plasmid copy number and growth feedback effects (Tan et al, 2009 in Cardinale & Arkin, 2012). Heavy host burden can sometimes result in plasmid rejection or recombination especially between repeated palindromic terminator sequences (Bjedov et al, 2003). Algar et al, (2013) demonstrated and modelled the drop in constitutive expression of mCherry from plasmids or genomic copies when GFP expression under a strong arabinose inducible promoter P_{BAD} (Siegele & Hu,1997) is induced. However sufficient 'quantification' of the burden at the extent where synthetic biologists can predict its impact on the their designed circuits does not exist to date. Given the limitations of expressing complex networks within the same host the current interest is turning towards splitting large networks across cell populations or cell consortia (Brenner et al, 2007 & 2008; Shong et al, 2012; Amos, 2014).

Unlike electronics that can use the same material for wiring, every 'connection' within synthetic circuits has to be mediated by a different component (TF, signalling molecule or RNA molecule). Hence for building large and complex networks a large number of these components must be made available. In fact the number of reliable and well characterised components not only are very limited, but also most of them are native regulators of the most commonly used 'chassis' (e.g. LacI in E. coli). This restricts the compatibility of these circuits with most available chassis in order to eliminate crosstalk to and from cellular operations. As discussed in Chapter 4, the same exact regulators (LacI, TetR, CI, AraC etc) have been used in almost every bacterial synthetic genetic circuit that has been constructed, something that restricts the compatibility of a circuit with the host, but also with other synthetic circuits. To combat the above a recent study has shown the successful use of an 'editable' or 'diversifiable' DNA binding site domain of a TAL effector (Blount et al, 2012). Using the specific protein a large number of orthogonal inducible TFs were created and were able to regulate their cognate engineered promoters. More promising potentials for addressing the limitations in regulatory components is found in the increasing research around sRNA, riboswitches and RNA secondary structure prediction (Chappell et al, 2013) since a large number of these regulators can be used within the same host. Another advantage of using RNA for regulation is also the fact that it is a translation-free mechanism something that can significantly ease the burden effects discussed above. Finally, RNA is produced and degraded faster than proteins (Average mRNA half life is 5 min compared to a protein which is several hours (Bionumbers database)), meaning that more dynamical behaviour can be achieved in synthetic networks without the need or risk of overloading the protease machinery.

Small synthetic biological circuits that have been developed during the 'first wave' of synthetic biology (Purnick & Weiss 2009) have been characterised in isolation from one another. Naturally, when coupled together a 'loading' effect is caused from the downstream circuit to the preceding one since the connecting species is also part of the upstream circuit. This 'retroactivity' effect is essentially caused due to the fact that the pool of that specific species is now shared between the upstream circuit and the connecting mechanism of the two and this effect was shown to affect the temporal dynamics and in some cases to alternate the original characterised behaviour of the circuits in isolation e.g. loss of oscillatory behaviour (Del Vecchio et al, 2008; Del Vecchio, 2010; Jayanthi et al 2013 ). More recent results however have modelled, analysed and demonstrated
Chapter 1. Introduction

how ‘insulation’ circuits can be used to eliminate the retroactivity on the upstream network. The insulators work by a high gain or a fast (time scale separated) negative feedback controlled phosphorylation cascade between the connected circuits (Jayanthi & Del Vecchio, 2011).

Adding to the above, the perturbations in the environmental conditions that a host is operated in and considering the stochastic nature of gene expression, it is easily understood that for a given synthetic network to operate predictively and robustly, principles from control theory need to be applied. This can be achieved by means of engineering autoregulation within circuits with the incorporation of negative feedback controllers (Becskei & Serrano, 2000; Simpson et al, 2003; El-Samad et al, 2005; Dublanche et al, 2006; Oyarzún & Stan, 2012). Closed-looped systems that interface cell populations with in silico control feedback have also been proposed recently (Milias-Argeitis et al, 2011). Even further, Proportional-Integral (PI) controllers have been investigated on a theoretical level (Klavins 2010; Ang et al 2010) but have only recently been implemented by Khammash M. group (personal communication - unpublished results). Their basic architecture consists of a negative feedback loop and a positive amplification, where an input inducer can be used as a setpoint.

1.2.3. User Input - Induction Approaches

Most genetic networks incorporate some access points for external input by the user, in order to activate, deactivate or finely control their operation. This control can be achieved at different levels e.g. inducers and growth conditions. Conveniently, most transcriptional regulators found in nature are often associated with their natural ligands, or co-factors that can modify their physiological activity. The overwhelming majority of gene expression control approaches involves the use of a small molecule ligand (e.g. a sugar, amino acid or antibiotic) that can regulate by inhibiting or enhancing the effect a TF has. However other types of induction have been demonstrated previously like temperature denaturation (Gardner et al, 2000) or UV (Lusic et al, 2007) induced expression. More recently two component systems have been reported to regulated gene expression in response to optical light wavelengths, discussed more extensively in section 1.3.

For achieving control over biological processes it is clear that the chemicals or other inducers need to fulfil some requirements, in terms of their orthogonality, toxicity to the host and stability.

The input signal needs to be orthogonal in its effect so it can bind and alternate the activity of one type of regulator. For example aTc and IPTG have been used to induce reliably only TetR and LacI, respectively, without any reported specific binding to other molecules. Rhamnose on the other hand, is a ligand to both RhaR and RhaS regulators (Egan & Schleif, 1993) something that needs to be taken into account if both regulators are used in a synthetic circuit. For other examples like temperature or UV it is more easily understood that they can interfere with cell regulation and have global effects. Another level of crosstalk was addressed by Sung et al, 2007, where IPTG was shown to be inhibiting to the arabinose mediated induction of AraC.

The inducers need to be non-cytotoxic over a wide range of concentrations. No matter if it is chemical, temperature or light, all inputs (or outputs) have an operational range that will not affect cell metabolism and growth. However, extremely high concentrations of IPTG (Wilson et al, 1981; Kosinski et al, 1992), aTc
(Gossen & Bujard, 1993), temperature and UV will have moderate to even lethal cytotoxic effects to the organism. In fact, the most commonly used reporter, GFP has some cytotoxic effects (Siegele & Hu, 1997). The ideal operational range for a given inducer should be enough to reach saturation levels of induction for normal concentrations of a particular regulator without any impact on the cell’s housekeeping function. Inducers have to be stable and non-metabolisable by the host organism so that their concentrations remain constant for long lasting experiments or applications. While cells cannot metabolise IPTG or aTc, sugar based inducers like arabinose are catabolised by living bacteria. This means for a time course assay the effective concentration of that inducer will be changing throughout the course of the assay. In addition, bacteria can change their primary carbon source to sugars used as inducers. The resources needed to produce the enzymes for different carbon sources is by itself a process that can affect doubling time and affect the transient dynamics of synthetic genetic networks. However, other non-metaboliseable inducers like aTc have been shown to be temperature and light sensitive which means they will degrade over the course of time even under normal experimental conditions. IPTG on the other hand is a relatively stable and a non-metabolisable inducer. The cell metabolism or essential house-keeping mechanisms need to be insensitive to the particular molecule. For example, using UV as an inducer will activate the stress response pathway of the host that apart from cleaving λ-CI (Galkin et al, 2009) will have numerous other effects on the normal cell regulation. Arabinose in high concentrations can trigger a change of the metabolism of the cell so that it takes up and uses arabinose as a primary carbon source. The above will certainly have an impact in terms of available resources and burden on the cells resulting to unintended shifts in the behaviour of the synthetic genetic network of interest. Preferably the inducer needs to have a fast effect. This is partially a property that is related at first with the ligand's affinity to its cognate regulator but also whether the molecule can freely diffuse through the cell membrane or needs a specific transportase to cross it. The latter is normally a slower process. e.g. aTc is a small molecule that can be diffused freely through the membrane as opposed to long carbon chain variants of HSL molecules involved in quorum sensing (Pearson et al, 1999) or lactosides like the commonly used IPTG where the LacY permease regulates the active efflux of the molecule (Hansen et al, 1998). Other types of induction like light and temperature can be uptaken by the relevant sensory mechanisms instantly compared to chemicals. Finally, in most cases it is preferable for a user-input to be reversible on its effect. Considering a typical regulator that undergoes a conformational change when induced, it will be faster and less energy demanding for the cell if this regulator was able to recover its original state when the input is removed, rather than replacing the entire regulator pool with newly expressed proteins. For chemically induced systems reversibility can not be easily of efficiently achieved. Practically, for any assay or application running in a static environment, the only way for the user to remove this signal is to replace the growth medium with one which is free of the specific molecule. On the other hand, for non isolated systems, previous studies have shown how microfluidics can be employed to cycle through different chemical inducer concentrations (Mondragón-Palomino et al, 2011). Light induction on the other hand has the major advantage that any induction can be applied or reversed
instantly compared to the time scales of the other cell processes (Levskaya et al, 2005; Levskaya et al, 2009; Tabor et al, 2011; Davidson et al, 2013).

Other interesting characteristics that an induction can have involves the uniformity of the induction (i.e. across a culture) and the spatiotemporal resolution of the induction. The frequency for which a chemical inducer can be applied and removed ranges from hours to several minutes, while light pulses can cycle at fractions of a second (several Hz) (Levskaya et al, 2009). A light or lazer beam’s diameter can be condensed to \( \mu m^2 \) while localisation of chemical inducers in the cytoplasm is not possible due to diffusion, at least in their native form. Chemicals are expected to suffer from heavier stochastic effects in the non well-stirred intracellular environment, something that light induction overcomes by the uniform photon density in a beam of light generated by electronic components or lazer. In addition light equipment (LEDs,lazer etc) are found in electronic setups that can be easily coupled and controlled by computers in order to apply more complex and finely tuned induction patterns, as opposed to chemicals. Automating the latter would require, either a robotic setup or a microfluidic platform with electronic pumps something that is more costly compared to simpler DIY electronic interfaces that can control an LED.

The above, summarised in Table 1.2.1., suggest that light is the induction approach which holds the most useful properties for fine control over genetic networks. Light is the fastest inducer, constant and uniform in its application, it allows the user to have an ultra-fine control over the space to be applied and timing of the application and it comes at a low cost compared to purified and synthesised chemicals. Finally, it is reversible and can be easily automated and driven by a computer setup. The down side for this particular type of induction, is the limited number of well characterised bacterial light responsive systems, the fact that UV will have a wider impact than just the intended regulator or cytotoxic effects and also the fact that the two component systems already characterised can have long absorption spectrum tails suggesting that they can be activated to some extent by light wavelengths far off their peak absorbing wavelength (Tabor et al, 2011).

Table 1.2.1.  Induction Methods and Relative Comparison of their Properties

<table>
<thead>
<tr>
<th>Type of induction</th>
<th>Examples</th>
<th>Stability</th>
<th>Effect</th>
<th>Orthogonality</th>
<th>Spatial Resolution</th>
<th>Temporal resolution</th>
<th>Cost</th>
<th>Automation cost</th>
<th>Reversibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organic molecules</td>
<td>Tryptophan, lactose, arabinose, rhamnose</td>
<td>Poor</td>
<td>Low</td>
<td>Average</td>
<td>Low</td>
<td>Low</td>
<td>Average</td>
<td>High</td>
<td>No</td>
</tr>
<tr>
<td>Chemically modified molecules</td>
<td>IPTG, aTc</td>
<td>Average</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>No</td>
</tr>
<tr>
<td>Temperature</td>
<td>CI high temperature sensitivity</td>
<td>High</td>
<td>Average</td>
<td>Poor</td>
<td>Low</td>
<td>Average</td>
<td>Low</td>
<td>Average</td>
<td>No</td>
</tr>
<tr>
<td>Light</td>
<td>Optical light, UV</td>
<td>High</td>
<td>High</td>
<td>Average for optical, poor for UV</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Yes</td>
</tr>
<tr>
<td>Other (osmolarity, gas, pressure)</td>
<td>EnvZ osmolarity sensor, Nitrogen</td>
<td>Average</td>
<td>Low</td>
<td>Poor</td>
<td>Low</td>
<td>Low</td>
<td>Average</td>
<td>High</td>
<td>Yes</td>
</tr>
</tbody>
</table>
1.3. Optogenetic Control of Gene Expression in bacteria

Light induction as discussed above can potentially be a powerful tool for gaining dynamic control over genetic networks, something that has been conveyed by the applications and circuits that have been proposed in bacteria during the last decade. Starting with the first synthetic light receptor in *E. coli* (Leyskaya *et al.*, 2005), the same research group demonstrated how multiple such receptors can be used to gain multichromatic control within the same host (Tabor *et al.*, 2011) and more recently a blue responsive TCS by Möglich *et al.* (2009). The ultra fine resolution of light induction allowed implementations of the ‘game of life’ on a membrane of a single cell (Leyskaya *et al.* 2009) and an edge detection program in bacteria by combining quorum sensing and light induction (Tabor *et al.*, 2009). In addition, a light inducible closed feedback control loop was implemented first in yeast and then in bacteria by Milias-Argeitis *et al.*, (2011). A more recent publication demonstrated how a two state model can compute light induction patterns for living cells to perform pre-defined temporal programs with unprecedented accuracy, even with an open-loop system (Olson *et al.*, 2014). In the same work the authors harness the advantages of this fine temporal control to demonstrate how a sinusoidal input light function can be used to perform frequency analysis for characterising the response characteristics of genetic circuits. The various light input systems that have been used in a synthetic biology context are discussed below.

1.3.1. Two Component System Light Receptors

Across all kingdoms of life it was realised that sensory proteins especially the ones found in TCS share very conserved domains called Per Arnt Sim (PAS) (Möglich *et al.*, 2010; Möglich & Moffat, 2010). These domains allow for ‘interchangeability’ between the outer sensory domain and the intracellular actuator *i.e.* a Histidine Kinase (HK) or Phosphatase (PA) domain. Using this, Levskaya *et al.*, (2005) demonstrated the first synthetic light receptor functioning in *E. coli*. The receptor was a chimeric protein consisting of the intracellular HK from the native EnvZ osmolarity sensor and the outer membrane light sensory domain from Cph1 light receptor from *Synechocystis* sp. The Cph1 domain function relies on light absorption by a small molecule called Phycocyanobilin (PCB) which acts as the chromophore. The particular chromophore is produced enzymatically from heme, by heme oxygenase 1 (Ho1) and Feredoxin: Phyconyanobilin Oxydoreductase (PcyA), enzymes that were previously expressed in bacteria by Gambetta & Lagarias (2001). By default, the holophytochrome (phytochrome with the chromophore bound) in its ground state is auto-phosphorylated by ATP on its HK domain, consequently binding and phosphorylating OmpR, the natural osmoregulator of *E. coli*. Upon red light absorption the chromophore is excited causing a conformational change such that the auto-phosphorylation by ATP is no longer possible, therefore the phosphorylated levels of OmpR fall. Phosphorylated OmpR (OmpR-p) is responsible for regulating the expression of two porins, OmpC and OmpF, transcribed under the PompC and PompF promoters, respectively. At low levels of OmpR-p, only PompF is
active. As OmpR-p levels increase, PompF is repressed but transcription from PompC is upregulated. The engineered light TCS exhibited a 6-fold increase in expression activity between dark and light (Leyskaya et al, 2005). The CcaS green light receptor was also successfully cloned from Synecocystis sp. and expressed in E. coli. Its cognate regulator is CcaR which acts as an activator for the Pcpc-G2 promoter. CcaS phosphorylates CcaR upon green light illumination while red light drives it to a state that can actively de-phosphorylate CcaR-p via its phosphatase domain. Conveniently, for synthetic biology purposes, the particular light receptor uses the same chromophore (PCB) as Cph8, so that expressing the necessary enzymes to produce it can facilitate the operation of both receptors within the same host. The system exhibited a 1.5-fold difference when co-expressed with Cph8 (Tabor et al, 2011) and similarly Cph8 showed a lower 2.5-fold induction compared with its isolated version. Both TCS systems are shown diagrammatically in Figure 1.3.1 and are discussed more extensively in Chapter 3.

A very similar approach was more recently followed by Möglich et al, (2009) to re-engineer the FixL sensor and its cognate response regulator FixJ, involved in the nitrogen starvation and nitrogen fixation regulation of Bradyrhizobium japonicum. The authors replaced the FixL heme-binding sensory domain with the Light-Oxygen-Voltage (LOV) domain from YtvA (discussed below) found in Bacillus subtilis via their PAS homologous parts. The resulting synthetic sensor YF1, showed a ~1000-fold net HK activity rise and when expressed in E. coli exhibited a 70-fold induction of LacZ, expressed from a PFixK2 promoter which is regulated by FixJ-p. Like Cph8, the active state of YF1 is in the dark, while blue light (430 nm) forces a drop in the HK activity.

1.3.2. Alternative Light Based Gene Regulation Mechanisms

YcgF is an alternative blue light receptor found in E. coli and can be potentially used for regulating synthetic genetic networks (Tschowri et al, 2009) (Figure 1.3.2 a). The intermediate regulator YcgE is responsible for regulating YcgZ along with multiple genes involved in the biofilm formation mechanism (Tschowri et al, 2012). This system consists of a receptor and an intermediate regulator. However, this is not a two component system by the classic definition as it does not rely on the phosphorylation of any of its components. Instead the YcgF receptor acts as a direct anti-repressor by binding YcgE through its MerR-like domain, depleting its free pool and preventing access to the DNA binding sites. YcgF photosensitivity relies on the BLUF, flavin adenine dinucleotide (FAD) binding domain. YtvA from Bacillus subtilis can sense UV and blue light through LOV its domain (Möglich et al, 2007) shown in Figure 1.3.2(b). Although the mechanism is not entirely clear, it was shown that YtvA acts as part of the environmental stress response pathway, signalling through the RsB family of regulators that finally activate or deactivate the σB transcription factor (Gaidenko et al, 2006).

Photochemical induction, shown in Figures 1.3.2 (c-e) was previously used to control gene expression (Gardner & Deiters, 2012). It involves a photosensitive chemical group able to 'cage' and inactivate other regulating molecules. These can be protein regulators, other small molecules like chemical inducers or nucleic acids. Upon light exposure the photocaging molecules are cleaved from their targets restoring their biological activity.
Figure 1.3.1. The TCS light receptors. (a) The Cph8 receptor has the HK domain of EnvZ and the light sensory domain of Cph1 and (b) the CcaS light receptor. Both Receptors use the same chromophore, PCB, which can be produced enzymatically from Heme. The enzymes responsible for this are Ho1 and PcyA. OmpR can be phosphorylated and activate the PompC promoter when Cph8 is at its ground state, while upon red light illumination the HK activity is OFF. The CcaR regulator can be phosphorylated by CcaS upon green light illumination and activate the Ppcc-G2 promoter. Both systems were shown by Tabor et al, (2011) to be functional within the same host.
For protein based systems, the photocaging molecule is normally targeted onto non-natural amino acids such as photocaged analogs of tyrosine, cysteine and serine. When these acids are incorporated into the catalytic site of an enzyme they can reduce or block its activity, until light removes the caging groups. Another example is a photocaged T7 polymerase where it was shown to be completely inactive, unable to transcribe from the T7 promoter (Chou et al., 2010). Upon a brief UV exposure the ortho-nitrobenzyl (ONB) group caging a tyrosine in the T7Pol active site was released and its activity was restored. Examples of photocaged small molecules include doxycycline which can induce TetR (Cambridge et al., 2009), erythromycin (Gardner et al., 2011) and IPTG (Young & Deiters, 2007). More recently a photocaged IPTG molecule was demonstrated to give precise triggering over bacterial gene expression regulated by LacI (Binder et al., 2014). The latter was based on caging the IPTG molecule with a 6-nitropiperonal (NP) group that was relatively simple to be chemically synthesised. Upon UV irradiation the NP group was cleaved off, releasing the biologically active IPTG residue that binds and inactivates LacI, releasing repression of a GFP expressing T7 based promoter. Finally various strategies using photocaged nucleotides have been proposed in vitro (Pinheiro et al., 2008, Tang et al., 2010, Prokup et al., 2011) and limited examples in vivo in mammalian cells (Govan et al., 2011, Ceo & Coh, 2012).

Apart from photoactive groups there are examples of proteins or small molecules that are naturally sensitive to light. This property can be harnessed to achieve photo regulation on synthetic genetic networks. Such is the λCI repressor which is sensitive to UV light. As part of a genetic toggle switch (a bistable system discussed extensively in Chapter 5) Kobayashi et al., (2004) showed how UV irradiation was able to flick the switch towards the LacI site. Anhydrotetracycline (aTc) which is used as an inducer for TetR, is known to be a relatively unstable molecule under UV or blue light exposure (López-Penalver, 2010). Hence TetR repression activity can be potentially recovered when an aTc containing host is illuminated.

Although photocaging strategies were shown to be fast and precise without the need of expressing large genetic constructs, they lack of reversibility. Once the caging molecule has been cleaved, it is not possible to recage the target molecule when light is removed, as opposed to the two component systems discussed above. A promising direction towards reversible analogs is discussed by Krauss et al. (2011) where photosensitive cross linkers (like azobenzene derivatives) can be used and cause a photo-isomerisation on specific enzymes. However there are limited implemented examples and applications of this type up to this date. Such examples are a photoswitchable ribonuclease (Hamachi et al., 1998), ATPase, (Yamada et al., 2007) and a peroxidase (Muranaka et al., 2008).

1.4. Motivation and Objectives

In short, the current project was motivated by the prospect of applying a forward engineering framework approach to understand in depth, characterise and re-engineer the TCS light input modules to achieve control, in a predictable manner, over previously proposed circuits like the toggle switch (Gardner et al., 2000) and the...
Figure 1.3.2. Other optogenetic regulation strategies. (a) The native E. coli repressor (YcgE)-antirepressor (YcgF) blue light receptor (Tschowri et al, 2012). (b) The B. subtilis stress response receptor YtvA (Möglich et al, 2007). (c) A Photocaged IPTG molecule (Binder et al, 2012). (d) Photodegradation of light sensitive molecules like aTc. (e) The photocaged T7 polymerase (Chou et al, 2010). (f) Photocaging groups (NPP) inhibiting mRNA silencing. UV irradiation cleaves NPP and the siRNA blocks translation (Mikat & Heckel, 2007). (g) NPOM photocaging groups prevent the formation of dsDNA binding sites that can act as a decoy for transcription factors. Upon light exposure the dsDNA is formed and the transcription factor (TF) nF-kB is aggregated on the decoy binding sites depleting the free TF pool that can activate the actual promoter (Govan et al, 2011).
dual feedback oscillator (Stricker et al, 2008)

The project was structured on two levels. First I attempted to use the engineering cycle, which is yet to be fully exploited for synthetic biology, and identified some of the limiting factors regarding the current methodologies that prevent its full utilisation, as opposed to other engineering disciplines. I attempted to design and derive a model for the system of interest. I explored the potentials of a mechanistic modelling approach in order to understand and identify possible biological dials that can affect the performance of the proposed systems and also use a rigorous model reduction approach to help parametrise and fit the model to the experimental data. Finally the implementation of the physical system by obtaining, cloning and assembling the genetic parts according to the initial design followed by their experimental characterisation. By feeding some of the data back to the model I attempted to refine its predictive abilities.

The second parallel level is structured in regards to the advantages of using light as an inducer, as discussed above. In this work, I aimed to characterise some of the proposed light input systems i.e. Cph8 and CcaR in order to better understand their function and predictably use them to achieve control, over two well known synthetic genetic networks. To achieve that, I attempted deriving a mechanistic model (which has not been proposed to date) and used it to identify potential ways for improving the relatively small fold induction that was reported on Cph8 and CcaR TCS (Levskaya et al, 2005; Tabor et al, 2011). Next, I designed and implemented an efficient and cost effective, in terms of burden and response time, ‘light and chemical integrator’ module for coupling the TCS with other circuits without the need of expressing additional connecting components. Finally, by using a forward engineering approach, based on the developed models and constructs, I attempted to re-engineer the toggle switch as an example of a static system where light can potentially shift the stable equilibrium points, thus, modulate the way that the switch is driven to one state or the other. The second system that was proposed to be controlled by light is an inherently unstable dynamic circuit i.e. an oscillator, where a very intriguing prospect was to use one light wavelength to tune the frequency and a different one to tune the amplitude of the oscillation. To the best of my knowledge no such genetic oscillator where frequency and amplitude are decoupled, exists to date.

In more detail the project is split between four subprojects or 'modules' where the outcome of each of them is designed to be a 'plug and play' physical circuit accompanied by the corresponding theoretical work. The modules and circuit complexity is progressively scaled across chapters, starting from the individual light receptors, to developing a novel, to some extent, coupling circuit of the TCS with other networks and finally the implementation of a light tuneable toggle switch and proposed design of the light tuneable oscillator. The module objectives are shown more analytically below:

**Module 1 - Light Receptors (Chapter 3)**

**Theoretical work** Derive a mechanistic NL ODE model for the Cph8, CcaS and multichromatic control system from first principles. Reduce rigorously the model based on time scale separation or other appropriate assumptions, in order to eliminate as many time variables and parameters as possible. Populate the parameter set with literature values, or fitted values based on available experimental data. Perform sensitivity analysis
and stochastic simulations

Experimental work Build or obtain a suitable light induction platform. Assemble and characterise the static and dynamic performance of the original (Levskaya et al., 2005 and Tabor et al., 2011) and reconstructed systems.

Module 2 - Coupling Module (Chapter 4)

Theoretical work Design and derive a model for an integrator device where the light receptor TCS output can be connected with the downstream circuits in the most effective way. The model can be an extension of what was derived from module 1. Fit the model with data once they become available.

Experimental work Build and characterise the designed device for both static and dynamic performance and assess whether it serves its design purpose.

Module 3 - Light Tuneable Toggle Switch (Chapter 5)

Theoretical work Re-design and derive a model for a toggle switch (Gardner et al., 2000) where the light input module can be used as an input. The model should be an extension of the modelling framework derived in Module 1 and 2 so that the parameter values obtained can be reused. The model should be fitted for further parametrisation should data become available.

Experimental work Assemble and characterise the light tuneable toggle switch. Assessment should include bistability assessment, timer response and chemical induction shifts versus light intensity.

Module 4 - Light Tuneable Oscillator (Chapter 6)

Theoretical work Re-design and derive a model for the dual feedback oscillator (Stricker et al., 2008) where light can be used to control its frequency and/or its amplitude independently. Again, the light input module should be based on what was derived from modules 1-2 and coupled with the published mechanistic model by Stricker et al. (2008). The published model should be (re-)evaluated in order to identify possible re-designs that allow for independent tuning of the amplitude and the frequency. Perform sensitivity analysis and stochastic simulations on the final model to assess the predicted robustness of the system. If the model outcome yields strong indications of a robust system proceed on the experimental implementation.

Experimental work If time permits, build a suitable platform to observe single cell dynamics using light as an input (e.g. a microfluidics platform). Obtain the Stricker et al. (2008) oscillator plasmids and modify them based on design. Assess and analyse the system using single cell tracking software.
In this chapter I present the methodology followed for the experimental and theoretical aspects of this work. Protocols and wet-lab methods are given in a general context as to what was used in the majority of cases but also details on the individual experiments. More specific details for particular constructs or assays are presented in the results Chapters 3-6, to give the reader a more direct picture of the data presented. Finally, tables showing all the bacterial strains, plasmids and other genetic parts, oligos and solutions used are given.
Chapter 2

Materials and Methods

2.1. Experimental Methods

2.1.1. Cloning and DNA Assembly

Throughout this work all of the required sequences were obtained from pre-existing plasmids with the exception of the downstream part of the Plac/ompR promoter (Chapter 4) which was synthesised by GenScript (www.genscript.com). After obtaining the required sequences they were either cloned, transformed and purified from a DH5α strain or amplified by Polymerase Chain Reaction (PCR) using the appropriate oligonucleotides shown in Table 2.1.1. The cloned sequences were assembled into the desired constructs by either restriction digestion followed by ligation or in some cases by In-Fusion® (Zhu et al, 2007) or Gibson assembly (Gibson et al, 2009). Specific small sequence modifications like RBS or promoters were introduced as overhangs in the corresponding oligonucleotides used for the initial amplification of parts. Below the standard protocols followed are given, while specific information about the construct architecture and exact cloning workflow is reported in the result chapters (3-6) for each construct.

Restriction Digestion Protocol

For all restriction endonuclease reactions the protocol, according to the supplier’s guidelines is shown in Table 2.1.1

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Cloning</th>
<th>Screening</th>
</tr>
</thead>
<tbody>
<tr>
<td>10X Buffer</td>
<td>5 μl (Final 1X)</td>
<td>1 μl (Final 1X)</td>
</tr>
<tr>
<td>BSA*</td>
<td>5 μl (Final 1X)</td>
<td>1 μl (Final 1X)</td>
</tr>
<tr>
<td>DNA</td>
<td>1 μg</td>
<td>250 ng</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>Up to the final volume</td>
<td>Up to the final volume</td>
</tr>
<tr>
<td>Restriction Endonuclease</td>
<td>1 μl (10 Units)</td>
<td>0.2 μl (2 Units)</td>
</tr>
<tr>
<td>Final Reaction Volume</td>
<td>50 μl</td>
<td>10 μl</td>
</tr>
</tbody>
</table>

* If required and not included in the buffer

Incubation temperature was normally at 37 °C unless stated otherwise by the supplier for the specific restriction enzyme. Incubation time was approximately 1.5 h for screening digestions and 2.5 h for cloning purposes. For all digestions carried out on linear PCR amplified DNA the incubation time was increased to 3.5 h. After incubation, the samples were normally run on an agarose gel for electrophoresis for purification. Otherwise the enzymes were heat inactivated according to the manufacturers guidelines (typically 60-70 °C for 30 min). All restriction enzymes were purchased from New England Biolabs (NEB) with some exceptions that were purchased from Promega.
Ligation protocol

Ligation reactions were carried out by using the standard T4 Ligase (NEB) for a slow overnight or rapid (Promega) on bench incubation. In both cases the standard protocols shown in Table 2.1.2. were very similar.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>T4 Ligase (NEB) Overnight</th>
<th>T4 Ligase (Promega) Rapid</th>
</tr>
</thead>
<tbody>
<tr>
<td>10X Buffer / 2X Rapid Ligation Buffer</td>
<td>1 μl</td>
<td>5 μl</td>
</tr>
<tr>
<td>Vector DNA</td>
<td>50 ng</td>
<td>50 ng</td>
</tr>
<tr>
<td>Insert DNA</td>
<td>Variable</td>
<td>Variable</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>Up to the final volume</td>
<td>Up to the final volume</td>
</tr>
<tr>
<td>T4 DNA Ligase</td>
<td>1 μl (1 Unit)</td>
<td>1 μl (1 Unit)</td>
</tr>
<tr>
<td>Final volume</td>
<td>10 μL</td>
<td>10 μL</td>
</tr>
</tbody>
</table>

Incubation was carried out at 16 °C overnight for the T4 DNA ligase from NEB, while for the rapid ligation protocol the incubation was at room temperature for 30 minutes. For all ligations 50 ng of vector was used and a 3:1 molar ratio for the insert to vector was calculated with the formula (2.1):

\[
\text{Insert (ng)} = \frac{3 \times \text{Vector (ng)} \times \text{Size of Insert (bp)}}{\text{Size of vector (bp)}} \quad (2.1)
\]

Polymerase Chain Reaction (PCR) Protocol

Unless otherwise specified for all PCR reactions the High-Fidelity Phusion® DNA polymerase was used according to the manufactures protocol in Table 2.1.3.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Cloning</th>
<th>Screening</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffer HF / GC 5X</td>
<td>10 μl</td>
<td>4 μl</td>
</tr>
<tr>
<td>dNTPs (10 mM)</td>
<td>1 μl</td>
<td>0.4 μl</td>
</tr>
<tr>
<td>DMSO*</td>
<td>1.5 μl</td>
<td>0.6 μl</td>
</tr>
<tr>
<td>Forward Primer (10 μM)</td>
<td>2.5 μl</td>
<td>1 μl</td>
</tr>
<tr>
<td>Reverse Primer (10 μM)</td>
<td>2.5 μl</td>
<td>1 μl</td>
</tr>
<tr>
<td>Template DNA</td>
<td>Variable</td>
<td>Variable</td>
</tr>
<tr>
<td>Sterile ddH₂O</td>
<td>Up to the final volume</td>
<td>Up to the final volume</td>
</tr>
<tr>
<td>Phusion® DNA Polymerase</td>
<td>0.5 μl</td>
<td>0.2 μl</td>
</tr>
<tr>
<td>Final volume</td>
<td>50 μl</td>
<td>20 μl</td>
</tr>
</tbody>
</table>

*DMSO was used to prevent secondary structure formation.
The template amount used was varied with respect to its total size and the size of the targeted amplicon. e.g. For a typical 3 kb vector and 1 kb targeted amplicon the amount used in the reaction was 1-10 ng. This amount was increased up to 100 ng in some cases where the PCR reaction was inefficient, returning nonspecific bands or very low yield of amplicon. In some cases for screening a large number of clones, colony PCR was employed. The template in this case was approximately 1 μl of the supernatant solution after heat lysis and centrifugation of a colony or a bacterial pellet. The thermal cycle program used was:

**Table 2.1.4.** Thermocycler program for PCR reactions

<table>
<thead>
<tr>
<th>Thermal Cycler Step</th>
<th>Temperature °C</th>
<th>Step Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Denaturation</td>
<td>98</td>
<td>30 s</td>
</tr>
<tr>
<td><strong>25 - 35 Cycles</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denaturation</td>
<td>98</td>
<td>10 s</td>
</tr>
<tr>
<td>Annealing</td>
<td>Variable</td>
<td>30 s</td>
</tr>
<tr>
<td>Elongation</td>
<td>72</td>
<td>30 s kb⁻¹</td>
</tr>
<tr>
<td>Final Extension</td>
<td>72</td>
<td>10 min</td>
</tr>
<tr>
<td>Storage</td>
<td>4</td>
<td>N/A</td>
</tr>
</tbody>
</table>

The annealing temperature varied between 45-65 °C and was set according to the predicted melting temperature (Tm) for each oligo. For optimisation purposes gradient PCR was employed using a series of annealing temperatures 5-10 degrees plus or minus from the melting temperature (Tm). All primers (oligonucleotides) were ordered and synthesised by either Invitrogen or Sigma Aldrich. Table 2.3.5 shows all of the oligos used in this work.

**In-Fusion® Cloning**

The vectors and inserts were amplified from a PCR amplification reaction using oligos such that both parts contain matching linker sequences as overhangs. These sequences were either introduced as non-coding neutral sequence fragments (Casini *et al*, 2014) via the oligos used for the PCR or the actual sequence of the target vector or insert was used. The length of the linker sequence used was 15 bp on either side. For all the In-Fusion® (Clontech) cloning reactions the protocol in Table 2.1.5 was used

**Table 2.1.5.** In-Fusion® Cloning reaction mix

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>5X In-Fusion® Reaction Buffer</td>
<td>2 μl</td>
</tr>
<tr>
<td>Vector DNA</td>
<td>100-400 ng</td>
</tr>
<tr>
<td>Insert DNA</td>
<td>50-200 ng</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>Up to the final volume</td>
</tr>
<tr>
<td>In-Fusion® Enzyme</td>
<td>1 μl</td>
</tr>
<tr>
<td><strong>Final volume</strong></td>
<td>10 μL</td>
</tr>
</tbody>
</table>
The reaction mix was then incubated for 15 min at 37 °C, followed by 15 min at 50 °C. The reaction mix was then diluted up to 50 μl volume with Tris-EDTA (TE) buffer (pH 8.0) and 1-2 μl were used for transformation.

**Plasmid DNA Purification**

For all plasmid DNA extraction the peqGOLD Plasmid Miniprep I kit was used. The procedure followed was according to the manufacturer’s guidelines. 5 ml cultures were grown overnight. For low copy number plasmids 10 ml of total culture was used.

**PCR Purification / Reaction Clean-up / DNA extraction from Agarose Gel**

All PCR purifications, agarose gel DNA extractions or reaction clean-ups were performed using Promega’s Wizard® SV PCR and gel clean-up kit according to the manufactures guidelines. In some cases where smaller elution volumes were needed in order to obtain higher concentrations of DNA the Qiagen Minelute reaction clean-up kit was used, according to the manufactures guidelines.

**DNA Concentration and Purity Determination**

Plasmid DNA concentration and purity were determined based on the 260/280 nm and 260/230 nm absorbance ratios measured on a Thermo Scientific NanoDrop™ 1000.

**DNA Sequencing**

For all DNA sequencing the prepaid voucher single read service from MWG Eurofins Operon or Source Bioscience was used. Samples were prepared according to the service's guidelines. Typical procedure involved the overnight growth of single colonies followed by plasmid purification. Typical plasmid concentration used for sequencing was 100 ng/μl.

**Agarose Gel Electrophoresis**

For all agarose gel electrophoresis runs, the molecular grade agarose (Sigma Aldrich) was used at 1% or 0.8% w/v in 1X Tris Acetate EDTA (TAE) buffer for up to 4 kb and up to 10 kb fragments, respectively. The gel agarose solution was supplemented with GelRed DNA stain (Cambridge Bioscience) for screening purposes or SYBR® safe (Invitrogen) for gel purification purposes. The samples were run for approximately 50 min at 100 V in a solidified gel and imaged on a UV imager (Bio-Rad) equipped with a camera or a blue box transilluminator (Clarechemical) for DNA band excision.
**Electrocompetent Cells Preparation**

Most transformations throughout the entire project were carried out by electroporation with competent cells prepared fresh on the day of the transformation, since the protocol for preparing them was fast and yielded high transformation efficiencies with the freshly prepared cells. In some cases the DH10b Gold (NEB) efficiency chemically competent cells were used.

The required strains were grown in 5 ml cultures in LB with appropriate antibiotics, overnight at 37 °C with shaking at 235 RPM. The cultures were then diluted 1:100 in fresh LB medium with appropriate antibiotics in 15 ml round bottom tubes to a final volume of 10 ml and left to grow at 37 °C with shaking (235 RPM) up to an OD$_{600}$ of 0.6.

The cells were pelleted by spinning at 4000 RPM for 20 min in a precooled centrifuge at 4 °C. The supernatant was quickly removed and tubes were briefly inverted on blue roll paper to remove most of the remaining media. The tubes were transferred in an ice-slurry waterbath and the pellets were resuspended in 1 ml of sterile ddH$_2$O by gently shaking on a bench-top shaker. After complete resuspension of the pellet the tubes were filled with sterile ddH$_2$O up to 10 ml and centrifuged for another 15 min at 4 °C at 4000 RPM. The cells were resuspended and washed once more with sterile ddH$_2$O. After the final spin the tubes were inverted on blue roll under sterile conditions for 20-30 s to remove as much from the remaining water from the tube walls as possible. The remaining pellet was resuspended in the residual water which amounts to approximately 50 μl for each 15 ml tube. The cells were stored on ice until transformation.

The process can be carried out with larger volume cultures and by replacing ddH$_2$O with sterile 50% glycerol the electrocompetent cells can be stored in a -80 °C freezer to be used on a different day. Summarising the protocol requires 2-3 washes with sterile ddH$_2$O in order to remove all the salts while keeping the cultures at or below 4 °C. Sterile conditions were needed throughout the execution of this protocol.

**Transformations**

**Transformation by Electroporation**

25 - 50 μl of electrocompetent cells were mixed with 0.5 - 1.5 μl of ligation mix (depending on the ligation buffer used) or plasmid DNA and transferred in a chilled 0.2 mm Bio-Rad electroporation cuvette followed by a single pulse in a Bio-Rad electroporator under the bacteria electroporation preset of the machine. Typical average time constants observed were between 4.2 ms to 5.5 ms. The cells were then diluted up in fresh LB broth or SOC medium and incubated for 1 hour at 37 °C with shaking (235 RPM).

**Transformation by Heat Shock**

25 - 50 μl of chemically competent cells were combined with 1 - 3 μl of ligation reaction or plasmid DNA and incubated on ice for 30 min. Next, the cell-DNA mixture was heat shocked at 42 °C for 30 s and left on ice for 2 min before adding fresh Super Optimal Broth (SOC) outgrowth medium or LB broth up to 1 ml and incubated for 1 hour at 37 °C with shaking (235 RPM).
Following either transformation protocol after the 1 hour incubation the cells were plated on pre-warmed LB-agar plates with appropriate antibiotics and incubated overnight at 37 °C. Typically, two plates were used per transformation for plating a small (5 μl) and a large volume (300 μl) of the recovered cells. Where the transformation efficiency was expected to be low the entire 995 μl remaining volume of the recovery culture was pelleted, resuspended in 50 μl of LB and plated.

2.1.2. Experimental Assays

Cph8 TCS and CcaS TCS LacZ Based Characterisation

Single colonies were picked from a freshly streaked plate and grown overnight. 1:100 dilutions were prepared the following day and the cultures left to grow up to an OD_{600} of 0.6. Each strain and construct combination was prepared in biological triplicates. A new 1:10 dilution up to 6 ml in 12 ml round bottom Eppendorf tubes which were split into two 3 ml replicas from which one was kept in the dark and the second was placed under continuous illumination with a 635 nm LED and grew for 4 hours. It was estimated empirically that within this time frame the cells have time to grow and reach a plateau, regarding the LacZ expression without reaching the stationary phase. In addition, keeping a low volume in the growing culture aims to keep the cells in aerobic growing conditions as much as possible since the Ho1 mediated reduction of Heme to biliverdin (BV) requires oxygen. The intensity based on the distance from the light source was estimated to be approximately 0.21 W/m², which is in excess of the published value of 0.1 W/m² where the light receptor is saturated at the P^o (inactive) state. For the static response assays 1.5 ml of samples were collected after the 4 hours of exposure in different light conditions, centrifuged and subjected to the Miller assay procedure discussed below. For assays that were conducted the following day, the samples were resuspended in 50% glycerol and stored at -80 °C.

For the dynamic response assays a similar culture incubation method to the static response assays was followed, but in this case the samples were collected at regular time intervals. The second dilution was done at a 1:10 ratio to bring the optical density of the culture down to 0.6. After four hours of growth, the cultures kept in the dark where transferred to continuous red light illumination and the ones growing in the presence of light were switched to the dark. Samples were collected every 1 hour for the first three hours of growth and then every 30 minutes for one hour before and after the switching point. For each time point, 1.5 ml of sample was transferred into a new 2 ml Eppendorf tube, centrifuged and re-suspended in 50% glycerol solution. This allowed storage of the samples at -80 °C since the activity assay was conducted on the following day (with no detectable loss of LacZ activity). During growth before switching, the cells were diluted twice at a 1:5 ratio at t = 2 h and t = 4 h in order to keep the cells at their exponential phase of growth. An additional dilution was carried out after the switching point at t=6 h.

Miller Assays

For all LacZ activity measurements a version of the standard Miller Assay protocol found in Levskaya et al,
(2005) was used. For the activity assay the samples were thawed on ice, pelleted and resuspended in 1 ml of Z buffer. 100 μl of each sample were transferred in a 96 well microplate for determining the OD600. The samples were then transferred in a new Eppendorf tube and diluted in Z-buffer supplemented with β-mercaptoethanol (BME) such as the final OD600 is 0.5. This, was found empirically to give a sensitive, but slow enough reaction to be quenched in a convenient time frame (approximately 4-10 minutes). Next 100 μl of chloroform and 50 μl of 0.1% SDS were added and mixed by vortex in order to lyse the cells. The samples were then incubated for about 5 minutes at room temperature. The lysis step was followed by the addition of 200 μL of 4mg/ml ONPG. The time for the ONPG addition into each sample was recorded.

After approximately 2-10 min a faint yellow colour appears (depending on the LacZ abundance) which indicates that the reaction has proceeded enough such that ONP levels become detectable by the microplate reader. At this point the addition of 500 μl CaCO₃ quenches the reaction but also helps to develop further the characteristic yellow colour for ONP. The time of the quenching was also recorded. Next the samples were centrifuged for 5 minutes at 13 000 RPM so that cell debris and the chloroform phase is separated from the aqueous fraction. 200 μl from the aqueous phase were transferred to a 96-well microplate and the absorbance at 400 nm and 520 nm was recorded.

Miller Units were calculated by the formula (2.2):

\[
MU = \frac{1000 \cdot (OD_{400} - 0.075 \cdot OD_{520})}{V \cdot dT \cdot OD_{600}}
\]  

(2.2)

where V is the volume (ml) used, of the cells resuspended in Z buffer with BME, dT the time difference (min) between the addition of ONPG and CaCO₃ quenching, while OD400, OD520 and OD600 are the optical densities measured as absorbance at 400, 520 and 600 nm respectively. It is important to note that for OD600 the value used was the one recorded from the resuspended culture, prior to the lysis step, while for 400 and 520 nm the values measured after the reaction were used.

**Oxalic Acid Based Heme Quantification**

This method is based on the oxalic acid mediated reduction of heme to protoporphyrin which has a characteristic fluorescent spectrum with maxima at 600, 617 and 670 nm (Schwartz, et al. 1983).

Overnight samples were diluted to an OD₆₀₀ of 0.1 in 5 ml cultures with antibiotics and grown in LB at 37 °C with shaking at 235 RPM. 1.7 ml samples were collected at regular time points. 200 μl from each sample was used to measure the OD₆₀₀ using the microplate reader. The remaining 1.5 ml of the samples were pelleted and resuspended in 0.75 ml of hot oxalic acid (2M) followed by a 40 min incubation on a dry block set at 90 °C. Then the samples were left to cool down to room temperature so that the oxalic acid can be precipitated by pipetting up and down. The precipitated oxalic acid was removed by centrifugation at maximum speed (13400 RPM) for 10 min and 500 μl of supernatant were transferred in a new 1.5 ml Eppendorf tube. The fluorometric assay was carried out in a FluoroMax-3 spectrofluorometer (JY-Horiba).
The excitation spectrum range that was obtained varied from 300 to 500 nm with an increment of 1 nm and the emission wavelength used was at 600 nm. By exciting at 400 nm the emission spectrum was obtained from 450 up to 750 nm with 1 nm increment.

**Acetic Acid Based Heme Assay**

Although the oxalic acid based assay showed to be a sensitive assay, it had narrow linear range between fluorescence and heme concentration. Furthermore, it was impractical for running a large number of samples. For quantification purposes the acetic acid based modified method was used (Van Den Berg *et al.*, 1988). This method allowed for measuring a larger amount of samples using the microplate reader and in addition it does not involve precipitation of any reagent that can render the solution cloudy.

**Sample preparation:** Overnight cultures grown at 37 °C with shaking (235 RPM) and in the presence of appropriate antibiotics were centrifuged and washed with ddH2O twice, in order to remove any traces of Heme in the extracellular medium. The pellets were resuspended in 50 μl of water, weighed and corrected for residual amounts of water and combined with 450 μl of 5:1 Isopropanol to HCl. The mixture was vortexed and centrifuged for 15 min at 10000 RPM. The supernatant was then transferred in a new 1.5 ml Eppendorf tube.

**Standards preparation:** Purified Heme chloride (Sigma Aldrich) was weighed and diluted to give a final concentration of 10 mM in NaOH (0.5 M) aqueous solution. Once Heme was in solution it was diluted further down to a stock concentration of 10 μM in isopropanol.

**Assay:** 50 μl of standards or samples were combined with 500 μl of glacial acetic acid in a new 2 ml Eppendorf tube. Next 100 μl of Fe.SO4 (0.12 M) and HCl (4.5 M) solution was added, mixed well with the samples by vortexing and left to incubate in a waterbath set at 60 °C for 30 min. At the end of the 30 min incubation the solution develops a distinct yellow tint. 1 ml of 1:1 ddH2O and isopropanol was then added and mixed with the sample. 200 μl from each sample were transferred on a 96 well-plate for fluorescent measurements. The measurements were carried out in a Gemini EM plate reader/fluorescent scanner. The excitation wavelength used was 400 nm and 597 nm for emission. Alternatively, fluorescent excitation-emission spectra were recorded for the range between 500-750 nm with 1 nm increments.

**Microplate Based GFP, mCherry and eCFP Spectra and Single Wavelength Readings**

For obtaining the eCFP spectrum, 25 ml cultures grown overnight in 50 ml Falcons were centrifuged and resuspended in 1X TAE or PBS buffer and lysed by sonication for cell wall disruption. Sonication parameters were 4X30 s pulses @ 20 μM amplitude and 2 min intervals for 6 min. The samples were kept on ice throughout the process. The samples were then centrifuged for 45 min at 4000 RPM at 4 °C. 1000 μl of the supernatant was transferred to a 1 ml cuvette and fluorescent excitation and emission spectra were recorded in a FluoroMax-3 spectrofluorometer (JY-Horiba) equipped with a monochromator. The increment used was 1 nm and the scanned range was 400-700 nm using 430 nm and 470 nm for excitation and emission spectra, respectively. Alternatively for live cell measurements, a Gemini EM plate reader (Molecular Devices)
equipped with a monochromator was used to obtain eCFP emission spectra scanned at 440-550 nm for eCFP (1 nm increments) using 433 nm for excitation. Similarly for GFPmut3b, the range scanned was 505-600 nm excited at 485 nm. For single wavelength reads for GFP or mCherry, 96-well plate assays were run on a Synergy™ HT platereader (BioTek) or POLARStar Omega platereader (BMG Labtech) plate reader using the excitation emission filter wavelengths shown in Table 2.1.6. For all samples the OD<sub>600</sub> was also recorded for normalisation. For eCFP and GFP the Gemini EM (Molecular Devices) was used according to Table 2.1.6, however the specific plate reader cannot measure absorbance or optical density, therefore all live cell samples measured for fluorescence were transferred for OD<sub>600</sub> measurements to one of the two other aforementioned plate readers.

Table 2.1.6. Excitation, emission and cutoff wavelengths used for plate reader based fluorescent protein measurements

<table>
<thead>
<tr>
<th>Protein</th>
<th>λ&lt;sub&gt;ex&lt;/sub&gt; (nm)</th>
<th>λ&lt;sub&gt;cutoff&lt;/sub&gt; (nm)</th>
<th>λ&lt;sub&gt;em&lt;/sub&gt; (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mCherry</td>
<td>585</td>
<td>600</td>
<td>610</td>
</tr>
<tr>
<td>GFPmut3b*</td>
<td>487</td>
<td>515</td>
<td>525</td>
</tr>
<tr>
<td>GFPmut3b</td>
<td>485</td>
<td>515</td>
<td>525</td>
</tr>
<tr>
<td>sfGFP</td>
<td>485</td>
<td>505</td>
<td>510</td>
</tr>
<tr>
<td>eCFP</td>
<td>433</td>
<td>475</td>
<td>477</td>
</tr>
<tr>
<td>eCFP*</td>
<td>420</td>
<td>475</td>
<td>477</td>
</tr>
<tr>
<td>Cph8:PCB</td>
<td>400</td>
<td>590</td>
<td>600</td>
</tr>
</tbody>
</table>

* Empirically estimated settings for minimizing the bleed-through fluorescence when GFPmut3b and eCFP are co-expressed

Normalisation, Baseline Correction and Error Propagation

For the normalisation of live cell fluorescence, the cell density values at 600 nm (OD<sub>600</sub>), measured using a Synergy™ HT platereader (BioTek) or POLARStar Omega platereader (BMG Labtech) and applying path length correction, were used according to (2.3). For Eppendorf or Falcon tube based assays, 100 μl samples were transferred to microplates and measured using the aforementioned platereaders.

\[
x_n = \frac{x}{OD_{600}} \quad (2.3)
\]

\[
x_c = x_n - b_n \quad (2.4)
\]

\[
\Delta x_c = \sqrt{(\Delta x_n)^2 + (\Delta b_n)^2} \quad (2.5)
\]
where $x_n$ is the normalised fluorescence value by the corresponding optical density of the cell population. $x_c$ is the corrected value of the latter by subtracting the baseline fluorescence $b_n$ (2.4) where, unless otherwise specified in the text, is a sample of a strain grown under the same conditions but does not express the specific fluorescent reporter. Autofluorescence of media was embedded in $b_n$. The combined uncertainty or standard deviation is given by $\Delta x_c$ (2.5). For the cases where a mean value is normalised by another mean (2.6) value the following formula (2.7) was used:

$$x_a = \frac{\langle x \rangle}{\langle a \rangle}$$

$$\frac{\Delta x_c}{\langle x_a \rangle} = \sqrt{\left(\frac{\Delta x}{\langle x \rangle}\right)^2 + \left(\frac{\Delta a}{\langle a \rangle}\right)^2}$$

### Flow Cytometer based assays

For flow cytometry, a FACScan cytometer (BD Biosciences) equipped with an Automated Multiwell Sampler (AMS) was used. Unless specified otherwise in the text, the standard procedure followed was as follows. Overnight cultures were diluted 1:100 and let to grow at 37 °C with appropriate antibiotics. This was followed by a second 1:10 dilution and grown under appropriate, assay specific conditions (chemical inducers, light, temperature, antibiotics and growth medium). Depending on the OD$_{600}$ during time course experiments an appropriate volume from each sample was diluted in 200 μl of ddH$_2$O into 96-well plates (Costar) and loaded on the AMS machine. For small number of samples, single tube reads were performed instead. The dilutions were such that the events per second recorded during the FACScan setup did not exceed 1500. The samples were scanned for Side Scatter (SSC), Forward Scatter (FSC), FL1 (for GFP) and FL2 (mCherry for the dual reporter light toggle switch). For all AMS or single tube reads a 30 s time or 150 000 events limit was set. For tube based reads the time limit was extended to 45 s.

FACScan data were collected and processed using CyFlogic v1.2.1 (http://www.cyflogic.com/) or FlowJo V10 (TreeStar http://www.flowjo.com/). All samples were gated based on the FSC and SSC dot plot in order to select the cells with regular size and shape, while discarding events corresponding to small dust particles or abnormal size and shape cells (that normally correspond to pairs of dividing cells that are accounted as a single event). For bimodal fluorescence distributions additional gates were used. Bimodality due to plasmid loss or mixed populations of dead or compromised cells is very common (Lehtinen et al., 2004) in FACScan data for *E. coli*, especially after long growth where the population is degrading the antibiotic fast, such as some cells can still survive even if they have ejected the plasmid. This population has a distribution peak which lies at lower levels than non induced cells. The statistics exported from the above analysis were the mean, geomean and variance of the fluorescence distribution.
Fluorescent Microscopy

For microscopy experiments a Ti Eclipse Inverted Fluorescent Microscope (Nikon) was used. The microscope was equipped with 4X, 10X, 20X and phase 60X Oil objectives, a motorised stage and fitted incubator. The filter sets for fluorescence included a GFP-B set with λex 480 nm (bandwidth 460-500 nm), λem 535 nm (BW 520-560 nm), a Cy3 λex 545 nm (BW 512-552 nm) λem 610 (BW 565-615 nm) and DAPI λex 360 nm (BW 340-380 nm) λem 460 (BW 435-485 nm). The sets were used for GFP, mCherry and CFP, respectively.

For single cell and microcolony imaging cells at their exponential phase of growth were transferred on agar pads. For agar pad preparation 1 ml of LB-agar with appropriate antibiotics or inducers was sandwiched between 2 microscopy slides and left to solidify. From the resulting LB-Agar strip, small bands were cut and 1 μl of cell culture of cells at exponential growth was placed on the pad and let to dry for 15 min at 37 °C. Next a cover slip was placed and nail varnish was used to seal the pad in order to prevent evaporation. For microcolony formation, the microscope’s incubator was set at 37 °C. For the light toggle switch images, the green fluorescence was recorded with 1 s exposure time using the GFP-B filter set and eCFP was recorded for 3 s exposure, using the DAPI filter set. The exposure times were determined empirically.

Plac/ompR and Plac/ccaR Promoter Characterisation

Standard sample preparation procedure like the one discussed in the Cph8/CcaS characterisation was used for all assays on the dual input promoters. Overnight cultures were grown overnight at 37 °C with shaking (235 RPM) in LB medium supplemented with Ampicillin and Chloramphenicol for reporter and TCS plasmids respectively. The samples were diluted 1:100 in fresh LB medium with appropriate antibiotics and grown for 2 hours. A second 1:10 dilution was followed and grown under different inducing conditions.

For the steady state response 14 ml tube based assay, triplicates of samples were grown under 4 different conditions, no induction, IPTG (1 mM) only, light only (>0.5 W/m²) or both. The samples after 4 h of growth were measured using the FACScan protocol described above. For the induction surface plots the samples were diluted into 96-well microplates and grown for 4 h in the light incubator shown in Figure 2.3.2 (b) by applying a gradient of light intensity (0-1W/m²) from the control software of the incubator and an IPTG 0-10 mM on the horizontal axis (columns 1-6). Half the microplate (Rows 7-12) was inoculated with control the control strains under identical light and chemical induction conditions. Finally, for the time course assay the cells were diluted at a 1:3 ratio at t=2 h, t=4 h and t=6 h during the assay in order to keep the cells in the exponential phase of growth. The measurements for each time point were performed using the FACScan machine.

Light Tuneable Toggle Switch Characterisation

For the light toggle switch assays, the standard overnight growth protocol was used for growing individual colonies picked from Petri dishes. Two dilutions, 1:100 and 1:10 before the beginning of all assays were performed. For all assays the overnight cultures were supplemented with 5 ng/μl aTc in order to initiate the assays from the low state, with the exception of the time course and stability assay where 1 mM IPTG was
used to push half the samples to the high state. The plate layouts used are shown in Figure 2.1.1. For the chemical induction response bar charts, layout C was used where 2 replica plates were prepared. One was kept in the dark and one at approximately 20 cm from the red LED device (~ I=0.5 W/m²). Layout B was used for the Hill fits as a function of IPTG, where one plate was inoculated with samples and another one with the control strains. Each plate was half covered with aluminium foil for growing cells in the dark. Layout A was used for the two dimensional induction assay, where light intensity was varied by using neutral density (ND) filters, however the samples in this case were inoculated in duplicates. Finally, layout D was used for the time course, and pulse assays. All samples where measured on the Gemini EM platereader for fluorescence and Synergy™ HT platereader for OD₆₀₀. For the light pulse assay the samples where normalised by their fluorescence at 650 nm (excited at 500 nm) since the Gemini plate reader cannot read optical density for the overnight part of the assay. The specific wavelengths were found empirically to be fairly proportional to cell density but not affected by the two fluorescent reporters. Finally, for all microplate based assays on pre-pierced film was used to seal the plate in order to prevent evaporation but allow good aeration during growth.

2.2. Theoretical and Computational Methods

2.2.1. Computational Modelling

All the mass-action or phenomenological ODE models discussed in the text were written, simulated and partially analysed using the MATLAB's application SimBiology v3 (MATLAB 2011b) and v4 (MATLAB 2013a). The application allows for building by populating a list of reactions in the general form A + B <-> C + D, a table of species (that can be automatically generated from the reaction table) and a table of parameters. Additional entry tables used were the Events, where a time trigger can change the value of a specified species or parameter and an entry table for rules where 'initial assignments' (calculated prior to the simulation) or 'repeated assignments' (calculated for every time step during the simulation) can be computed. The compartment, doses and variants features were not used. All mechanistic models discussed in the text were built as purely mass-action reaction sets, while the reduced or phenomenological models as custom. An example of the full structure of the model on SimBiology's interface is given in Appendix A and B, due to their large size. For mass-action models, the parameter values found in the literature or modelling and parameter databases were unit converted to min⁻¹ for first order reactions and min⁻¹molecule⁻¹ for second order. All species units where in molecules.

Solvers and Simulations

For deterministic simulations the system was solved by Sundials or OD15s with absolute and relative tolerances set at 1.0E-8 and 1.0E-12 with a variable maximum time step (calculated by the algorithm). The simulation time units used was minutes while the C++ compiler in MATLAB was used to accelerate the simulations along with the parallel computing feature of the application. For stochastic simulations the SSA solver was used, but was only able to run models that are purely mass-action.
A number of in-built simulation functions were used. The time course simulation, where specified species can be tracked over time, the scan function where a time course simulation is repeatedly run by changing a set of parameters within a predetermined range and step and finally the ensemble runs where a stochastic simulation is run multiple times.

Once the model was simulated all the run specific information and the model structure are exported as an object in MATLAB’s workspace, from where custom scripts were used to analyse the simulations further. Custom scripts were used for generating the surface plots shown in Chapter 3.1. (Multichromatic TCS model), the surface plots in Chapter 4, and all the plots that include amplitude or period values in Chapter 6.

**Sensitivity Analysis**

For the sensitivity analysis, an inbuilt function of SimBiology was used. All or a group of parameters were
set as 'inputs' for perturbation. The 'output' species on which the sensitivity is measured, was the reporter species. The application records the time dependent sensitivity of a species, that can be not normalised, half normalised with respect of the species amount at the specific time point, or fully normalised according to 2.8:

\[ S_x^k(t) = \frac{k}{x(t)} \left| \frac{dx(t)}{dk} \right| \]  

(2.8)

where \( S_x^k(t) \) is the normalised sensitivity at time \( t \) of species \( x \) for the parameter \( k \). The sensitivity values reported in the text is the time integral (2.9) of \( S_x^k(t) \) from \( t = 0 \) until \( t = \) final of the simulation.

\[ \int_0^t S_x^k(t) dt \]  

(2.9)

For the light tuneable oscillator work the sensitivity formulas used are discussed in the text (Chapter 6), since these terms are used under a different context, for defining how the amplitude or the period is changing with respect to light.

**Parameter Fitting**

The application has an inbuilt non linear regression, least squares based function for estimating unknown parameters. The estimation method used in this work was the individual fit (NLINFIT) where during an iterative loop, the error between the model and the data is minimised by optimising the values of a set of parameters. For the parameters to be estimated, an initial guessed value had to be set along with the transformation function (By default this was logarithmic). The error was calculated by a constant error model of the form \( Y = f + a*e \), where \( f \) is the model's function \( i.e. \) the output for the response species at time \( t \), \( e \) a normal distribution (Gaussian) term and \( a \) a constant parameter which is calculated by the algorithm. Other error models included proportional or combined, that were used in cases where the standard deviation in the data seemed to be increasing with the mean value of the output.

A major limitation of SimBiology's parameter fitting algorithm was that for releases earlier than MATLAB 2014b, it does not allow for a predefined range within the parameters can be fitted. This resulted in some cases, to estimations far off any reasonable biological values since of the multiple local error minima in higher dimensional models. An alternative method utilised was the 'squeeze and breath' algorithm proposed by Beguerisse-Diaz et al, (2012). This algorithm is based on the generation of random parameter values or 'seeds' in the hyper-dimensional space defined by the fitting parameter set. In each iteration of the routine, regions where the seeds exhibit high error are excluded reducing the boundaries for the region in which new seeds are generated. The process is repeated until the error minimisation function converges to the minimum possible error. The range constraints for each parameter is implemented by penalising with high error, values that fall outside that range or do not satisfy a prespecified relation. Parameters that were not expected to change between the light tuneable toggle switch in this work and the classic toggle switch, were estimated by Dr Mariano Beguerisse-Diaz on experimental unpublished data of a genetic toggle switch version kindly provided by Dr James Arpino. The collaboration with the aforementioned researchers aimed to identify biologically reasonable ranges for various parameters.
Phase Planes and Bifurcation Diagrams

The bifurcation diagrams shown in Chapter 5, were generated with MatCont, a continuation algorithm for MATLAB. The phenomenological models used for these are discussed in Chapter 5. The phase planes were drawn by a custom script while the model simulated trajectories were overlaid on the specific plots.

Curve Fitting

All curve fitting shown as hill fits in the text was performed using the MATLAB's Cftool. The general equation form used for all fitting, is shown in (2.10) for activating and (2.11) for repressing hill functions.

\[ f(x) = \frac{ax^n}{k^n + x^n} + c \]  \hspace{1cm} (2.10)

\[ f(x) = \frac{ak^n}{k^n + x^n} + c \]  \hspace{1cm} (2.11)

where \( a \) is the maximum expression rate or levels, \( k \) is the binding, activation or repression coefficient, \( c \) is the basal expression rate or levels, while the \( n \) exponent accounts for the cooperativity (for regulated promoters) or the steepness of the function. \( a \) and \( c \) initial values where set according to the data, while the two other parameters where determined by the cftool. For all, the fitting range was set from 0 to infinity.

For the LED device curve fitting of intensity versus distance, the inverse square law function (2.12) was used.

\[ f(d) = a \frac{1}{d^n} \]  \hspace{1cm} (2.12)

Finally for the standard heme concentration curve (in Chapter 3.2) a linear model (2.13) was used to fit the linear range of the curve.

\[ f(x) = a \cdot x + b \]  \hspace{1cm} (2.13)

2.2.2. Data Analysis and Illustrative Software Used

The experimental and simulation data were analysed in Excel 2010 (Microsoft) or scripts written in MATLAB (Mathworks). For flow cytometry data, FlowJo V10 (TreeStar) and Cyflogic was used. Softmax Pro 6.4 was used to export Gemini EM (Molecular Devices) data. The majority of plots in this work were generated in MATLAB using plottools or in some cases in Excel. Plots for steady state performance of the light responsive TCS were drawn using Prism (GraphPad). All diagrams, and graphs (after exporting from MATLAB or Excel) illustrations were build, edited and finalised in Illustrator CS5 (Adobe) and InDesign CS5 (Adobe). ODE equations were written and manipulated in Mapple 16 (Maplesoft) and then imported into SimBiology.
### 2.3. Materials

#### 2.3.1. Light Induction Source

*Light Emitting Diodes (LED):* Custom made LED devices were used to illuminate the light sensitive samples. The circuit diagram shown in Figure 2.3.1 was kindly provided by Dr Mark Warren and Dr Anne Fintzpatrick. The electronic components given in Table 2.3.1 were used to build a red (635 nm), blue (540 nm) and green LED device (460 nm) and a Far-Red (730 nm) device. These LEDs were used for 14 ml tube based assays (Figure 2.3.2 (a)) or microplate assays according to Figure 2.3.2 (c).

*Light Incubator:* Dr Eric Davidson provided for certain time periods a shaking incubator equipped with fully programmable LED array. Each well can be illuminated by either green or red emitting diode in a continuous or pulsed manner. This device is shown in Figure 2.3.2 (b) and the specific setup was used to obtain the surface plots shown in Chapter 4.2.

#### 2.3.2. Strains, Genetic Material and Solutions

Strains used in this work are shown in Table 2.3.2, while Table 2.3.3 shows a list of BioBricks™ (Registry of Biological Parts) used to reconstruct the Cph8 TCS. A list of plasmids used in this work is given in Table 2.3.4 followed by a list of the oligos (Table 2.3.5) used for the cloning work discussed in the text. All the oligos were ordered by Invitrogen or Sigma with the exception of the Plac/ompR downstream part (GenScript). Finally a list of buffers, solutions and antibiotics used is given in Table 2.3.6.

#### Table 2.3.1. List of Electronic components for building the LED devices

<table>
<thead>
<tr>
<th>Part Description</th>
<th>Manufacturer</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>LED Driverr 1000 mA Constant- DC/DC LDU2430S1000</td>
<td>XP Power</td>
<td>Farnell Element14</td>
</tr>
<tr>
<td>Adaptor 12VAC, 1A, UK, 2.1 AC-12100BS2.1</td>
<td>Stontronics</td>
<td>Farnell Element14</td>
</tr>
<tr>
<td>RED-LED High Power (10 Watts) LZ4-20R110</td>
<td>LED Engin</td>
<td>Mouser Electronics</td>
</tr>
<tr>
<td>Far Red -LED High Power (0.5 Watts) LZ1-10R305</td>
<td>LED Engin</td>
<td>Mouser Electronics</td>
</tr>
<tr>
<td>BLUE-LED High Power (10 Watts) LZ4-20B210</td>
<td>LED Engin</td>
<td>Mouser Electronics</td>
</tr>
<tr>
<td>Green - LED High Power (14.4 Watts) 897-LZ440G100</td>
<td>LED Engin</td>
<td>Mouser Electronics</td>
</tr>
<tr>
<td>JACK SOCKET, DC RAPC722X</td>
<td>Switchcraft</td>
<td>Farnell Element14</td>
</tr>
<tr>
<td>STRIPBOARD, EUROCARD 451058</td>
<td>Kelan</td>
<td>Farnell Element14</td>
</tr>
<tr>
<td>Heat Sinks ECC-00531-01-GP</td>
<td>Ideal Power</td>
<td>Mouser Electronics</td>
</tr>
<tr>
<td>Soldering kit B57NZ, Wiring</td>
<td>Mapling</td>
<td>Mapling Electronics</td>
</tr>
</tbody>
</table>
Figure 2.3.1. Circuit diagram of the LED devices. They consist of a 9 V, 1000 mA AC/DC transformer which is connected to the power adaptor on the circuit board. The current is directed to the LED through a 1000 mA DC/DC driver. b) Image of the circuit boards for the specific device. c) Image of the 10W 540nm, 635nm and 460 nm device under operation.
Chapter 2. Materials and Methods

Figure 2.3.2. Incubator setup for fixed and variable light intensity. (a) For 14 ml tube based assays a custom made rack on which the cultures are mounted at various distances from the light source is placed in a regular incubator with transparent walls. The intensity curves are corrected for approximately 1% absorption from the plexi glass walls of the incubator. Aluminium foil or carton paper is used as a light shield on the walls of the incubator apart from a small slit from which the LED source shines through the incubator. (b) An incubator equipped with a programmable LED array suitable for microplate based assays was kindly provided by Dr. Eric Davidson for certain time periods. (c) For microplate based assays where the microplate light incubator was not available, neutral density light filters were used to adjust the light reaching the cultures.
### Table 2.3.3. Strains used in this work

<table>
<thead>
<tr>
<th>E. coli strain</th>
<th>Genotype</th>
<th>Reference</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>JT2</td>
<td>RU1012 (ΔPompC-lacZ)</td>
<td>Tabor et al (2011)</td>
<td>Dr Eric Davidson</td>
</tr>
<tr>
<td>EPB238</td>
<td>λ- rph-1 ψ(ompR'-yfp') envZ- Δ(lacI lacZYA)</td>
<td>Batchelor &amp; Goulian (2006)</td>
<td>Prof M. Goulien</td>
</tr>
<tr>
<td>JS006</td>
<td>MG1655 (ΔaraC ΔlacI)</td>
<td>Stricker et al (2008)</td>
<td>Prof J. Hasty</td>
</tr>
</tbody>
</table>

### Table 2.3.4. BioBricks™ used in this work. (MIT Registry of Biological Parts)

<table>
<thead>
<tr>
<th>Part</th>
<th>Description</th>
<th>Length in bp</th>
<th>Distribution Plate/Well</th>
<th>Vector</th>
<th>Resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBa_S03419</td>
<td>Cph8 gene with RBS</td>
<td>2281</td>
<td>2009 Pl.2 22K</td>
<td>pSB1AC3</td>
<td>A/C</td>
</tr>
<tr>
<td>BBa_I15009</td>
<td>pcyA gene for PCB biosynthesis</td>
<td>750</td>
<td>2009 Pl.1 20F</td>
<td>pSB2K3</td>
<td>K</td>
</tr>
<tr>
<td>BBa_I15008</td>
<td>ho1 gene for PCB biosynthesis</td>
<td>726</td>
<td>2009 Pl.2 13J</td>
<td>pSB2K3</td>
<td>K</td>
</tr>
<tr>
<td>BBa_R0082</td>
<td>PompC promoter</td>
<td>108</td>
<td>2009 Pl.1 16K</td>
<td>pSB1A2</td>
<td>A</td>
</tr>
<tr>
<td>BBa_E0420</td>
<td>eCFP with RBS and terminator</td>
<td>878</td>
<td>2010 Pl.1 8K</td>
<td>pSB1A2</td>
<td>A</td>
</tr>
<tr>
<td>BBa_B0015</td>
<td>Combination of BBa_B0010 and BBa_B0012</td>
<td>129</td>
<td>2009 Pl.1 23L</td>
<td>pSB1AK3</td>
<td>A/K</td>
</tr>
<tr>
<td>BBa_B0014</td>
<td>Combination of BBa_B0011 and BBa_B0012</td>
<td>95</td>
<td>2009 Pl.2 24C</td>
<td>pSB1AK4</td>
<td>A/K</td>
</tr>
<tr>
<td>BBa_B0034</td>
<td>RBS</td>
<td>12</td>
<td>2009 Pl.1 2M</td>
<td>pSB1A2</td>
<td>A</td>
</tr>
<tr>
<td>BBa_J23110</td>
<td>Constitutive promoter</td>
<td>35</td>
<td>2010 Pl.1 20C</td>
<td>BBa_J61002</td>
<td>A</td>
</tr>
<tr>
<td>BBa_J23109</td>
<td>Constitutive promoter</td>
<td>35</td>
<td>2009 Pl.2 2G</td>
<td>BBa_J61002</td>
<td>A</td>
</tr>
<tr>
<td>BBa_I20260</td>
<td>GFPmut3b reporter</td>
<td>919</td>
<td>2009 Pl.2 17F</td>
<td>pSB3K3</td>
<td>K</td>
</tr>
<tr>
<td>BBa_I52001</td>
<td>ccdB and nonfunctional pUC19 derived high copy origin</td>
<td>1090</td>
<td>2009 Pl.1 5G</td>
<td>pSB3CS</td>
<td>C</td>
</tr>
</tbody>
</table>
### Table 2.3.5. Plasmids used in this work

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Description</th>
<th>Origin</th>
<th>Resistance</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pL-Cph8FD</td>
<td>Re-assembled constitutive Cph8 system with PCB production genes</td>
<td>pMB1</td>
<td>Amp/Kan</td>
<td>This work</td>
</tr>
<tr>
<td>pL-Cph8FD(Cm)</td>
<td>Re-assembled constitutive Cph8 system with PCB production genes</td>
<td>ColE1</td>
<td>Cm</td>
<td>This work</td>
</tr>
<tr>
<td>pL-Cph8FD(3c)</td>
<td>Re-assembled constitutive Cph8 system with PCB production genes</td>
<td>p15A</td>
<td>Cm</td>
<td>This work</td>
</tr>
<tr>
<td>pPCB(A)</td>
<td>PCB production: Ho1 and PcyA genes</td>
<td>p15A</td>
<td>Spec</td>
<td>Dr E. Davidson</td>
</tr>
<tr>
<td>pPCB(S)</td>
<td>PCB production: Ho1 and PcyA genes</td>
<td>p15A</td>
<td>Amp</td>
<td>Dr E. Davidson</td>
</tr>
<tr>
<td>pCph8</td>
<td>Cph8 gene</td>
<td>ColE1</td>
<td>Cm</td>
<td>Dr E. Davidson</td>
</tr>
<tr>
<td>pPlOPCB</td>
<td>Plac/ompR : GFP and constitutive expression of PCB genes</td>
<td>p15A</td>
<td>Amp</td>
<td>This work</td>
</tr>
<tr>
<td>pPlcPCB</td>
<td>Plac/CcaR : GFP and constitutive expression of PCB genes</td>
<td>p15A</td>
<td>Amp</td>
<td>This work</td>
</tr>
<tr>
<td>pJT118</td>
<td>CcaS/CcaR and PccpG2:LacZ</td>
<td>ColE1</td>
<td>Cm</td>
<td>Dr E. Davidson</td>
</tr>
<tr>
<td>pJT122</td>
<td>CcaS/CcaR and PccpG2:LacZ and Cph8</td>
<td>ColE1</td>
<td>Cm</td>
<td>Dr E. Davidson</td>
</tr>
<tr>
<td>pJT106b</td>
<td>PomC:CI and λPL:LacZ</td>
<td>pSC101</td>
<td>Cm</td>
<td>Dr E. Davidson</td>
</tr>
<tr>
<td>pIKE107</td>
<td>Toggle switch- PTrc-2:TetR, GFP reporter and Pteto1:LacI</td>
<td>ColE1</td>
<td>Amp</td>
<td>Prof J. Collins</td>
</tr>
<tr>
<td>pTAK117</td>
<td>Toggle switch- PTrc-2:CI and λPL:LacI</td>
<td>ColE1</td>
<td>Amp</td>
<td>Prof J. Collins</td>
</tr>
<tr>
<td>pRG021</td>
<td>Toggle switch- PTrc-2:TetR:RFP and Pteto1:LacI:GFP</td>
<td>ColE1</td>
<td>Amp</td>
<td>Dr J. Arpino</td>
</tr>
<tr>
<td>pPLORG21</td>
<td>Toggle switch- Plac/ompR:TetR:RFP and Pteto1:LacI:GFP</td>
<td>ColE1</td>
<td>Amp</td>
<td>This work</td>
</tr>
<tr>
<td>pRG21PloPCB</td>
<td>Toggle switch - Plac/ompR:TetR:RFP and Pteto1:LacI:GFP and PCB genes</td>
<td>p15A</td>
<td>Amp</td>
<td>This work</td>
</tr>
<tr>
<td>pIKEPlo</td>
<td>Toggle switch - Plac/ompR:TetR, GFP reporter and Pteto1:LacI</td>
<td>ColE1</td>
<td>Amp</td>
<td>This work</td>
</tr>
<tr>
<td>PIKEPlo(dT)</td>
<td>Toggle switch- Plac/ompR:CI, GFP reporter and Pteto1:LacI</td>
<td>ColE1</td>
<td>Amp</td>
<td>This work</td>
</tr>
<tr>
<td>A/N Description DNA Sequence 5' -&gt; 3' bp</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Cph8:PCB(FD) casette Forward TCGAATTCTTCTAGAGaagagagaaatactagatggtggccaccacgta 49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Cph8:PCB(FD) casette Reverse ATTACCGCCTTTTGGTAGAGC 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 PCB sequencing primers 1 TTTGGCCATTTCCTCTCC 19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 PCB sequencing primers 2 GAAGGTGAGCCAGTGTGA 18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Cph8 sequencing primers 1 CCCCAAGCAATGGCCGCC 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Cph8 sequencing primers 2 CCTTCTTTTGTCACTGCT 19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Synthesised: Plac ompR Downstream fragment TCGAATTCTTCTAGAGGTCCTTTTGGTAGAGC 177</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Plac/ara-1 Upstream fragment Forward TCGAATTCTTCTAGAGGCGGATACATATTTGAATGGAC 41</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 Plac/ara-1 Upstream fragment Reverse CTTCCTGCAGTACTAGTACTTTTCCTTACGCACAAGAG 45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Cph8(FD)/3C5:PCB casette Forward tacggcgcgccaggaattctttacggctagctcagtcctaggtacaatgctagcttATAAAGAGGAGAAagtacGATGAGTGCAC 99</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 PCB casette Forward ctcATggtaccTTTCTCCTCTTTAATctagtagctagcattatatcctagcatgctagctagctagctagctagtaATACCAGGCTTCTGAGTGCAC 93</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 PCB casette reverse 2 CTCAGAAGCGGCGGAGGCAGTCAGATATCATATTCAATTCTCAAGTGAAGC 62</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 CFP deletion from pLCph8(3C5) Forward tagtagatctcctgacagctagctcagtcctaggtataatgctagctactagATTAAAGAGGAGAAagtacCATGAGTGTAACT 86</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 CFP deletion from pLCph8(3C5) Reverse GAGCGAGGAAGCGGctgcagAattaatTCCCCTAGGTCAGGCGGCGGATT 51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 PpcpG2 cloning into Plac/ara Insert Forward ttacttacgacacctcgagacagtcagagggtatttattgAGCCCATTGTGCTTTTCTCT 60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 PpcpG2 cloning into Plac/ara Insert Reverse gctagcgaactacacgactggatactgacttttcacaccgTCATAGATAAATTTAAACCTTAA 70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 PpcpG2 cloning into Plac/ara Vector Reverse caataaatctccctctagtctctgagtgtgtaagtaaCCTCTTGACATTCCTCCTTG 60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 PpcpG2 cloning into Plac/ara Vector Reverse GAcgggtagaaagtaagtctagctcgtagttcaAGCATGACATTTTATCCATA 60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 Plac ompR cloning into pRG021 Forward TTATTGTCTCACTAGTGGATACATATTTGAATGGACGCTTCTG 44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 Plac ompR cloning into pRG021 Reverse CGGTGTTCTTCTGAGGAAATTTGGATACGCTTCTCATAATGGATGGC CACA 74</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 PLaclompR cloning into pIK107 Reverse 1 GAATTTCTCCATACCGGCGTCTTTTGGTAGGGAATTTGCTTCAATTCCACGAGCTCAGATTCCAGAATAATC 85</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22 PLaclompR cloning into pIK107 Reverse 2 CATAACCGGATTTCCTCTGTGAAGGTGAGCCAGTGTGAATTTATCGCTCACAAATCCACAGTCCATTCTCC 60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23 PLaclompR cloning into pIK107 Forward TTATTGCACTGCGGCGGCACCATATATTTGAAT 33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Table 2.3.7. Solutions and buffers**


<table>
<thead>
<tr>
<th>Substance</th>
<th>Work Solution Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na₂HPO₄·7H₂O</td>
<td>60 mM</td>
</tr>
<tr>
<td>NaH₂PO₄·H₂O</td>
<td>40 mM</td>
</tr>
<tr>
<td>KCl (1M)</td>
<td>10 mM</td>
</tr>
<tr>
<td>MgSO₄ (1M)</td>
<td>1 mM</td>
</tr>
<tr>
<td>β-mercaptoethanol (BME)*</td>
<td>50 mM</td>
</tr>
</tbody>
</table>

*Solved in H₂O and adjusted to pH 7.0, Stored at 4 ºC. * added on the day of use


<table>
<thead>
<tr>
<th>Substance</th>
<th>Work Solution Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na₂HPO₄·7H₂O (0.06M)</td>
<td>60 mM</td>
</tr>
<tr>
<td>NaH₂PO₄·H₂O (0.04M)</td>
<td>40 mM</td>
</tr>
</tbody>
</table>

*Solved in H₂O and adjusted to pH 7.0

3. **ONPG solution**

<table>
<thead>
<tr>
<th>Substance</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONPG</td>
<td>4 mg/ml</td>
</tr>
<tr>
<td>Phosphate buffer</td>
<td>0.1 M</td>
</tr>
</tbody>
</table>

*Prepared fresh before the Miller assay.

4. **DNA oligos Annealing Buffer**

<table>
<thead>
<tr>
<th>Substance</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris-HCl</td>
<td>10 mM</td>
</tr>
<tr>
<td>NaCl</td>
<td>50 mM</td>
</tr>
<tr>
<td>EDTA</td>
<td>1 mM</td>
</tr>
</tbody>
</table>

*Solved in H₂O and Adjusted to pH 8.0

*The oligos are heated for 1 min at 98 ºC and left to anneal overnight at room temperature

5. **Minimal Media M9 (Chappell, 2013)**

<table>
<thead>
<tr>
<th>Substance</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycerol</td>
<td>0.40 %</td>
</tr>
<tr>
<td>M9 salts</td>
<td>1X</td>
</tr>
<tr>
<td>Casamino acids</td>
<td>0.20 %</td>
</tr>
<tr>
<td>CaCl₂</td>
<td>0.1</td>
</tr>
<tr>
<td>Thiamine hydrochloride</td>
<td>1 mM</td>
</tr>
<tr>
<td>MgSO₄</td>
<td>2 mM</td>
</tr>
</tbody>
</table>

*Solved in H₂O

6. **Antibiotics (1000X solutions)**

<table>
<thead>
<tr>
<th>Substance</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ampicilin (Amp)</td>
<td>100 mg/ml</td>
</tr>
<tr>
<td>Kanamycin (Kan)</td>
<td>50 mg/ml</td>
</tr>
<tr>
<td>Chloramphenicol (Cm)</td>
<td>34 mg/ml</td>
</tr>
<tr>
<td>Spectinomycin (Spec)</td>
<td>100 mg/ml</td>
</tr>
</tbody>
</table>
The successful design and implementation of large and complicated synthetic genetic networks, that are controlled by light requires prior knowledge of the performance characteristics of the 'light input' module. Hence a starting point for this work was the detailed and precise characterisation of the light receptors used, both on a theoretical and experimental level. This enabled to some extent to follow a bottom-up approach towards realising a light controlled network design, that incorporates the static and dynamical properties of the light input modules allowing to roughly predict the behaviour of the physical system. In the following chapter I discuss in detail the structure of the model for the Cph8 receptor, the reduction process followed and how this model was shown able to capture the basic behaviour of the system in both qualitative and quantitative level. Finally I present the experimental results and model modifications regarding the Multichromatic control system that comprises both the Cph8 and CcaS receptors.
3.1. The Cph8 Two Component System Model

3.1.1. System Overview

As stated in Chapter 1 introduction (Section 1.3.1) Cph8 is a chimeric transmembrane receptor (Levskaya et al, 2005) with its intracellular domain originating from the *Escherichia coli* EnvZ osmoregulator (Cai & Inouye, 2002) and the outer-membrane domain from the Cph1 phytochrome (Hübschmann et al, 2001) from *Synechocystis* sp. More specifically it retains only the Histidine Kinase domain (HK) but not the phosphatase domain (PA) from EnvZ and the Cph1 photo-sensory domain from the Cph1 receptor from *Synechocystis* sp. This allows for the receptor to bind Phycocyanobilin (PCB) and render a photosensitive outer membrane domain which once on ground state (Pr) the intracellular HK domain can bind and phosphorylate OmpR.

Phosphorylated OmpR (OmpR-p) has increased affinity to the PompC and PompF promoters of the natural OmpC and OmpF osmoregulatory porins in *E. coli* (Head et al, 1998). Upon red light induction (635 nm) the active phytochrome will transit to the excited state (Pfr) which results to a conformational change of the HK domain rendering it unable to phosphorylate OmpR. The drop in the intracellular levels of OmpR-p will drive the transcription rate from PompC back to basal levels but raise the activity of the PompF. While the natural EnvZ system utilises both PompF and PompC promoters only the latter was used in this work in line with the original work from Levskaya et al (2005) who implemented the Cph8 chimeric receptor.

This promoter is particularly interesting in the way that binds OmpR-p. It comprises of three operators each having a primary and secondary binding site (a total of six binding sites per promoter). Six OmpR-p molecules are required for a fully occupied and active promoter. Furthermore the order of the binding appears to be specific and follows a hierarchy both across the operators but also the two sub-operator binding sites as shown in Figure 3.1.1 (Yoshida et al, 2006). This binding pattern is called the “galloping” model in parallel to the order that a horse will take its steps.

The chromophore PCB, can be provided externally, but Gambeta & Lagarias (2001) showed that it is possible to produce the specific chromophore in bacteria or yeast when the enzymes Ho1 and PcYA are expressed in the cell. More specifically Ho1 (Heme oxygenase 1) is responsible for the initial reduction of Heme to biliverdin (BV, an open porphyrin ring) and PcYA (Phycocyanobilin: ferredoxin oxidoreductase) for an additional reduction step of Biliverdin to PCB. The PCB biosynthesis, its absorption spectra and structure is shown in Figure 3.1.2.
The Galloping Model

**Figure 3.1.1.** Diagrammatic overview of the galloping model. (a) Top: Phosphorylation and de-phosphorylation of OmpR by the EnvZ Histidine Kinase domain (HK) and Phosphatase domain (PA) respectively. (a) An OmpR-p molecule will first bind the secondary (F1b) site of the operator. Intermolecular non-covalent interaction promotes dimerisation with a second OmpR-p molecule which eventually will be stabilised on the primary site F1a. (b) The same applies on the PompC operator sites (e.g. C1: C1a, C1b) where C2 will be occupied place after both sides of C1 are bound and C3 will follow (Red numbers indicate the order of binding). The figure was adapted from Yoshida et al. (2006)

**Figure 3.1.2.** (a) The biosynthesis pathway for Phycocyanobilin in *Synechocystis* sp. (Reconstructed from MetaCyc biochemical pathway database) (b) The absorption spectra of Pr and Pfr states of the holophytocrome bearing PCB (Figure from Miller et al. (2006)) (c) Structure of PCB and domain profile for Cph1. Figure adapted from Essen et al (2006). The Cys 259 on which PCB binds covalently is indicated
3.1.2. Biochemical Model

Based on the literature information available, a set of biochemical reactions that take part in the functional receptor formation, the phosphorylation of OmpR and finally expression of the reporter gene from the PompC promoter was derived. These reactions are shown diagrammatically in Figure 3.1.3 This diagram can be split into three main modules shown in Table 3.1.1.

- The PCB production module
- The light sensing module
- The reporter module

This kinetic scheme was based on a set of initial assumptions:

- Proteins and small molecules (e.g. Heme, PCB) are degrading much slower than the cell division time hence their degradation rate approximates the dilution rate due to cell division. Based on this, the enzymes, like in their free form, they were assumed to degrade with their substrates as a complex instead of modelling separate degradation reaction of the enzyme and a different one for the possibility of the substrate to degrade while bound on the enzyme.
- The ultrafast dynamics at a picosecond scale of the PCB isomerisation (Heyne et al, 2002) were omitted since the rest of the downstream modelled state transitions progresses on a microsecond or slower scale.
- ATP-mediated auto-phosphorylation reactions of the receptors were assumed to proceed at a constant rate and the system was assumed to operate under chemostatic conditions in terms of aeration, stage of growth and availability of resources. The auto-phosphorylation of HK by ATP reaction was assumed embedded in the phosphorylation of OmpR.
- Finally dimerisation of Cph8 is not modelled since there is no sufficient information in the literature and a 1:1 stoichiometry of OmpR and Cph8 HK domain was reported by Yoshida et al (2002)

Table 3.1.1. The three identified modules of the Cph8 system and their components*

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
<th>Intermediate Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCB production</td>
<td>Heme</td>
<td>PCB</td>
</tr>
<tr>
<td>Light sensing</td>
<td>PCB, Light</td>
<td>OmpR-p</td>
</tr>
<tr>
<td>Reporter</td>
<td>OmpR-p</td>
<td>LacZ</td>
</tr>
</tbody>
</table>

* All the components on the table are presented with their assigned model name (e.g. gCph8 refers to the cph8 gene) shown in Appendix A model species
**Figure 3.1.3.** The network diagram of the Cph8 receptor model. The PCB production module is shown at the bottom left side. The different states of the Cph8 receptor shown at the top are ‘cph8’ as the free form, ‘Pfr’ active Cph8-PCB complex form and ‘Pr’ inactive Cph8-PCB complex form. Finally the different states of the ompR regulator, PompC promoter and the LacZ reporter are shown at the centre and bottom right. Blue thick lines represent the main flux of the system when Cph8 is starting from its ground state, able to bind and phosphorylate OmpR. Thin grey lines represent inverse reactions and degradations. Black lines account for transcription, translation and state transitions. The \( \text{ompR} \) gene (\( g\text{ompR} \)) is the only native gene in \( E. \text{coli} \), the Pompc promoter upstream the LacZ gene was integrated into the genome of the RU1012 (KanR:ΔenvZ strain) (Utsumi et al., 1989) while the rest were cloned into plasmid vectors with either medium copy number (p15A: 15-20) to high (ColE1: 50, pMB1: 200). The model was adjusted for the different combinations of vectors and genes. \( h\text{o1} \) and \( p\text{cyA} \) were set in a biocistronic casette sharing the same promoter and same ribosome binding site (RBS). The diagram was designed in CellDesigner v4.0.1.
Chapter 3. The Light Receptor Characterisation

PCB production module:

The PCB producing genes, ho1 and pcyA are transcribed in a bicistronic configuration under the same promoter with a rate constant $t_r$, which embeds the time for RNA polymerase binding, formation of the closed complex and elongation time of the mRNA. Both bear identical copies of Ribosome Binding Sites (RBS) which will bind initiate and elongate translation with a rate constant $t_f$. RNA polymerase, Ribosomes, NTPs, tRNA and cell resources in general are not explicitly modelled.

$$gPCB \xrightarrow{tr} mPCB + gPCB$$
$$mPCB \xrightarrow{dp} mPCB + Ho1 + PcyA$$

Ho1 can bind Heme reversibly form a complex and catalyse the reduction to BV with a rate constant $k_h$ and $c_h$ respectively. Similarly PcyA binds and reduces BV into PCB with rate constants $k_b$ and $c_b$.

$$Ho1 + Heme \xrightarrow{k_h \over k_a} Ho1c \xrightarrow{c_h} BV + Ho1$$
$$PcyA + BV \xrightarrow{k_b \over k_a} PcyAc \xrightarrow{c_b} PCB + PcyA$$

With the exception of gPCB everything is decayed from the system as follows:

$$mPCB \xrightarrow{dp} \emptyset$$

$$BV \xrightarrow{dp} \emptyset$$
$$PCB \xrightarrow{dp} \emptyset$$

$$Ho1 \xrightarrow{dp} \emptyset$$
$$Ho1c \xrightarrow{dp} \emptyset$$
$$PcyA \xrightarrow{dp} \emptyset$$
$$PcyAc \xrightarrow{dp} \emptyset$$

where $d_r$, $d_p$, and $d_h$ are the decay rate constants for mRNA the enzymes and the billins. This decay constants are the sum of the molecule degradation rate and dilution rate due to cell division.
The Cph8 module:

Transcription, translation and degradation for Cph8 and OmpR is summarised in:

\[
g\text{Cph8} \xrightarrow{\nu_r} m\text{Cph8} + g\text{Cph8} \\
g\text{OmpR} \xrightarrow{\nu_r} m\text{OmpR} + g\text{OmpR} \\
m\text{Cph8} \xrightarrow{\nu_t} f\text{Cph8} + m\text{Cph8} \\
m\text{OmpR} \xrightarrow{\nu_t} \text{OmpR} + m\text{OmpR}
\]

\[
f\text{Cph8} \xrightarrow{d_p} \emptyset \\
m\text{Cph8} \xrightarrow{d_{pr}} \emptyset \\
\text{OmpR} \xrightarrow{d_p} \emptyset \\
m\text{OmpR} \xrightarrow{d_{pr}} \emptyset
\]

Where \(tr, tl\) and \(dp\) and \(d\) are rate constants for transcription translation, Protein and mRNA degradation respectively. Index \(r\) refers to the Cph8 receptor and \(o\) to the OmpR regulator.

Furthermore free \(f\text{Cph8}\) can bind PCB covalently (Miller et al. 2006) and obtain the Pr light sensitive form of the receptor docked on the cell membrane. In addition dimerisation of Cph8 molecules is not shown explicitly since the stoichiometry for OmpR to EnvZ(HK) is 1:1 when the system is phosphorylating OmpR (Yoshida et al, 2002). For this system the transition between active (Pr) and inactive (Pfr) form of the holophytochrome is a function of red light \(f(LR)\) and reversibly of far-red light \(f(LFR)\) plus the rate constant \(k\) which account for the system returning from the excited (Pfr) state to the ground state (Pr) in the absence of light.

\[
f\text{Cph8} + PCB \xrightarrow{k} C\text{ph8}_{pr} \\
C\text{ph8}_{pr} \xleftarrow{k + f(LFR)} C\text{ph8}_{pfr}
\]

\[
C\text{ph8}_{pr} \xrightarrow{d_{pr}} \emptyset \\
C\text{ph8}_{pfr} \xrightarrow{d_{pr}} \emptyset
\]

Finally for the OmpR:Cph8(HK) complex formation and phosphorylation the following reactions were considered:

\[
\text{OmpR} + C\text{ph8}_{pr} \xrightarrow{k} C\text{ph8c} \xrightarrow{\gamma} C\text{ph8}_{pr} + \text{OmpR}
\]

\[
\text{OmpR} \xrightarrow{\alpha} \text{OmpRp} \\
\text{OmpRp} \xrightarrow{d_p} \emptyset \\
C\text{ph8c} \xrightarrow{d_p} \emptyset
\]

where \(\alpha\) and \(\alpha_{-1}\) account for the phosphorylation and dephosphorylation by auxiliary interactions with other kinases and phosphatases in the host cell (Groban et al, 2009).

*The functions are being investigated and defined during model reduction (Section 3.1.3)*
The Reporter Module:

OmpR-p has an increased affinity to the PompC promoter compared to OmpR. The PompC promoter has three operators C1,C2,C3 and their affinity order to OmpR-p is C1>C2>C3 resulting to a hierarchical manner of binding as described in Figure 3.1.1 (Yoshida et al. 2006, Head et al, 1998). Each operator holds two binding sites where the binding of the primary was assumed as the rate limiting step while the occupation of the secondary one and dimerisation of OmpR-p was considered to be the fast step with the same rate constant $s$ across all three operators. To model the biochemical reactions the promoter:gene species $R_{[b,a]}$ was introduced where $b$ represents the index of primary sites occupied and $a$ the index of secondary sites (right). The limitation that the $(i+1)^{th}$ $b$ site can only bind OmpR-p once the $i^{th}$ corresponding secondary site is occupied was used in order to model a strict hierarchical order for binding. For example $R_{2,1}$ indicate that the C1b and C2b (Cib are the primary sites as shown in Figure 3.1.1) while only the C1a secondary site is occupied. The next OmpR to bind the promoter will occupy the C2a site.

$$OmpRp + R_{i-1,i-1} \xrightarrow{h_{b,a}} R_{i,i-1} + OmpRp \xrightarrow{h_{a,i}} R_{i,i}$$

$i \in \{1, 2, 3\}$

The promoter $R$ in the current model can exist in 7 different occupancy states with affinity order, and total number of OmpR-p bound on:

$$R_{0,0} > R_{1,0} > R_{1,1} > R_{2,1} > R_{2,2} > R_{3,2} > R_{3,3}$$

$$N_{\text{OmpR}:R_{b,a}} = a + b$$

$a, b \in \{1, 2, 3\}$

The binding of OmpR-p to nonadjacent sites due to their smaller contribution to the overall behaviour of the system and for reducing the complexity of the system, was neglected. Hence the binding in this model is strictly hierarchical. Once the last secondary site $R_{3,3}$ is occupied, LacZ* is transcribed and translated according to:

$$R_{3,3} \xrightarrow{r} mLacZ + R_{3,3}$$

$$mLacZ \xrightarrow{d_t} LacZ + mLacZ$$

$$mLacZ \xrightarrow{d_t} \emptyset$$

$$LacZ \xrightarrow{dp_c} \emptyset$$

* In a variant of the model LacZ was replaced with eCFP and its corresponding constants and gene copy number used in the case of eCFP experimental measurements
The corresponding set of Ordinary Differential Equations (ODEs) for the mass-action model is shown below:

**PCB module:**
\[
\begin{align*}
\dot{H}o1 &= I_l - k_s[Ho1][Heme] + c_i[Hoi1c] + k_a[Hoi1c] - d_p[Ho1] \\
\dot{Hoi1c} &= k_s[Ho1][Heme] - c_i[Hoi1c] - d_p[Hoi1c] - k_a[Hoi1c] \\
\dot{B}V &= c_i[Hoi1c] + k_a[PcyAc] - k_s[PcyAc][BV] - d_s[BV] \\
\dot{P}cyA &= I_l - k_s[PcyA][BV] + c_i[PcyAc] - d_p[PcyA] + k_a[PcyAc] \\
\dot{P}cyAc &= k_s[PcyA][BV] - c_i[PcyAc] - d_p[PcyAc] - k_a[PcyAc] \\
\dot{PC}B &= c_i[PcyAc] - d_p[PCB] - k_s[fCph8][PCB] \\
mPCB &= tr_g[PCB] - d_p[mPCB] \\
\dot{H}eme &= k_s[Hoi1c] - k_s[Ho1][Heme]
\end{align*}
\]

**Receptor Module:**
\[
\begin{align*}
\dot{fCph8} &= I_l[mCph8] - k_s[fCph8][PCB] - d_p[fCph8] \\
\dot{C}ph8_{pr} &= k_s[fCph8][PCB] - (k_i + f(LFR))[Cph8_{pr}] + k_s[Cph8c] - [Cph8_{pr}](k_s[OmpR] + f(LR) + d_p) \\
\dot{C}ph8_{pr} &= -(k_i + f(LFR))[Cph8_{pr}] + f(LR)[Cph8_{pr}] - d_p[Cph8_{pr}] \\
\dot{C}ph8c &= k_s[OmpR][Cph8_{pr}] - k_s[Cph8c] - d_p[Cph8c] - c_i[Cph8c] \\
\dot{OmpR} &= I_l[mOmpR] - k_s[OmpR][Cph8_{pr}] + k_s[Cph8c] - d_p[OmpR] - a_i[OmpR] + a_o[OmpRp] \\
\dot{OmpRp} &= c_i[Cph8c] + a_i[OmpR] + \sum_{i=1}^{3} b_i[R_{1,i}] + \sum_{i=1}^{3} s_i[R_{1,i}] - [OmpRp] \left( \sum_{i=1}^{3} b_i[R_{2,i-1}] + s_i[R_{2,i-1}] \right) + d_p + a_o \\
mCph8 &= tr_g[Cph8] - d_p[mCph8] \\
mOmpR &= tr_g[OmpR] - d_p[mOmpR]
\end{align*}
\]

**Reporter Module:**
\[
\begin{align*}
\dot{R}_{0,0} &= +b_1[R_{1,1}] + b_0[OmpRp][R_{0,0}] + d_p \left( \sum_{i=1}^{3} (R_{1,i} + R_{i,1}) \right) - b_1[OmpRp][R_{0,0}] + b_0[OmpRp][R_{0,0}] - d_p[R_{0,0}] \\
\dot{R}_{1,0} &= -s_1[R_{1,1}] + s_0[R_{1,1}] - b_1[R_{1,0}] + b_0[OmpRp][R_{0,0}] - d_p[R_{1,0}] \\
\dot{R}_{1,1} &= +s_1[R_{1,1}] + b_1[R_{1,1}] + b_2[R_{2,1}] - b_1[OmpRp][R_{1,1}] - d_p[R_{1,1}] \\
\dot{R}_{2,1} &= -s_2[R_{2,1}] + b_1[R_{2,1}] - b_2[R_{2,1}] + b_0[OmpRp][R_{1,1}] - d_p[R_{2,1}] \\
\dot{R}_{2,2} &= +s_2[R_{2,2}] + b_2[R_{2,2}] - b_1[R_{2,2}] - b_0[OmpRp][R_{2,2}] - d_p[R_{2,2}] \\
\dot{R}_{3,2} &= -s_3[R_{3,2}] + b_3[R_{3,2}] - b_1[R_{3,2}] - b_0[OmpRp][R_{3,2}] - d_p[R_{3,2}] \\
\dot{R}_{3,3} &= +s_3[R_{3,3}] + b_3[R_{3,3}] - b_2[R_{3,3}] + b_0[OmpRp][R_{3,3}] - d_p[R_{3,3}] \\
mLacZ &= tr_g[R_{3,3}] - d_p[mLacZ] \\
\dot{LacZ} &= I_l[mLacZ] - d_p[LacZ]
\end{align*}
\]
3.1.3. Model Reduction

From an early stage it was clear that deriving and working with the proposed mass-action model was a challenging task due to the large number of reactions and time variables. In addition such a model cannot be easily identified in terms of constants and parameter values along with the fact that any mathematical manipulation of a multidimensional system with such a large number of non-linearities extends beyond the purposes and capacity of this work. Therefore, the model was reduced.

Model reduction can potentially decrease the number of dimensions and non-linearities, allow the use of lumped parameters (e.g. Dissociation constants $k_\text{D}$) and potentially highlight the most critical parameters to investigate further. In the following reduction process it is demonstrated how starting from the mass-action model and by introducing assumptions based on the current understanding of the system, the model was effectively reduced to the least possible number of time variables. The assumptions used can be summarised in the following categories:

- Quasi-Steady State Approximations (QSSA), based on time scale separation or quantity scale separation
- Identification of conserved quantities e.g. Total concentration of an enzyme or receptor
- Symmetry for the promoter’s dynamics
- Other case specific assumptions

The introduction of these assumptions and estimation of some of the parameters was carried out in parallel since the former requires some prior knowledge of rate constant values and initial conditions. On the other hand, the estimation was not always possible for kinetic rate constants individually but only when are combined into a practically measurable quantity that can be found in the literature. In addition the theoretical system is only examined when the cells are in the exponential phase of growth in line with the experimental conditions applied for time course runs. As a consequence the dilution rate due to cell division was assumed globally constant.

The objective for each of the three modules was to reach an expression which relates the INPUT and OUTPUT (table 3.1.1). It is understood that collapsing the relatively large set of ODE’s into three expressions is only possible when a large number of assumptions is introduced. These inevitably have an effect on the predicted dynamics of the system yet the steady state should have been approximately equal for both the mass-action model and the reduced one.

Finally it is worth pointing out that the system is stable and all molecules converge and rest to their steady state under constant light conditions. The experimental data shown in the experimental section of this chapter (Section 3.2) suggested that the transition from the basal level of expression to the maximum (for the LacZ reporter), lasted approximately ninety minutes to two hours.
The conversion of Heme to PCB (Phycocyanobilin) is catalysed via two enzymatic reactions with one intermediate substrate (BV). The rate of change of the final product PCB is

$$\dot{PCB} = c_p[\text{PcyAc}] - d_a[PCB] - k_i[fCph8][PCB]$$  \hspace{1cm} (3.1)

and the objective was to eliminate the intermediate species so the production term of (3.1) can be re-written in terms of the initial substrate (Heme). These requires the derivation of an expression for PcyAc which is also part of the system

$$\{\text{PcyAc, BV, Ho1c, Ho1, Heme}\}$$  \hspace{1cm} (3.2)

Heme is a vital component for the general bacterial function and it is tightly regulated within the host. Cases where an exogenous “Heme sink” was introduced, like cytochrome b5 that can bind Heme but not metabolise it, was shown to trigger upregulation of heme biosynthesis such that the amounts of free heme in the system are quickly recovered (Woodard & Dailey, 1995). Thus, the total amount of heme (free and bound) is elevated. In fact Ho1 can bind heme fast but the oxygen mediated catalysis step is slow, hence a similar Heme-sink effect occurs. This was confirmed in the experimental data (section 3.2.4). According to the same data the total estimated amount of Heme, even in cells that are not expressing Ho1 is much higher than the total predicted amount of Ho1 at equilibrium. Based on the above and taking in consideration the experimental conditions regarding the growth conditions and the fact that the system was examined only when the cells have exited the lag phase of growth, the following assumptions were introduced:

**Assumption 3.1:** Heme quantity is much higher than the total Ho1 and considering also the upregulation in heme production when Ho1 is expressed the free Heme levels will be in excess and Ho1 has a negligible effect on its pool. Thus Heme is treated as a constant.

**Assumption 3.2:** The relatively high and constant levels of free Heme and its relatively high affinity for Ho1 will result into the fast complex formation of Ho1c. Adding to the latter, the slow conversion step to BV, the total amount of Ho1 was assumed to be in the Ho1c form.

**Assumption 3.3:** Due to the fact that both enzymes are constitutively expressed under the experimental conditions used the total amount of enzymes (Ho1, Ho1c, PcyA, PcyAc) and consequently BV will be at steady state throughout the experimental time course.

**Assumption 3.4:** The two enzymes Ho1 and PcyA that catalyse the first and second reaction respectively are transcribed under the exact same promoter and same strength ribosome binding sites. In addition they have similar gene sizes, in base pairs (ho1 = 738 bp and pcyA = 768 bp). Hence in the physical system the RNA Polymerase binding rate, closed complex formation, ribosome binding rate, translation initiation rate and elongation time for both was assumed approximately equal, thus their total amount at any given time was assumed equal.
Chapter 3. The Light Receptor Characterisation

\[
\begin{align*}
A(3.4), A(3.3) & \Rightarrow \\
\begin{cases}
\dot{\Ho} = \Ho c +\Ho &\text{tot} \\
\dot{\PcyA} = \PcyAc + \PcyA &\text{tot}
\end{cases} \\
\Rightarrow \Ho = \PcyA \approx E
\end{align*}
\]

The system (3.2) can then be reduced into the form

\[
\begin{align*}
&\Ho = k_h[Ho][Heme]-c_h[Hole]-d_p[Hole]-k_h[Ho] \approx 0 \\
&\text{tot} \\
&BV = c_h[Hole] + k_h[PcyAc]-k_h[PcyAc][BV]-d_p[BV] \approx 0 \\
&PcyAc = k_h[PcyA][BV] - c_h[PcyAc] - d_p[PcyAc] - k_h[PcyAc] \approx 0
\end{align*}
\]

where the enzymes and intermediate products are on their equilibrium. Based on (3.2),(3.3) Ho1 can be re-written,

\[
\Ho = E - \Ho
\]

a variation of the classic Michaelis-Menten equation for Ho1c can then be derived

\[
A(3.2), QSSA \quad [\Ho c] = \frac{[E][Heme]}{[Heme]+K_{m1}}
\]

\[
K_{m1} = \frac{k_{-h} + c_h + d_p}{k_h}
\]

\(K_m_1\) is the Michaelis-Menten constant variant that includes the degradation rate (units in molecules). Adding both sides of the remaining equations of (3.4) and substituting with (3.6) yields

\[
PcyAc = \frac{c_h[Heme][E][Heme] - d_p[BV]([Heme]+km)}{([Heme]+km)c_p}
\]

Hence when (3.7) substitutes PcyAc in (3.1), PCB can be expressed as

\[
PCB = \frac{c_h[Heme][E][Heme] - d_p[BV]([Heme]+km)}{([Heme]+km)} - [PCB](d + k [fCph8])
\]

Assumption 3.5: When A(3.2) and A(3.3) are taken into account, BV at steady state will be small enough because of its slow production rate by Ho1c and fast uptake and catalysis by PcyA. This results into

\[
BV \approx 0 \Rightarrow c_h[Heme][E] \gg d_p[BV]([Heme]+km)
\]

A(3.5) \Rightarrow PCB = \frac{c_h[Heme][E]}{([Heme]+km)} - [PCB](d + k [fCph8])

In other words A(3.5) and (3.9) suggest that the total amount of BV is instantly converted to PCB
At this point $E_{\text{tot}}$ could be defined further. Regardless of the state and interactions of its constituent components ($H_1/H_1c$ or PcyA/PcyAc) the production and depletion of $E_{\text{tot}}$ at steady state because of $A(3.3)$ can be described by

$$g_{\text{PCB}} \rightarrow_{r_p} m_{\text{PCB}} + g_{\text{PCB}}$$

$$m_{\text{PCB}} \rightarrow_{d_p} E_{\text{tot}} + m_{\text{PCB}}$$

$$m_{\text{PCB}} \rightarrow_{d_p} \emptyset$$

$$E_{\text{tot}} \rightarrow_{d_p} \emptyset$$

$$\dot{m}_{\text{PCB}} = tr_p[g_{\text{PCB}}] - d_p [m_{\text{PCB}}] \approx 0$$

$$\Rightarrow [m_{\text{PCB}}] = \frac{tr_p[g_{\text{PCB}}]}{d_p}$$

$$\dot{E}_{\text{tot}} = tl_p[m_{\text{PCB}}] - dp_p[E_{\text{tot}}] \approx 0$$

$$\Rightarrow [E_{\text{tot}}] = \frac{tl_p tr_p [g_{\text{PCB}}]}{d_p dp_p} = T_p[g_{\text{PCB}}], \quad T_p = \frac{tl_p tr_p}{d_p dp_p}$$

(3.10)

Where $T_p$ is the dimensionless parameter for the production rate of either enzyme. By substituting (3.10) into (3.9) the PCB production function was expressed only in terms of gene copy number ($g_{\text{PCB}}$) and Heme with the relevant parameters. The depletion function is a function of the protein degradation rate and the binding to fCph8 which is discussed in the light receptor module reduction. As a way to validate the reduction process, numerical simulations were performed in order to compare the reduced module model with its mass-action and stochastic counterparts. A summary for these simulations is shown in Figure 3.1.4, for the impact of reduction on the dynamics. An expected deviation in the dynamics due to the time scale separation assumptions was observed for the overall PCB, though showing minimal deviation from the steady state. In fact, this deviation was greater for higher values of Heme in the cell. Finally an additional factor that affects the steady state difference between the reduced and the full model is the gene ($g_{\text{PCB}}$) copy number. To address this assumption $A(3.5)$ had to be ignored in order to reach an expression for $B_V$. The exact solution for $B_V$ steady state was derived when (3.6) and the equivalent Michaelis-Menten solution for PcyAc were injected into the rate of change of $B_V$ in (3.4). This yields a quadratic equation and its analytical solution for $B_V$ is shown in Appendix A (Table of Rules). This solution can be used directly in (3.8). The numerical time course simulation when (3.8) is used, shown in Figure 3.1.5, showed that the steady states of the full and reduced model matched.

**The Light Receptor Module Reduction**

The light receptor module is where the phosphorylation levels of the OmpR pool is regulated within the host. The aim in this section was to obtain an expression of OmpR-p as a function of light. Based again on the physical system's structure the following assumptions were introduced.

**Assumption 3.6:** The $cph8$ gene was expressed constitutively under the J23109 promoter thus under the experimental conditions used, the receptor amount was assumed constant. The same applied for OmpR which was expressed from its native single genomic copy such that the OmpR regulator levels are maintained constant and approximately at $\sim 3000$ copies per cell (Cai & Inouye, 2002).
Chapter 3. The Light Receptor Characterisation

Figure 3.1.4. Simulations for the mass-action, reduced and stochastic model Based on (3.9). The figures show the constraints and limits of the reduction of the PCB module were evaluated numerically. The computational model was set up such that it can calculate simultaneously both reduced and full model using the same rate constants. The downstream network starting from the fCph8-PCB association reaction was maintained in the full mass-action form so fCph8 is the only dynamic variable that PCB is depended on. Equation (3.9) was used in this case were BV is small enough to neglect it based on A(3.5). The dynamics shown in figure (a) were faster for the reduced model as a result of the assumptions that the enzymes are already at steady state. Nevertheless the steady state level of PCB showed minimal deviation from the mass-action model which was the direct result of A(3.5). The transfer function (b) and (c) for varying amounts of initial Heme concentration and gPCB copy number respectively showed that the deviation is minimal when either of the two are kept low.
Assumption 3.7: The formation of the free fCph8 and PCB complex is fast because of the large and not limiting amounts of PCB. This was supported by the predicted expression levels of Cph8 under the estimated parameter set where Cph8’s steady state amounts are in the range of hundreds of molecules (in line with the natural EnvZ levels (Cai & Inouye, 2002)), while PCB lies within the region of thousands. This allowed from an early stage to neglect the reaction where free fCph8 docked on the membrane and be constitutively ON due to the absence of PCB to drive the conformational change. However any significant leakiness was assumed to be accounted for in the $a_1$ rate constant for the auxiliary phosphorylation reaction.

Assumption 3.8: Photon absorption by PCB proceeds at a time scale of femtoseconds (Dasgupta et al, 2009) while the conformational change from Pr to Pfr form and OmpR-p phosphorylation at the scale of microseconds (Möglich & Moffat, 2010). This is orders of magnitude faster than transcription and translation of the reporter, hence the ratios of Pr/Pfr and OmpR/OmpR-p were assumed to be adjusting instantly upon light condition change. Based on that a QSSA on the Cph8c, Cph8Pr/Pfr and OmpR can be applied.

In the rate of change of OmpR-p equation

$$OmpR_p = c_i[Cph8c] + a_i[OmpR] - [OmpRp](a_1 + dp_o) + J$$

$$J = \sum_{i=1}^{J} \left( b_i[R_{i,1}] + s_i[R_{i,1}] - [OmpRp] \left( b_i[R_{i,1}] + s_i[R_{i,1}] \right) \right)$$

(3.11)
Chapter 3. The Light Receptor Characterisation

Cph8c can be expressed at steady state based on A(3.6)-A(3.8) in terms of Cph8 copy number, PCB and light. Applying the QSSA for the Cph8 system of ODEs:

\[
\begin{align*}
\dot{fCph8} &= tl_f[mCph8] - k_i[fCph8][PCB] - dp_f[fCph8] \approx 0 \\
\dot{Cph8}_{py} &= k_i[fCph8][PCB] + \left( k_i + f(LFR) \right)Cph8_{py} + k_+Cph8c - \left( Cph8_{py} \right) \left( k_i[OmpR] + f(LR) + dp_p \right) \approx 0 \\
\dot{Cph8}_{py} &= -\left( k_i + f(LFR) \right)Cph8_{py} + f(LR)Cph8_{py} - dp_pCph8_{py} \approx 0 \\
Cph8c &= k_k[OmpR][Cph8_{py}] - k_+[Cph8c] - dp_p[Cph8c] - c_+[Cph8c] \approx 0
\end{align*}
\] (3.12)

Solving the third equation from (3.12) for \( Cph8_{py} \) (3.13) was obtained

\[
A(3.8) \left[ Cph8_{py} \right] = \frac{f(LR)[Cph8_{py}]}{(k_i + f(LFR)) + dp_p} \Rightarrow [Cph8_{py}] = L[Cph8_{py}]
\] (3.13)

\[
\Rightarrow L = \frac{f(LR)}{(k_i + f(LFR)) + dp_p}
\]

Where \( L \) is the master function of light which includes the combined effect for both red and far-red light. Next the conserved quantities \( Cph8_{PCB} \) were defined as the total holophytochrome (Cph8 bound with PCB) and \( Cph8_{tot} \) as the total amount of the receptor.

\[
\begin{align*}
Cph8_{PCB} &= Cph8_{py} + Cph8_{py} + Cph8c \\
Cph8_{tot} &= fCph8 + Cph8_{PCB}
\end{align*}
\] (3.14)

Substituting (3.13) into (3.14) and the resulting equation into the ODE for Cph8c in (3.12) yielded

\[
QSSA \quad [Cph8c] = \frac{[OmpR][Cph8_{PCB}]}{km_2(L + 1) + [OmpR]} \Rightarrow km_2 = \frac{c_+ + k_+ + dp_p}{k_+}
\] (3.15)

Finally the steady state for \( fCph8 \) from (3.12), \( mCph8 \) and \( Cph8_{tot} \) in terms of \( gCph8 \) as in (3.10) was obtained and solved for \( Cph8_{PCB} \)

\[
\begin{align*}
[fCph8] &= tl_f[mCph8] - k_i[PCB] + dp_f \Rightarrow [Cph8_{PCB}] = \frac{tr[fCph8][gCph8]}{d, dp_p}[Cph8_{PCB}] \\
[mCph8] &= tr[gCph8] + dp_f \Rightarrow [Cph8_{PCB}] = \frac{tr[tl_f[mCph8]]}{d, dp_p}[Cph8_{PCB}] \\
Cph8_{tot} &= tl_f[mCph8] \Rightarrow [Cph8_{PCB}] = [Cph8_{tot}] - [fCph8]
\end{align*}
\]
where $Tr$ and $kmr$ are the dimensionless combined parameters of the rate constants carried from the corresponding biochemical reactions.

Using (3.16), (3.15) becomes

$$Cph8c = \frac{T_r [gCph8][PCB]}{km_r + [PCB]}$$

where $T_r = \frac{tr}{d_r dp_r}$, $km_r = \frac{dp_r}{k_i}$

Next, the conservation law for OmpR was introduced:

$$OmpR_{tot} = OmpRp + OmpR + Cph8c = \frac{T_r [gCph8][PCB]}{(km_r (L + 1) + [OmpR])(km_r + [PCB])} + \sum_{i=1}^{3} \left[ 2 \cdot iR_{[i]} + (2i - 1)R_{[i-1]} \right]$$

This allowed for the elimination of OmpR. However when term $a$ in (3.18) for Cph8c was substituted with (3.17), a quadratic equation was obtained and consequently it yielded a large complex solution. For further simplification the following assumption was used

**Assumption 3.9:** The Cph8c complex formation and the phosphorylation reaction are fast and the total amount of free histidine kinase domains is much smaller compared to the population of OmpR.

$$Cph8c \ll OmpR_{tot} - OmpRp - OmpR - \sum_{i=1}^{3} \left[ 2 \cdot iR_{[i]} + (2i - 1)R_{[i-1]} \right]$$

This reduction was done in parallel with the reporter module reduction as in both cases a QSSA was applied for the promoter, based on the assumption A(3.10). $b$ was defined at this point as the unknown function $B \{R_{tot}, OmpR-p, a, b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8, b_9, b_{10}, a\}$. 

$$A(3.9) \Rightarrow OmpR \approx OmpR_{tot} - OmpRp - \sum_{i=1}^{3} \left[ 2 \cdot iR_{[i]} + (2i - 1)R_{[i-1]} \right]$$

Term $b$ in (3.18) represents the amount of OmpR-p molecules bound on the promoter for each of the different occupancy states.
The Reporter module

The final system of equations that can be reduced is the one that describes the transient states of the PompC promoter from completely unoccupied to the active state \( R_{3,3} \) occupied by six OmpR-p molecules.

\[
\begin{align*}
\dot{R}_{0,0} &= +b_s[R_{1,0}] - b_b[\text{OmpRp}][R_{0,0}] + d_{po} \sum (R_{i,j} + R_{j,i-1}) \\
\dot{R}_{1,0} &= -s_s[R_{1,0}][\text{OmpRp}] + s_s[R_{1,1}] - b_b[R_{1,0}] + b_b[R_{0,0}] - d_{po} [R_{1,0}] \\
\dot{R}_{1,1} &= +s_s[R_{1,1}][\text{OmpRp}] - s_s[R_{1,1}] + b_b[R_{1,0}] - b_b[R_{0,0}] - d_{po} [R_{1,1}] \\
\dot{R}_{2,1} &= -s_s[R_{2,1}][\text{OmpRp}] + s_s[R_{2,2}] - b_b[R_{2,1}] + b_b[R_{0,0}] - d_{po} [R_{2,1}] \\
\dot{R}_{2,2} &= +s_s[R_{2,2}][\text{OmpRp}] - s_s[R_{2,1}] + b_b[R_{2,2}] - b_b[R_{0,0}] - d_{po} [R_{2,2}] \\
\dot{R}_{3,2} &= -s_s[R_{3,2}][\text{OmpRp}] + s_s[R_{3,3}] - b_b[R_{3,2}] + b_b[R_{0,0}] - d_{po} [R_{3,2}] \\
\dot{R}_{3,3} &= +s_s[R_{3,3}][\text{OmpRp}] - s_s[R_{3,3}] - d_{po} [R_{3,3}] 
\end{align*}
\]

The aim was to eliminate the intermediate steps and express the promoter in terms of either the vacant or active state and total amount of promoter. Again a QSSA was introduced along with the relevant conversation laws.

**Assumption 3.10:** The phosphorylation and binding of OmpR-p on Ri, is in the order of milliseconds while the LacZ expression proceeds in minutes. Hence the intermediated states of the promoter were assumed quasi stationary.

\[
A(3.10) \quad \left\{ \dot{R}_{0,0}, \dot{R}_{1,0}, \dot{R}_{1,1}, \dot{R}_{2,1}, \dot{R}_{2,2}, \dot{R}_{3,2} \right\} \approx 0
\]

**Assumption 3.11:** The promoter given a saturating amount of OmpR-p will go fully from the inactive to fully occupied state such as its total amount is

\[
A(3.11) \quad R_{\infty} = R_{33} + R_{00} \quad (3.20)
\]

**Assumption 3.12:** The degradation rates \( d_{po}[R_{ij}] \) are negligible for the promoter reaching the fully active state given A(3.11) is not violated.

Then by simple elimination by substitution, of the intermediate occupancy states of the promoter along with the elimination of \( R_{\infty} \) via the conservation law (3.20), (3.21) was obtained

\[
A(3.10) A(3.11) A(3.12) \quad R_{33} = \frac{[\text{OmpRp}][R_{\infty}]}{[\text{OmpRp}] + Kd_o} \quad (3.21)
\]

\[
Kd_o = \frac{b_s b_p b_s}{b_b b_p s_1^3} \quad (3.22)
\]
Where $K_d$ is the combined dissociation constant for PompC (Head et al, 1998). However this simplistic approach was considered relevant only for the range of OmpR-p levels that A(3.11) and A(3.12) are valid. In fact Cai & Inouye (2002) reported that from the large pool of OmpR even in the active state of EnvZ, only a small fraction of it (10% ~350 molecules per cell) will actually exist in its phosphorylated OmpR-p form. In contrast with the native system the synthetic network can be full filing that assumption because

- The RU1012 strain is a ΔenvZ knockout and the Cph8 receptor does not retain the Phosphatase activity. This will naturally shift the ratio in favour of OmpR-p when the system is active.

- The Cph8 receptor is expressed from a high copy plasmid. This will yield more available HK domains for phosphorylating OmpR-p at higher rates than the natural system.

- Other unknown factors like the aerobic conditions and auxiliary phosphorylation/dephosphorylation mechanism might be affected in the knockout strain.

Nevertheless, there is not enough experimental data to conclude about the validity of these assumptions. The main consequence of them is that the reporter in the model will be expressed eventually from the total amount of the promoter as opposed to a wider distribution of promoter occupancy states which will allow only a fraction of the total copy number to express it.

Alternatively to avoid the risk of using assumptions that are valid under very specific conditions, $R_{iii}$ can be re-derived by following the same approach but taking in account only the QSSA assumption A(3.10). The conservation law then becomes

$$R_{tot} = R_{3,3} + R_{3,2} + R_{2,2} + R_{2,1} + R_{1,1} + R_{0,0}$$

$$[R_{3,3}] = \frac{s^R b_{3,1} b_{3,1}[OmpRp]^3}{\sigma_j [OmpRp]^6 + \sigma_1 [OmpRp]^4 + \sigma_2 [OmpRp]^3 + \sigma_3 [OmpRp]^2 + \sigma_4 [OmpRp] + \sigma_5 [OmpRp]} \Rightarrow \sigma_j = G_j \left( s, t, b_{1,2,3,4}, b_{1,2,3,4}, dp_{o, o} \right), \quad j = \{1, 7\}$$

Where the "load" on each promoter in the reduced steady state equation for $R_{3,3}$ is indicated under terms of the denominator in (3.23). Likewise (3.21) both the numerator and the denominator are of the sixth order. The constants $q_i$ are given by polynomial functions of all the association/dissociation and degradation rate constants. The full equation is given in Appendix A (rules) and the numerical simulations obtained are shown in Figure 3.1.6. The latter shows that OmpR-p and PCB:Cph8 complex formation levels are identical for the reduced model and the corresponding quantities of the full mass-action model.

The final output of the module and the model in general is given by the rate of change of the reporter LacZ, where the transcription was reduced to eliminate the mRNA as a dynamic variable similarly to the previous modules.

$$\dot{LacZ} = \frac{tr_{bl}^T_{-3,3}}{d_{z}} [R_{tot}] - dp_{o, o} [LacZ]$$

The expressions for the steady state of each promoter, obtained during (3.23) derivation, can be used to obtain the corresponding polynomial B for (3.18) but also J in (3.11) which are both functions of OmpR-p of the sixth order, total promoter copy number $R_{tot}$ and rate constants (Appendix A, Rules).
Figure 3.1.6. (a) Time Course run for OmpR-p and its transfer function for varying the (b) Red light Intensity and the (c) gene copy number of gPCB. The light receptor module that is responsible for the regulation of the Phosphorylated levels of OmpR shows an excellent agreement between the reduced and the mass-action model as it matched the steady state levels but also showed a very good agreement for the dynamics. This suggested that within the initial parameter set obtained from literature that these numerical simulations were based on, most of the assumptions used for the reduction up to this point were satisfied.
Finalising the reduction - Connecting the modules

After the reduction process the model was reduced down to three ODEs,

\[
P_{\text{PCB}} = \frac{E_{\text{PCB}} - \frac{d_{\text{tr}}}{d_{\text{h}}} g_{\text{PCB}}}{( \text{Heme} + km_i)} \]  

(3.25)

\[
\text{OmpR} = \frac{c_r T_1 g_{\text{Cph8}} (\text{PCB})}{(km_l(L + 1) + ([\text{OmpR}](km + [\text{PCB}])) + \frac{a_1 [\text{OmpR}]}{T_1} - (\text{OmpR})(a_1 + dp_1) + J (R_{\text{OmpR}}, s, s_1, b_1, b_2, b_3, b_4, d_p)} \]  

(3.26)

\[
\text{LacZ} = \frac{d_{\text{tr}}}{d_{\text{h}}} \frac{[R_{\text{tot}}]}{[\text{R}_{\text{tot}}]} - \frac{dp_2 [\text{LacZ}]}{\text{Depletion}} \]  

(3.27)

Up this point the reduced model exhibited a fair agreement with the mass-action and Stochastic model in terms of the Steady state values for the three variables however the dynamics were affected. This is the typical consequence when applying Quasi-Steady-State Approximations based on time scale separation. However, a two or three state model is significantly simpler to handle theoretically than a twenty five state model. Yet a final reduction step can yield a two state model so the basic function of the system can be fitted to a simpler function. In order to reach that, more “aggressive” assumptions had to be introduced such that an analytical expression for OmpR-p’s equilibrium point was reached.

**Assumption 3.13:** PCB can also be considered at steady state since all its interacting variables are treated at their equilibrium points based on the assumptions A3.1- A3.5.

PCB at steady state can be solved and written as

\[
A(3.14) \quad PCB = \frac{1}{2} \left( W_{\text{PCB}} - W_{\text{RP}} - km_i \right. \left. - \sqrt{W_{\text{PCB}}^2 + (-2W_{\text{RP}} + 2km_i)W_{\text{PCB}} + (W_{\text{RP}} + km_i)^2} \right) \]  

(3.28)

\[
W_{\text{RP}} = \frac{t_1 a_1 g_{\text{Cph8}}}{d_{\text{h}}} , \quad W_{\text{PCB}} = \frac{C_p T_p g_{\text{PCB}} [\text{Heme}]}{(\text{Heme} + km_i)}, \quad C_p = \frac{c_p c_p}{d_{\text{h}} (c_p + dp_1)} \]  

Where the parameters \( W_{\text{RP}}, W_{\text{PCB}} \) have a constant value since gene copy numbers and Heme is assumed to be constant while \( C_p \) is the dimensionless parameter describing the combined catalysis reaction from the enzymes. In line with Michaelis-Menten kinetics the product of \( C_p T_p \) and \( g_{\text{PCB}} \) would be the Vmax for the reaction. (3.28) is the solution for PCB when A(3.5) is considered valid. The full solution of (3.8) when A(3.5) is neglected is given in the Appendix A. In either case, the real root of BV and PCB at steady state can be substituted into (3.26) giving a final two state reduced model.
Chapter 3. The Light Receptor Characterisation

Alternatively a reasonable approximation can be written simply as

\[ A(3.14) \quad PCB \approx \frac{C_T gPCB}{[Heme] + km_i} \quad fCph8 \approx 0 \]

where the following assumption was taken into account:

**Assumption 3.14:** PCB is produced fast enough reaching a steady state in the region of thousands of molecules, such that it is always in excess compared to the hundreds of molecules of fCph8. Therefore the binding on the Cph1 domain will not change the PCB concentration significantly.

At this point the function \( f(LR) \) and \( f(LFR) \) and the term \( L \) in equation (3.13) were defined. For simplicity reasons photons were treated as molecules (particles) that can interact with the chromophore’s (PCB) double bond at a ratio of 1:1 (Heyne et al., 2002). The exact mechanics of the multistep transition from the step of photon absorption to the conclusion of the conformational change was not explicitly modelled as these steps proceed with rate constants at the range of fempto- and pico-seconds. In the initial mass-action model the process is collapsed into one step with the photon being absorbed by the holophytochrome and a state transition of the holophytochrome occurs. The light input is measured in light Intensity with Watt/Meter Square as units. The mass-action model was parametrised in terms of molecules hence a converting function of the Intensity into photons per area per unit of time was needed. The photon flux (\( \Phi \)) is related to Intensity and can be written in units of photons per \( m^2 \) per s (Goff et al., 1992). Using the latter and assumption 3.15

**Assumption 3.15:** Having not explicitly modelling the ultrafast multistep dynamics of photo isomerisation the functions \( f(LR) \) and \( f(LFR) \) are assumed as linear functions of the form \( f(I) = u \cdot I \) which return a rate constant with units of \( \text{min}^{-1} \).

the light functions can be expressed as (3.29) and (3.30)

\[
\begin{align*}
Cph8_{Pfr} & \xrightarrow{f(\lambda_{640}) \text{ min}^{-1}} Cph8_{Pfr} \\
\Phi & = \frac{\lambda}{h v} \Rightarrow I \\
\end{align*}
\]

\[
\begin{align*}
f(I_{640}) = z_{640} \frac{A}{h v} & \Rightarrow i = \frac{I_{640}}{z_{640} A} \frac{h v}{I} \tag{3.29} \\
\end{align*}
\]

\[
\begin{align*}
f(I_{730}) = z_{730} \frac{A}{h v} & \Rightarrow I = \frac{I_{730}}{z_{730} A} \frac{h v}{I} \tag{3.30} \\
\end{align*}
\]

where \( z \) is the dimensionless scaling parameter used to describe the efficiency of a given photon, reaching, being absorbed and triggering the photo-isomerisation of PCB. \( A \) is the approximate area of the photon absorbing \( C_{15}=C_{16} \) double bond on PCB (in \( m^2 \)), \( \lambda \) is the light wavelength in m, \( h \) is Plank’s constant in J\text{-}\text{s}, \( v \) the speed of light (m\text{-}\text{s}^{-1}) and \( I \) is the intensity in J\text{-}\text{s}^{-1}\text{m}^{-2}.

\[82\quad \text{Marios Tomazou, PhD Thesis}\quad \text{Imperial College London, September 2014}\]
In the combined function for light $L$ that occurs in the reduced model (3.13) it is worth pointing out the following. $L$ equals the ratio of Pfr to Pr form

$$[\text{Cph8}_{Pfr}] = \frac{f(L) [\text{Cph8}_{Pfr}]}{(k_i + f(LFR)) + dp_r} \Rightarrow [\text{Cph8}_{Pfr}] = L [\text{Cph8}_{Pfr}]$$

$$\Rightarrow L = \frac{f(LR)}{(k_i + f(LFR)) + dp_r} = \frac{u_{640} I}{(k_i + u_{730} I) + dp_r} = \frac{[\text{Cph8}_{Pfr}]}{[\text{Cph8}_{Pr}]}$$

(3.31)

$[\text{Cph8}_{Pfr}] \approx [\text{Cph8}_{PCB}]$ if $u_{640}[I_{640}] \gg (k_i + u_{730}[I_{730}]) + dp_r$

$[\text{Cph8}_{Pfr}] \approx [\text{Cph8}_{PCB}]$ if $u_{640}[I_{640}] \ll (k_i + u_{730}[I_{730}]) + dp_r$ or $u_{640}[I_{640}] \approx 0$

$[\text{Cph8}_{Pfr}] < [\text{Cph8}_{Pr}]$ if $u_{640}[I_{640}] = (k_i + u_{730}[I_{730}]) + dp_r$ or $u_{640}[I_{640}] = u_{730}[I_{730}] \Rightarrow k_i + dp_r$

$[\text{Cph8}_{Pfr}] < [\text{Cph8}_{Pr}]$ if $u_{640}[I_{640}] \approx u_{730}[I_{730}] \Rightarrow u_{640}[I_{640}], u_{730}[I_{730}] \approx k_i + dp_r$

The information available in the literature about the native Cph1 and EnvZ receptors might not apply entirely on the Cph8 chimera as the functional protein structure has been modified. One of these parameters that might have been affected is $k_l$ (the rate constant for the phytochrome transition to the ground state Pr from Pfr when either light wavelength is absent). Figure 3.1.7 shows three possible cases. If $k_l + dp_r$ is larger than $u$ then the system tends to be independent of FarRed light. In addition if given that FarRed light will have an effect on the receptor, if $k_l \ll dp_r$ suggests that in the absence of light the receptor remains locked on the Pfr state and the population is only recycled back to the Pr form through dilution and expression of new receptor at ground state. The only way to decipher for the exact mechanism is by real time measurement of the two states which would require purification of the holophytochrome and advanced equipment able to perform ultrafast time resolution measurements. However this extends beyond the scope of this project.

**Figure 3.1.7.** Transfer function for $L$ versus light for different relations between the light independent terms. The range of values for $I$ for the two light wavelengths were within the published values were the system goes from completely inactive to saturation in terms of the reporter output. Although the function $L$ is not saturating for increasing amounts of light, the availability of either state of the receptor is what saturates the final output. It is implicit that the dilution rate will have a minor effect as either $k_l$ or $f(LFR)$ are of orders of magnitude faster steps (milliseconds compared to minutes-hours for cell division)
Chapter 3. The Light Receptor Characterisation

In order to evaluate the entire reduction process and validity of the assumptions a number of numerical simulations was performed for different light conditions but also for varying gene copy numbers such as to investigate the limits of the assumptions used: Figures 3.1.4 - 3.1.6, 3.1.10-3.1.12

The general conclusion is that the steady state values are conserved fairly accurately for the general output as opposed to the dynamics that are showing a large degree of deviation in some cases. However this deviation is minimal when the system is operating at low copy number of reporter as shown in Figure 3.1.11. The amount of receptor and PCB producing genes appear not to be affecting either the dynamics or steady state values. However these simulations were carried out using a parameter set obtained from the literature. The model was re-evaluated after experimental results became available in order to estimate the value of parameters that the model is more sensitive to.

Refining and Expanding the Model

Load and Retroactivity Aspects

The structure of the proposed model was based on the original system demonstrated by Levskaya et al (2005). However the reconstructed single plasmid version (discussed in Chapter 3 Section 3.2) incorporates apart from the single genomic copy of the PompC:LacZ gene, an additional PompC:eCFP gene on the ColE1 based origin of replication of pL-Cph8FD. The binding mechanics of OmpR-p on these additional copies of the promoter are governed exactly by the same rate constants used for the genomic copy of the promoter, however they are introducing approximately twelve hundred additional binding sites for OmpR-p. This ‘sink’ is expected to have an effect on the pool of free OmpR-p resulting into a retroactive effect (Del Vecchio et al, 2008). If this effect is significantly strong it is understood that LacZ will be expressed at a lower rate when OmpR-p becomes limited. In order to obtain a functional model that can account for this effect when the system is connected to n downstream genetic devices, single or multicopy the reduced model was re-derived with an additional set of promoter states \( R^{n}_{tot} \) where \( n \) is the index of the additional copy of the promoter that can control the same or a different downstream gene. The corresponding equation to (3.26) yields an additional term which is the function J but with \( R^{n}_{tot} \) the copy number of the new additional promoter:gene. Similarly in (3.19) an additional term \( B^{n} \) occurs.

\[
\Rightarrow OmpR \approx OmpR^{n}_{tot} - OmpRp - \left( B^{1} + B^{2} + ... B^{n} \right) \tag{3.32}
\]

\[
\begin{align*}
OmpRp &= \frac{c_{o}}{\left( km_{2}(L+1) + \left[ OmpR \right][PCB]\right) + \left( \frac{\left[ OmpR \right][PCB]}{km_{1} + \left[ PCB \right]} \right)} \tag{3.32} \\
&+ \frac{a_{o}[OmpR]}{km_{1} + \left[ PCB \right]} J^{1} + J^{2} + ... J^{n} \tag{3.33}
\end{align*}
\]

\[
B^{n} \left( R^{n}_{tot}, OmpRp, s, s_{a}, b_{1}, b_{2}, b_{3}, b_{4}, b_{5}, m_{1}, m_{2}, d_{m}, d_{p}, \right) \}
\]

\[
J^{n} \left( R^{n}_{tot}, OmpRp, s, s_{a}, b_{1}, b_{2}, b_{3}, b_{4}, b_{5}, m_{1}, m_{2}, d_{m}, d_{p}, \right) \}
\]

\( n = \{1, 2... N\} \)
These terms point to a simple way to expand the model to a system where a number of \( n \) promoters are being used. However the system output for either of the promoters will be decreased as the total copy number of either promoter rises. Figure 3.1.8 shows numerical simulations that were used to determine how retroactivity increases as a function of promoter copy number and where the optimal point is for maximum total protein expression. The numerical simulations results suggested the relations in (3.34) and (3.35)

\[
\text{OmpR}_{\text{tot}} = \text{OmpR} + G_{\text{tot}}
\]

\[
G_{\text{tot}} = \sum_{i=1}^{n} B_i
\]

For 1 promoter \( R_{\text{tot}}^1 \) at steady state

\[
R_{\text{tot}}^1 \approx \infty \Rightarrow \text{LacZ}^1 \approx 0
\]

\[
6R_{\text{tot}}^1 \approx \text{OmpR}_{\text{tot}} \Rightarrow \max \text{LacZ}^1
\]

\[
R_{\text{tot}}^1 \approx 0 \Rightarrow \text{LacZ}^1 \approx 0
\]

For \( n \) promoters \( R_{\text{tot}}^n \) at steady state

\[
G_{\text{tot}} \approx \infty \Rightarrow \text{LacZ}^n \approx 0
\]

\[
6G_{\text{tot}} \approx \text{OmpR}_{\text{tot}} \Rightarrow \max \sum_{i=1}^{n} \text{LacZ}^i
\]

\[
G_{\text{tot}} \approx 0 \Rightarrow \text{LacZ}^n \approx 0
\]

Where (3.34) and (3.35) suggest that the optimal copy number should be approximately six times less than the total Phosphorylated OmpR-p in order to get the maximum possible output levels under induced conditions. In case of multiple downstream networks the condition is the same for the combined output from all the PompC controlled genes. However it was not possible to derive the above conclusion analytically.

**System Behaviour in the Absence of PCB**

An additional extension of the model aimed on capturing correctly the behaviour of the physical system when the PCB producing enzymes are not expressed \((\text{gPCB} = 0)\). The physical apophytochrome (Cph8) in the absence of chromophore will still dock on the membrane and will phosphorylate OmpR but it will be insensitive to light (Levskaya et al., 2005). In other words the system is constitutively ON unless PCB is bound on it to absorb light and cause the conformational change to render it inactive. In the mass-action model the addition of the following reactions was needed

\[
f\text{Cph8} + \text{ompR} \xrightleftharpoons[k_2]{k_1} f\text{Cph8c} \xrightarrow{co} f\text{Cph8} + \text{ompR}
\]

\[
f\text{Cph8c} \xrightarrow{dp} \emptyset
\]

\[
f\text{Cph8c} + \text{PCB} \xrightarrow{k_1} \text{Cph8c}
\]
Chapter 3. The Light Receptor Characterisation

Figure 3.1.8. Retroactive effect of the PompC promoter copy number on the reporter output. (a) In a one reporter system, at steady state, increasing the total number of OmpR-p binding sites by increasing the copy number of the promoter will initially result to a higher LacZ output. However once the number of OmpR-p sites, six for each promoter copy, approach the total number of phosphorylated OmpR-p molecules the steady state levels of LacZ reach their maximum. Increasing the promoter copies even further will result to a drop of LacZ. (b) The same applies in a two reporter/gene system under the same promoters where the maximum combined levels for both molecules is approximately at the point where the total amount of binding sites from both promoters approaches the total amount of available OmpR-p. However this observation was not shown analytically and appeared to be not applicable when basal expression and auxiliary OmpR-p phosphorylation was considered.
The mass-action ODEs are corrected as follows (marked with A are the new occurring terms)

\[
P_{\text{CB}} = c_r [P_{\text{CyAc}}] - d_r [P_{\text{CB}}] - k_i \left( f_{\text{Cph8}} + \frac{[f_{\text{Cph8c}}]}{[P_{\text{CB}}]} \right) [P_{\text{CB}}]
\]

\[
f_{\text{Cph8}} = \frac{t_l [m_{\text{Cph8}}] - k_i \left( f_{\text{Cph8}} \right) [P_{\text{CB}}] - d_p \left( f_{\text{Cph8}} \right) - k_i \left( \text{OmpR} \cdot f_{\text{Cph8}} \right) + k_i \left( f_{\text{Cph8}} \right) + c_r \left( f_{\text{Cph8c}} \right)}{d}
\]

\[
f_{\text{Cph8c}} = \frac{k_i \left( \text{OmpR} \cdot f_{\text{Cph8}} \right) - k_i \left( f_{\text{Cph8c}} \right) - d_p \left( f_{\text{Cph8c}} \right) - c_r \left( f_{\text{Cph8c}} \right) - k_i \left( f_{\text{Cph8c}} \right) [P_{\text{CB}}]}{d}
\]

\[
\text{Cph8c} = \frac{k_i \left( \text{OmpR} \cdot [\text{Cph8}_{i_0}] - k_i \left( f_{\text{Cph8c}} \right) - d_p \left( \text{Cph8c} \right) - c_r \left( \text{Cph8c} \right) + k_i \left( f_{\text{Cph8c}} \right) [P_{\text{CB}}]}{d}
\]

\[
\text{OmpR} = \frac{t_l \left( m_{\text{OmpR}} \right) - k_i \left( \text{OmpR} \cdot [\text{Cph8}_{i_0}] \right) + k_i \left( \text{Cph8c} \right) - d_p \left( \text{OmpR} \right) - a_r \left( \text{OmpR} \right) + a_r \left( \text{OmpR} \right) - k_i \left( \text{OmpR} \cdot f_{\text{Cph8}} \right) + k_i \left( f_{\text{Cph8}} \right)}{d}
\]

\[
\text{OmpR}_p = c_r \left( [\text{Cph8c}] + \frac{[f_{\text{Cph8c}}]}{[P_{\text{CB}}]} \right) + a_r \left( \text{OmpR} \right) + \sum \frac{b_i}{[\text{R}_{i_0}]} + \sum \frac{c_i}{[\text{R}_{i_0}]} - \left( \sum \frac{b_i}{[\text{R}_{i_0}]} + 1 \right) \left( [\text{Cph8}] + d_p + a_r \right)
\]

The conservation laws, steady states and OmpR-p rate become

\[
\text{Cph8}_{i_0} = \frac{f_{\text{Cph8}} + f_{\text{Cph8c}} + \text{Cph8}_{i_0} + \text{Cph8}_p + \text{Cph8c}}{f_{\text{Cph8c}}}
\]

\[
\text{OmpR}_{i_0} = \text{OmpR} + \text{OmpR} + \text{Cph8c} + \frac{\sum B_i}{f_{\text{Cph8c}}}
\]

\[
\Rightarrow f_{\text{Cph8}}_{i_0} = t_l \left( m_{\text{Cph8}} \right) - f_{\text{Cph8}}_{i_0} \left( k_i \left( [P_{\text{CB}}] + d_p \right) \right)
\]

At Steady State

\[
f_{\text{Cph8}}_{i_0} = \frac{t_l \left( m_{\text{Cph8}} \right)}{\left( k_i \left( [P_{\text{CB}}] + d_p \right) \right)} = \frac{t_l \cdot t_r \cdot [g_{\text{Cph8}}]}{d_r \left( k_i \left( [P_{\text{CB}}] + d_p \right) \right)}
\]

\[
\text{Cph8}_{\text{pCB}} = \frac{T_r \left( g_{\text{Cph8}} \right) \left( [P_{\text{CB}}] \right)}{\left( [P_{\text{CB}}] + k_m \right) \left( k_i \left( [P_{\text{CB}}] + \left( f_{\text{Cph8}} \right) \right) \right)}
\]

\[
f_{\text{Cph8c}} = \frac{\left( [\text{OmpR}] \cdot [\text{f_{\text{Cph8c}}}] \right) k_r \cdot t_r \cdot d_p}{\left( [\text{PCB}] \right) + \left( k_m \right) \left( \left( [\text{PCB}] \right) + \left( f_{\text{Cph8}} \right) \right)}
\]

\[
\text{Cph8c} = \frac{\left( k_m \cdot T_r \cdot [\text{OmpR}] \cdot [g_{\text{Cph8}}] \right) \left( \left( L + 1 \right) \cdot d_p \right) + \left( [\text{OmpR}] \right) \left( [\text{PCB}] \right) k_m + \left( k_m \right)}{\left( [\text{PCB}] + k_m \right) \left( k_m \cdot \left( L + 1 \right) + \left( [\text{PCB}] \right) \cdot \left( f_{\text{Cph8}} \right) \right) + \left( [\text{PCB}] \right) \cdot \left( k_m \right) + \left( [\text{PCB}] \right) \cdot \left( f_{\text{Cph8}} \right) + \left( k_m \right)}
\]

\[
k_m = \frac{k_i}{k_i}
\]

\[
\text{OmpR}_p = \frac{c_r \left( [\text{Cph8c}] \right) + c_r \left( \frac{[f_{\text{Cph8c}}]}{[P_{\text{CB}}]} \right) + a_r \left( [\text{OmpR}] \right) - \left( a_r \left( [\text{OmpR}] \right) \right) + \sum \frac{b_i}{[\text{R}_{i_0}]} + \sum \frac{c_i}{[\text{R}_{i_0}]} - \left( \sum \frac{b_i}{[\text{R}_{i_0}]} + 1 \right) \left( [\text{Cph8}] + d_p + a_r \right)}{f_{\text{Cph8c}}}
\]

Interrogating the model for varying the copy numbers of genes and light Input, the limits were the reduced model is true for the steady state with respect to the mass-action model and within the standard deviation bounds for the stochastic counterpart were investigated. The only deviation found was when $G_{\text{tot}}$ increases for values larger than the optimal output (approximately 55 copies for the specific parameter set). The term generating that inconsistency is the Equation (3.19) for the OmpR steady state (Figure 3.1.11)
To correct for that the assumption A(3.9) was discarded, considering only the QSSA (the phosphorylation is fast), introduced new conserved quantities that allowed to find an analytical solution for OmpR at steady state with standard Michaelis-Menten kinetics. The equation becomes quadratic and was solved for the real root (3.38)

\[
\frac{C_{\text{ph8}_{\text{tot}}}}{C_{\text{ph8c}_{\text{tot}}}} = \frac{f_{\text{ph8}} + C_{\text{ph8p}} + C_{\text{ph8p2}} + C_{\text{ph8c}} + f_{\text{ph8c}}}{C_{\text{ph8c}_{\text{tot}}}}
\]

\[\Rightarrow C_{\text{ph8c}_{\text{tot}}} = k_2[OmpR]C_{\text{ph8}_{\text{tot}}} - [C_{\text{ph8c}_{\text{tot}}}](dp_r + k_c + c_r)\]

At Steady State

\[C_{\text{ph8c}_{\text{tot}}} = \frac{[OmpR][C_{\text{ph8c}_{\text{tot}}}]}{[OmpR] + km_2}\]

\[\Rightarrow OmpR_{\text{tot}} = OmpR + OmpR_{\text{p}} + C_{\text{ph8c}_{\text{tot}}} + \sum_{i=1}^{5} B' \Rightarrow \text{solve, w.r.t. OmpR}\]

\[
OmpR = \frac{-[OmpR] - \sum_{i=1}^{5} B' - km_2 - [C_{\text{ph8}_{\text{tot}}} + [OmpR_{\text{tot}}] + \frac{[C_{\text{ph8c}_{\text{tot}}} + [OmpR_{\text{tot}}] + \frac{1}{2}[C_{\text{ph8c}_{\text{tot}}} + [OmpR_{\text{tot}}] + \frac{1}{2}[OmpR_{\text{tot}}] + \frac{1}{2}[OmpR_{\text{tot}}] + \frac{1}{2}[OmpR_{\text{tot}}]}}{[OmpR] + km_2}}{3.38}
\]

**Basal Expression**

Finally the model was expanded to include basal expression or leakiness from the R \(ii\) promoter.

**Assumption 3.16:** For each of the unoccupied states of \(R_{ii}^{*}\), polymerase may still bind and transcribe the gene with a much lower rate constant \(b\) which is proportional to the tr \(j\) (b\(v^*\)tr \(j\))

For the non fully occupied states let \(R_{\text{tot}}^{1} = R_{3,2}^{1} + R_{2,2}^{1} + R_{1,2}^{1} + R_{1,1}^{1} + R_{0,0}^{1} = R_{\text{tot}}^{1} - R_{3,3}^{1}\)

\[
R_{\text{tot}}^{1} \xrightarrow{b=v^*tr_j} mLacZ^{1} + R_{\text{tot}}^{1}
\]

\[\Rightarrow LacZ^{1} = \frac{tl_{tr_j}}{d_z} [R_{3,3}^{1} + v(R_{\text{tot}}^{1} - R_{3,3}^{1})] - dp_z [LacZ^{1}] = \frac{tl_{tr_j}}{d_z} \left((1 - v) \frac{R_{3,3}^{1}}{R_{3,3}^{1}} + v[R_{\text{tot}}^{1}]\right) - dp_z [LacZ^{1}]\]

With this the model reduction was concluded. The final version used for all the following analysis and experimental data fitting was given by the rates for OmpR-p (3.39) and LacZ (3.40)

The numerical simulations shown in Figures 3.1.10-3.1.12 were carried out using the SimBiology 4 package in MATLAB and numerical data were processed by custom scripts. A report of the computational model is given in Appendix A. Most of the runs were carried simultaneously for the mass-action and reduced model using the Sundials solver while the corresponding stochastic model was run under the same initial conditions using the stochastic solver which is an implementation of the Gillispie algorithm. Time courses, Transfer function and sensitivity analysis were re-evaluated after the experimental data shown in section 3.2.
3.1.4. Model Simulations

A summary of the performance of the reduced and mass-action model with numerical simulations under different conditions along with sensitivity analysis is shown in Figures 3.1.10-3.1.12. In addition the following points regarding the model reduction are summarised:

- The reduced model within the biological acceptable range of values for the inputs evaluates in complete agreement with the mass-action model, the steady states for the output for LacZ.
- It accounts for the retroactive effect of one or multiple genes that are controlled by a PompC promoter that can be either on a single genomic copy or multiple vectors.
- For low total copy number of promoter the dynamics of both models overlay.
- For high copy numbers the steady states match but there is a slight loss in the dynamics during the transitions from low to high output or inversely, with the reduced model being faster. The apparent deviation is in the order of minutes.
- A two state model was reached from the reduction of the initial ~26 dimensional mass-action version. At the same time the main terms and parameters that represent possible biological ‘dials’ are maintained. This can be used to fine tune the system for optimal conditions depending on the application.
- All the parameters in the resulting reduced model are combinations of rate constants that can be tracked back to their corresponding biochemical reactions. Hence all the parameters have a physical meaning and no artificial time delays or scaling parameters where used with the exception of \( z \) in the light function (3.29).
- From the reduced version of the model the system shares similarities with the function of a low-pass filter in electronics.
Figure 3.1.9. Diagram of the reduced model

Figure 3.1.10. Time course simulation with induction at t=700. The reduced model predicts almost identical dynamics and steady state levels for LacZ under a single PompC copy, compared with the full mass-action and stochastic counterpart. The simulation is shown for t>500 since the mass-action/Stochastic models need to evolve to match the initial conditions of the reduced one before having a meaningful comparison.
Effect of PompC Promoter Copy number ($R_{tot}$) to the Dynamics of the Reduced and Mass-Action Model

The deviation in the transition dynamics between the reduced and mass-action when light is applied, is larger for an increasing number of promoter copy number. In particular as the copy number is equal or greater than the optimal point for maximum LacZ expression (shown in Figure 3.1.8) is as high as 100 minutes, as opposed to tens of minutes when the copy number is kept lower than ~50. This indicates that when OmpR-p becomes limited the QSSA assumption $A(3.10)$ is violated.

The logarithmically decreasing function of LacZ steady state levels for an increasing amount of light.

**Figure 3.1.11.** (a) The deviation in the transition dynamics between the reduced and mass-action when light is applied, is larger for an increasing number of promoter copy number. In particular as the copy number is equal or greater than the optimal point for maximum LacZ expression (shown in Figure 3.1.8) is as high as 100 minutes, as opposed to tens of minutes when the copy number is kept lower than ~50. This indicates that when OmpR-p becomes limited the QSSA assumption $A(3.10)$ is violated. (b) The logarithmically decreasing function of LacZ steady state levels for an increasing amount of light.
Chapter 3. The Light Receptor Characterisation

Figure 3.1.12. Sensitivity analysis results. (a) In order to obtain a qualitative picture of how sensitive the LacZ expression levels are to the large amount of parameters that the mass-action model was based on, four groups were identified. In group A all the rate constants associated with transcription, translation and degradation processes are shown to be the ones affecting the most the final output. Apart from the obvious that LacZ expression rates will have a relatively high sensitivity degree, the expression of Cph8 (tr,tlr,dpr,dr) and OmpR (tro,tlo,do,dpo) related rate constants are shown to be lying on low to intermediate relative sensitivities. Overall LacZ appears to be robust for changes on any of the terms related to PCB in any of the four groups. In group B (enzymatic reactions) the phosphorylation rate constant for OmpR (co) is the one with the most impact followed with the association-dissociation of OmpR to Cph8 (k2,k2r). Auxiliary dephosphorylation (ar) and secondary PompC operator binding rates (s,sr) are the ones with the higher values in group C, though 6-fold lower than tlz. From the copy number group D, Cph8 and OmpR copy numbers have low sensitivity while PCB related species (BV,Heme and gPCB) seem not to affect LacZ levels when perturbed. On the contrary the gene copy number for Pompc::LacZ holds the highest relative sensitivity. (b) The fact that during the model reduction many of the rate constants were collected into lumped parameters there is no real way to compare the two models in terms of relative sensitivity. However in the attempt to obtain values for all the parameters the sensitivity matrix plot reveals similar trends for both the initial mass-action rate constants and their corresponding lumped parameters (e.g. trz,tlz,dz compared with Tz). In all groups the values shown is the normalised sensitivity (dimensionless).
**Multichromatic Model Extension**

To include the CcaS operation in the same host the following additional reactions were introduced

**CcaS expression and state transition**

\[
gCcaS \xrightarrow{\nu_g} mCcaS + gCcaS \\
mCcaS \xrightarrow{\nu_m} fCcaS + mCcaS \\
fCcaS + PCB \xrightarrow{k_k} CcaS_{Pg} \\
CcaS_{Pg} \xrightarrow{f(GL)} CcaS_{Pr}
\]

**CcaR expression and auxiliary phosphorylation**

\[
gCcaR \xrightarrow{\nu_g} mCcaR + gCcaR \\
mCcaR \xrightarrow{\nu_m} CcaR + mCcaR \\
CcaR \xrightarrow{a} CcaRp
\]

**CcaR:CcaS complex formation including phosphatase activity**

\[
CcaS_{Pr} + CcaR \xrightarrow{k_{k_1}} CcaSc \xrightarrow{c} CcaS_{Pr} + CcaRp \\
fCcaS + CcaRp \xrightarrow{k_{k_2}} fCcaSc \xrightarrow{c} CcaSc + CcaR \\
CcaS_{Pg} + CcaRp \xrightarrow{k_{k_2}} CcaS_{Pg} \xrightarrow{c} CcaS_{Pg} + CcaR
\]

**CcaR:CcaS complex formation including phosphatase activity**

\[
CcaSc \xrightarrow{f(GL)} CcaSc
\]

**Degradation**

\[
mCcaS \xrightarrow{d_m} \emptyset \\
mCcaR \xrightarrow{d_m} \emptyset \\
\{fCcaS, fCcaSc, CcaS_{Pr}, CcaSc, CcaS_{Pg}, CcaSc\} \xrightarrow{d_p} \emptyset \\
\{CcaR, CcaRp\} \xrightarrow{d_p} \emptyset
\]

**CcaRp and PcpcG2 (P,) promoter interaction and reporter expression**

\[
CcaRp + P_0 \xrightarrow{g} P_1 \\
CcaRp + P_1 \xrightarrow{g} P_2 \\
\{P_1, P_2\} \xrightarrow{d_p} P_0 \\
P_2 \xrightarrow{n} mLacZ + P_2 \\
\{P_0, P_1\} \xrightarrow{v\epsilon r} mLacZ + \{P_0, P_1\}
\]
Similarly to the Cph8 model the multichromatic extended version was reduced only to include the rates of the phosphorylated state of CcaR (CcaRp) and the reporter. All other variables were solved for their steady state values and substituted into the two remaining ODEs.

\[
CcaS_0 = \frac{f_{CcaS} + f_{CcaSc} + CcaS_0 + CcaSc + CcaS_y + CcaSc = T_{[gCcaS]}}{CcaS_y}\]

\[
B_i = g^*[CcaRp][P_{ri}](2g[CcaRp] + dp_i + gr_i)
\]

\[
P_i = g^*[CcaRp] + g(g_i + 2dp_i)[CcaRp] + (dp_i + g_i)^2
\]

\[
H = g^*[CcaRp][P_{ri}][2g[CcaRp] + dp_i + g_i]
\]

\[
[CcaR] = \frac{1}{2} \left( -k_i[CcaS_0] - [CcaR] + [CcaR_0] - km_i + \sqrt{k_i[CcaS_0]^2 + 2k_i([CcaRp] + B_i + km_i - [CcaR_0])[CcaS_0] + ([CcaRp] + B_i - km_i - [CcaR_0])^2} \right)
\]

\[
[fCcaScr] = \frac{dp_i[CcaR][CcaS_0]}{(k_i[PCB] + dp_i)(km_i[PCB] + [CcaR] + km_i)}
\]

\[
[fCcaSc] = \frac{d_{iL}[gCcaS] + t_{LR}[gCcaS] + t_{LR}km_i[PCB][gCcaS]}{(d_{iL}[km_i][PCB] + dp_i)(k_i[CcaR] + k_i[PCB] + dp_i + km_i)}
\]

\[
[CcaS_y] = \frac{k_i[PCB][CcaS_0][k_i[CcaR] + k_i[km_i + f(RL) + l_i)] + G[CcaR](f(RL) + l_i))}{(G + 1)(k_i[PCB] + dp_i)(k_i[CcaR] + k_i[km_i + f(RL) + l_i][CcaRp] + k_i[km_i^2 + (k_i)[CcaR] + f(GL) + f(RL) + l_i)km_i + f(GL)[CcaR])}
\]

\[
[CcaSc] = \frac{k_i[PCB][CcaS_0][k_i[km_i + km_i(CRea * k2 + f(GL) + f(RL) + l_i)] - (CcaR)G[f(RL) + Gk_i - f(GL))]}{(G + 1)(k_i[PCB] + dp_i)(k_i[CcaR] + k_i[km_i + f(RL) + l_i][CcaRp] + k_i[km_i^2 + (k_i)[CcaR] + f(GL) + f(RL) + l_i)km_i + f(GL)[CcaR])}
\]

The PCB expression (3.28) was corrected to include its depletion from CcaS. In order to include the CI repression the Rii (PompC) promoter was set to express the CI repressor which can bind on the Pλ and repress LacZ expression in compliance with the published multichromatic system (Tabor et al, 2011). The following reactions were considered

\[
R_{1,3} \xrightarrow{v_3} mCI + R_{1,3}
\]

\[
\{R_{0,0}, R_{1,0}, R_{1,1}, R_{2,1}, R_{2,2}, R_{3,2}\} \xrightarrow{v_{3r}} mCI + \{R_{0,0}, R_{1,0}, R_{1,1}, R_{2,1}, R_{2,2}, R_{3,2}\}
\]

\[
mCI \xrightarrow{d_3} CI + mCI
\]

\[
CI + P\lambda_0 \xrightarrow{r_{CI}} P\lambda_1
\]

\[
CI + P\lambda_1 \xrightarrow{r_{CI}} P\lambda_2
\]

\[
P\lambda_0 \xrightarrow{v_{3r}} mLacZ + P\lambda_0
\]

\[
\{P\lambda_1, P\lambda_2\} \xrightarrow{v_{3r}} mLacZ + \{P\lambda_1, P\lambda_2\}
\]

\[
mCI \xrightarrow{d_3} \emptyset
\]

\[
CI \xrightarrow{d_3} \emptyset
\]

\[
\{P\lambda_1, P\lambda_2\} \xrightarrow{d_3} P\lambda_0
\]

The reduced version for the Pλ expressed LacZ is of the same form as P (PcpcG2) with the difference that instead of Pλ2 the steady state expression of unoccupied Pλ0 is used.
Figure 3.1.13. Time course Simulation of the multichromatic control system comprising LacZ expression under CcaS control, CI expression under Cph8 control and represses another copy LacZ. For clarity reasons the green light control LacZ is referred as GFP in this figure. In reality the only observable quantity is the total output for LacZ (Grey line). This does not allow direct measurement of the levels from each promoter however some of these can be calculated graphically. \( \alpha \) is the basal expression when light is OFF and should be approximately equal to the total basal expression from CcaS controlled LacZ since this gene is expressed from a high copy number plasmid as opposed to PompC:LacZ which is found on the low copy pSC101 (~5 copies). \( \beta \) is the difference of the maximum levels when both systems are ON with the levels of when only Cph8 controlled LacZ is ON. \( \gamma \) is the levels of CcaS controlled LacZ plus the cross activated by green light Cph8 controlled LacZ. From these we can simply calculate the following:

- Cph8 controlled LacZ levels under green light: \( \beta - \gamma \)
- Cph8 controlled LacZ levels under red light: \( \beta - \alpha \)
- CcaS controlled LacZ levels under both or only green light: \( \delta - \beta \)

Figure 3.1.14. Transfer function of red and green light at steady state
Chapter 3. The Light Receptor Characterisation

3.2. The Experimental System

3.2.1. The Light Source

For illuminating the bacterial cultures, a single LED circuit was build according to the circuit diagram shown in methods (Chapter 2 section 2.3.1). The total power output of the specific LEDs was chosen to be relatively high, 10W, compared to the required light intensity to activate the Cph1 and CcaS photoactive domains (0.1 W/m²), so the system can be assessed for a wide dynamic range of light intensities.

The power output, was measured using a power meter (kindly provided by the photonics group at Imperial College London) for varying displacement from the LED. As shown in Figure 3.2.1, the function of the Intensity for a varying distance from the source was fitted on a function of the form \( f(x) = a/x^n \) where \( a \) and \( n \) are coefficients calculated for each wavelength. However, the exponents were deviating from a strict inverse square law and ranged from 1.14 - 2.1. This is more likely due to human error in measurements, or abnormalities on the surface of the LED.

For an alternative way of varying the intensity, neutral light filters were acquired. The filters available according to the manufactures specifications had 5% 25% 50% 75% and 90% cutoff coefficients for light at any wavelength. All intensity values were corrected for an approximately 1% loss in intensity when light was shined through the 1 cm transparent Plexiglas incubator ceiling as described in Materials and Methods (section 2.1.3). Finally a customised incubator equipped with an LED array was available for a limited amount of time through the run of this project, kindly provided by Dr Eric Davidson.

3.2.2. The Light Receptor Construct

The original Cph8 device from Levskaya et al (2005) requires two plasmids in a ΔenvZ strain (RU1012). The pCph8 that harbours the Cph8 receptor under a Ptet0-1 promoter, a coE1 based origin of replication and a chloramphenicol resistance cassette. The second plasmid harbours the ho1 and pcyA genes in a bicistronic arrangement under the Plac/ara-1 promoter, a P15A origin and an Ampicillin resistance cassette.

Although the system operates under the basal expression from these promoters, the fact that they are regulated by TetR, LacI and AraC, is a significant limitation for coupling this system with other networks, that are most likely to be using the same regulators (Discussed in Chapter 4.1). This will result in crosstalk of the downstream network regulation, with the light sensing machinery. In addition, the use of two of the most widely used origins or replications, CoE1 and p15A, requires any network to be controlled by light to be cloned into one of them since origins within the same incompatibility group are not stable in the same bacterial host. Hence in order to build a ‘plug and play’ version of Cph8, the three required genes (cph8,ho1,pcyA) were reassembled into one vector (pB1AK3: CoE1, Ampicillin and Kanamycin resistance cassettes) under constitutive expression (J23109 and J23110 promoters for cph8 and ho1/pcyA genes respectively). Finally, the PompC promoter was cloned upstream of an ecfp gene, to act as a reporter. The plasmid map of pL-Cph8FC harbouring the above is shown in Figure 3.2.2.
Figure 3.2.1. For each LED the power output was measured for varying distance from the source. The devices approximate but are not matching the inverse square law. Data points are represented by the markers while fitting curves are shown in a continues blue line flanked by the 99% confidence bounds shown as black dashed lines. The arrow in the red light curve marks the distance used in the incubator for a 0.5 W/m² intensity used in Chapter 5 (section 5.3)
Chapter 3. The Light Receptor Characterisation

The system at this point was assayed for two reporters. Initially, the cultures were grown under red light and dark conditions followed by a lysis step by sonication and scanned for eCFP expression levels, using a fluorescence scanner (Fluoromax 3) as described in Methods (Section 2.1.2). As seen in Figure 3.2.3 (b) the cultures grown in the dark, showed a distinctive blue colour, as opposed to the brown-blue appearance of the ones grown under red light. However, the phytochrome-bound PCB, also appeared blue. Thus, in order to determine the exact eCFP levels, it was necessary to obtain the emission-excitation fluorescent spectra, as shown in Figure 3.2.3 (c) for each sample.

Due to the fact that the RU1012 strain bears a genome integrated PompC::LacZ fusion, all of the following measurements on the Cph8 device, were taken by evaluating the LacZ levels by Miller assays (LacZ activity with ONPG as a substrate Figure 3.2.4 (a)). This, allowed for relatively faster and more practical assays, as no eCFP specific filters, were available in the lab’s equipment ( plate readers, microscope and FACS scanner) at the point when the system was being assayed. Also, measuring the LacZ expressed from the genomic integrated PompC promoter, allowed for direct comparison with the originally published dual plasmid version of the system (Levskaya et al, 2005)

3.2.3. Cph8 Experimental Characterisation Results

After verifying its functionality, the reconstructed system was characterised for its static and dynamic response, using the genomic PompC:LacZ reporter in the RU1012 strain. Data were collected for both versions of the system (i.e. reconstructed and original dual plasmid). Some general aspects of growth conditions, for all light responsive systems throughout this work, were taken in account. First, all assays performed and generated higher fold differences, when assayed one or two days at the latest after the transformation of cells with the plasmids harbouring the light responsive systems. Longer delays, resulted in high variability between samples or even complete loss of light sensitivity. The second major environmental factor was aeration. Under anaerobic conditions, in tightly sealed tubes or microplates, the samples did not yield strong light sensitive responses. This is due to general metabolism shifts, for cells growing under anaerobic conditions, but also the fact that the PCB-producing enzymes, require sufficient levels of oxygen for the reduction of heme to PCB. Unless otherwise specified, for all static responses the measurement time point was set at 4 hours after induction.

Static Characterisation

Samples were collected and assayed for LacZ expression (Methods 2.1.2 Miller Assays). As shown in Figure 3.2.4 the original dual plasmid system from Levskaya et al (2005), yielded approximately 1250 MU (miller units) for the dark and 200 MU for light induced conditions. The high activity values in the dark, were consistent with the published figures (Levskaya et al, 2005) however the basal levels of expression were at least 3-fold higher than the reported ~50 MU. This might be due to a number of reasons that are discussed below.
Figure 3.2.2.  (a) Plasmid maps of the pCph8 chimeric light receptor introduced by Levskaya et al 2005 and the Ho1 and PcyA enzymes on the pL-PCB (Amp resistant) demonstrated by Gambetta and Lagarias to be producing a functional chromophore (PCB) (b) Plasmid map of pL-Cph8FD reconstracted single plasmid system with a 1% Agarose gel image of the insert and vector after a double digestion with EcoRI and PstI
Figure 3.2.3. Phenotypic characterisation of Cph8. (a) Quenched Miller assay reactions. The yellow colour is characteristic for ONP (o-nitrophenol) which is the product of ONPG (o-nitrophenyl-β-D-galactoside) cleavage by LacZ. For each pair the left sample was grown in the absence of light while the right sample was grown under continuous 640nm illumination (0.5 W/m²). The left pair is the reconstructed device (pL-Cph8FD) in RU1012 strain, the middle pair is the two plasmid version of the device and the rightmost pair is RU1012, with only the pL-PCB plasmid. (b) Lysates of RU1012 cultures, harbouring the pL-Cph8FD plasmid grown in the absence (left) or presence (middle) of light. The right sample is a negative control of bare RU1012. (c) The fluorescence spectral profile of pL-Cph8FD strains cell lysate. The characteristic emission peaks at 476 nm and 497 nm confirm that the observed fluorescence is due to eCFP expression and not the blue holophytochrome (Cph8-PCB complex). The light red curve shows the corresponding excitation spectra. The twin sharp peaks around 440 nm for emission and 480 nm for excitation, is the point that the monochromator wavelength matches the bandpass filter’s wavelength during the emission/excitation scans. The inlet shows the corresponding emission eCFP spectra, of non-lysed cells in LB growth medium for samples grown in the dark (black curve) and under saturating amount of red light, ~0.5 W/m² (Red curve).
Figure 3.2.4. (a) Steady state characterisation for LacZ activity (Miller assays) and comparison between the reconstructed device (pL-Cph8FD) and the original dual vector version of the system. It is clear that the latter exhibits a larger fold difference between the active and inactive state for the specific reporter. The pL-PCB (negative control) shows basal levels of expression, while the pL-Cph8 only plasmid (positive control) is active independently of the light conditions, due to the fact that no PCB is present to drive the conformational change upon excitation with 635 nm light. (b) LacZ expression levels as a function of 635 nm light. The red markers represent experimental data points while blue continuous curves is the curve fitting, for a general repression hill function of the form \([a*b^n]/(b^n+(I635)^n)+c\). The calculated coefficients for the specific fit are: \(a = 915.6, b = 0.006511, c = 218.4, n=2.44\)
Chapter 3. The Light Receptor Characterisation

- The cells go into anaerobic conditions to a degree that PCB is not sufficiently produced resulting into an insensitive to light receptor

- Auxiliary OmpR-p phosphorylation mechanisms are activated under stress conditions, either due to lack of oxygen, resources or even the continuous strong red light illumination. A previously studied contributor to the non-cognate phosphorylation of OmpR, is the CpxA receptor (Groban et al, 2009)

- Human error is also a factor to consider, since Miller Assays is a multistage biochemical reaction assay. pH of Solutions, degraded ONPG, accurate reaction timing, quenching and OD normalisation contribute to some extent to the error.

Similarly to the original system, the pL-Cph8FD from this work, exhibited red light induced inactivation, but the levels of LacZ under dark conditions were approximately half. Basal expressions shifted slightly higher. For the ON state, the results are consistent with the hypothesis and model prediction that for a finite amount of OmpR-p in the cell, a large copy number of PompC promoter will introduce an OmpR-p sink. OmpR-p then becomes limited and unable to fully occupy and activate all of the promoters present, resulting to retroactive effects. However, other possible reasons cannot be ruled out:

- The large copy number is introducing a burden effect on the cells by taking up too much resources. Hence, even if OmpR-p does not become limited, the over-expression of eCFP, receptor and PCB enzymes along with the plasmid replication process, will withhold the LacZ production from reaching its potential levels.

- The modified RBS sequence downstream the J23109 - J23110 can be potentially under-performing regarding the translation rates for the receptor and enzymes.

- The promoters themselves were chosen to be relatively weak, in order to avoid the previous point made about burden. Nevertheless, the reduction in transcription rate can be potentially lower for Cph8 and enzymes, compared to the basal expression from PTetO1 and Plac/ara-1 used by the original device.

The fact that the basal expression is still equal or even higher than the original device, is not supportive to the basic hypothesis about OmpR-p depletion. If OmpR-p was limited then basal expression should drop as well. On the other hand, this is a complexly regulated system since OmpR is the native osmoregulator. Additional regulatory mechanisms, or leaky expression, independent to OmpR, could not be ruled out at this point.

The negative control sample which expresses only the PCB production genes, showed no dark induced activation, but instead remained at basal levels, slightly lower than the basal levels of the full constructs. These suggests that Cph8, under the light conditions used, exhibits some leakiness in phosphorylation of OmpR, since when not present, the basal level dropped. However, the non-zero value for this strain, indicates the extent of LacZ expression due to auxiliary OmpR regulation and non-OmpR dependent basal expression from PompC. The positive control (Cph8 without PCB enzymes), also showed no sensitivity to light due to the absence of the chromophore . Hence, Cph8 remained in an active conformational state significantly higher than basal levels. The fact that the levels of LacZ did not approach the levels of the dark-induced holophytochrome, indicates that PCB can be a stabilising co-factor for the HK domain when in Pr state , compared to the corresponding apophytochrome.
The transfer function of LacZ at steady state, for varying the light intensity (640 nm) shown in Figure 3.2.4 (b), showed a more linear transition rather than a step function (x axis is in a logarithmic scale). This was also consistent with published plate based assays (Levskaya et al, 2005; Tabor et al, 2011). Based on the Hill-type fitting curve, obtained for the dual plasmid system, the dynamic range for the system extended from \(~100\) μW/m² up to \(0.05\) mW/m² for basal to maximum activity level, respectively. In general, the error in the measurements that was observed, can be partially due to the light-distance calibration that was prepared for these measurements, along with all the previously stated reasons.

**Dynamic response**

Figure 3.2.5 shows the time course run for pL-Cph8FD according to the method discussed in section 2.1.2. The apparent time the system needs to turn from basal levels to steady state, was approximately 2 hours while for the opposite, it was estimated approximately at 2-2.5 hours. These values were defined as the rise and descent times, \(\tau_r\) and \(\tau_d\) respectively, which is the time needed for the system to rise or descent to 95% of its maximum or basal value. Switching the system ON, lacZ activity reached the expected values at 600 MU (\(~1200\) molecules per cell, calculation discussed below) but for switching OFF, the levels rested at a higher value than the basal, by approximately 150 MU or 300 molecules per cell. Apart from human error, other reasons that might have contributed to this are

- Additional, or more frequent dilutions might have been necessary, not just to keep the cells at exponential phase of growth, but keep the cell density low enough so that oxygen and osmolarity changes due to cell metabolism are kept low. These factors might have an effect on the auxiliary and OmpR-p independent basal expression. On the contrary, if the dilutions were at a high enough ratio, resulting to low OD₆₀₀, or underestimated by the plate reader used, this can skew the calculations of Miller Units upwards, since the formula for calculating MU normalises for OD₆₀₀.
- Sampling time intervals can potentially introduce error due to the fact that the cells are exposed to ambient light conditions, instead of red and dark, for \(~5\) minutes during sample collection or dilution. This time period, can be enough for the Pfr form of Cph8 to turn to its Pr form, increasing OmpR-p phosphorylation beyond basal levels. Although not conclusive, such rise can be seen during the four time points before the OFF to ON switching shown in in Figure 3.2.5(b).
- The switching off function, can be in fact more complicated and longer than what was observed during the time frame of the experiment. One hypothesis could be, that the additional PompC operators controlling eCFP and act as an OmpR-p sink, can be gradually releasing bound OmpR-p when its levels exhibit a sharp decrease. Hence, for a limited amount of time, OmpR-p in the cytoplasm is buffered and resupplied by the 'PompC-stored' molecules of the 200 extra copies of PompC. This can be introducing a 'resisting force' to the switching OFF the system, that can be more significant when using high copy numbers of PompC.
- Finally, if the value at the last time point \((t=8h)\) was overestimated for any reason related to the Miller assay, can by itself be the reason for this deviation. For clarity, it is specified that the model was fitted on these results, so the predicted time evolution shown in Figure 3.2.5, holds no information on how the real
Chapter 3. The Light Receptor Characterisation

decreasing function should look like.

Measurements of longer time frames are needed in order to exclude the last point, but the dilution effect and OmpR-p sink hypothesis is a more challenging task to pinpoint and reach a more conclusive result. The Miller assay, is by itself a labour demanding and time consuming process compared to a fluorescent reporter, making sample collection at a smaller time interval difficult, in terms of exposing the samples to uncontrolled light conditions (long enough to have a significant effect on the observed measurements). However, as discussed below Miller Units allowed the estimation of molecule numbers for fitting the model.

3.2.4. Model Fitting - Parameter Estimation

Given the set of data obtained with the specific experimental setup, an attempt to estimate the parameters of the model, that were not found in the literature and shown to be critical according to the sensitivity analysis shown in the modelling section of this chapter (section 3.1), was performed. The only available time trace about the system, was the final output in terms of LacZ activity but no intermediate processes, like phosphorylation of OmpR-p or PCB and Cph8 expression, could be monitored with the available methods and equipment. Hence, the fitted rate constants for these intermediate reactions, are not necessarily representative of the true values, but were shown sufficient for a semi-quantitative model that served the purposes of this work.

A first important task for the fitting was to define how to correlate experimental and modelling data. Miller Units is an index of the LacZ activity measured, in terms of the amount of ONPG cleaved to ONP, in a given time period. The conversion of this activity to an average number of LacZ, was shown possible by Garcia et al, (2011). The authors have proposed the formula shown below:

\[ k = 138 \cdot 10^6 \frac{M_{ONP}^M}{\text{min} \cdot \text{LacZ} \cdot M_{ONP}^M}, \gamma = 0.0045 \frac{\text{M}_{\text{ONP}}}{\mu \text{M}_{\text{ONP}}}, \delta = 8.9 \cdot 10^8 \frac{N_{\text{cells}}}{\text{nm} \cdot \text{ONP}} \]

\[ \Rightarrow \frac{1}{2} \frac{MU \cdot ml}{min} \approx \frac{Tetramers}{N_{\text{LacZ}}} \frac{1}{N_{\text{cells}}} \]

\[ \Rightarrow 2MU \cdot ml \cdot min \approx \frac{N_{\text{LacZ}}}{N_{\text{cells}}} \]
Figure 3.2.5. Time course run for pL-Cph8FD. (a) The cultures were initially grown in the dark for the first 4 hours and then red light (>1 W/m²) was switched ON. LacZ expression started high, but dropped after the switching point. The model showed a good fit with the experimental data, apart from the endpoint measurement where LacZ levels appeared to be higher by approximately 200 molecules, from what the modelled curved showed. (b) The cultures started under red light (>1 W/m²), and after four hours, were switched to the dark. LacZ levels started low but showed a clear increase, once the light was off. Again, the model was shown able to capture this behaviour (post-fitting) and be in good agreement with the experimental data (within the s.deviation bounds of the stochastic model), however, the rise in lacZ was steeper for the experimental system compared to what the numerical simulations suggested.
(3.41) was used, which returns a ratio of 1:2 Miller Units to LacZ monomers, because the dimerisation/tetramerisation reactions for LacZ, were not explicitly modelled (under the assumption that the dissociation constant for LacZ multimers was very small). Of course, this formula was a rough approximation, but it offered a direct correlation between LacZ mediated ONPG cleavage to LacZ Molecules and any quantitative relation, even not accurate, was preferred over scaling the model’s output, in order to obtain a relative correlation with the experimental results. The authors (Garcia et al, 2011) specified though, that an accurate computation would require a calibration curve. However, the methods and resources required for this, fell out of the context of this study.

For populating the model values, a group of fixed values was determined. This included all the initial conditions and inputs shown in Table 3.2.1. These, are the average gene copy numbers, the total Heme concentration and light intensity. The latter, was calculated according to the fitting curves shown in Figure 3.2.1 (distance from the light source, corrected for the appropriate percentage of intensity in the cases were a neutral light filter was used).

Heme levels were determined experimentally, using the method proposed by Van Den Berg et al (1988) which was a modification of the HemoQuant test (Schwartz et al, 1983), that uses hot oxalic acid in high concentrations, in order to remove the iron from the heme group. The resulting protoporphyrin ring, which can be excited at 400 nm and emits strongly at 600 nm is used as the output. In the modified version, glacial acetic acid was used in the presence of FeSO$_4$ and lower incubation temperature as described in materials and methods (section 2.1.2). The acetic acid method was found, for the purposes of this work, to be more suitable for analytical purposes, since it returned a wider linear part for the fitting curve compared to the oxalic acid one. However, the latter seemed to be much more sensitive since it returned a relatively strong signal from cultures as small as 1 ml at exponential phase. On the down side, the reaction was inefficient for high concentrations of purified Heme-Chloride that was used as standard, as fluorescent values for 1 mM of heme, were observed to be lower than the cell samples, which suggested a biologically impossible heme concentration per cell.

This could occur due to a number of reasons. The oxalic acid at 2 M, was not soluble at room temperature so it had to be added warm at the start of the experiment, in both the incubated and non-incubated (baseline) samples. This, can result in a disproportional background heme reduction between samples and standards, as the cells are not lysed at this stage. Oxalic acid might be inefficient, compared to acetic acid, for reducing other substances found in the cells, resulting to a different micro-environment (compared to the standards), which can yield stronger fluorescence for the cell samples. Or similarly, it might be inefficient to react with the standards solved in DMSO. It is worth noting that the heme in the standard samples, seemed to be crashing out of solution during the precipitation step of the oxalic acid, after the incubation step, unlike the cell samples where no visible pigments could be detected on the pellets after spinning. Hence, for the quantification of Heme in RU1012 strains, the acetic acid method (Van Den Berg et al, 1988), was used as shown in Figure 3.2.6, while the qualitative comparison with Ho1 expressing strains, was carried out with the
Figure 3.2.6. Determination of the range of Heme levels in the strain RU1012 based on acetic acid method. The values are based on the fluorescences at 600 nm. A series of dilutions of known amounts of purified heme-chloride (Sigma Aldrich) was used to obtain a curve of standard concentrations. The fitting curve was calculated in MATLAB’s curve fitting tool. The high concentration samples greater than 100 μM were excluded from the fitting since they fall outside the linear part of the curve. The RU1012 strain and standards were assayed as described in materials and methods section 2.1.2. The calculated value for the entire sample was approximately 2.8 (±1.01) μM. And the calculated value for the total number of heme molecules per cell is 7935 (± 2848 ) molecules.

Figure 3.2.7. Relative fluorescent levels of Oxalic acid treated sample suggest that the total heme found in strains expressing the PCB producing enzymes (Ho1 and PcyA) is significantly higher than those of the bare strain. Quantification with this method was not possible.
oxalic acid mediated reduction (Figure 3.2.7). From the results, the number of calculated molecules of heme per cell was 7935 ± 2848 - bounds determined on the standard deviation of the fluorescence). Previously published values, showed that E. coli under aerobic conditions can produce heme at a rate of 0.2 μmol/g (dry weight) per hour (Ishida & Hino, 1972 in Hart et al., 1994). Given the typical dilution rate of ~0.017 min⁻¹ of an exponentially growing cell, with 40 min doubling time, the calculated heme amount at steady state was 75571 molecules/cell. This, was 10-fold higher from what was measured in the acetic acid assay. However, the reference value used, was the maximum possible rate, under ideal conditions, of a different E. coli strain (K-12 W1895) in different medium.

Gene copy numbers were determined according to the published values in literature shown in Table 3.2.1

<table>
<thead>
<tr>
<th>Species</th>
<th>Molecules (Original/Reassembled)</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>gCph8</td>
<td>50 / 250 ColE1 origin / ColE1 origin (pL-Cph8FD)</td>
<td></td>
</tr>
<tr>
<td>gPCB</td>
<td>15 / 250 p15A (pl-PCB ) / ColE1 origin (pL-Cph8FD)</td>
<td></td>
</tr>
<tr>
<td>gOmpR</td>
<td>1 Genome integrated</td>
<td></td>
</tr>
<tr>
<td>Rii (PompC)</td>
<td>1 / 1 + 250 Genome integrated / ColE1 (pL-Cph8FD)</td>
<td></td>
</tr>
<tr>
<td>Heme</td>
<td>8000 Experimentally determined (Figure 3.2.6)</td>
<td></td>
</tr>
<tr>
<td>OmpR-p</td>
<td>N/A Model determined at time t = 1000 min</td>
<td></td>
</tr>
<tr>
<td>LacZ</td>
<td>N/A Model determined at time t = 1000 min</td>
<td></td>
</tr>
<tr>
<td>Light</td>
<td>Variable Based on section 3.1.2</td>
<td></td>
</tr>
</tbody>
</table>

Due to the large number of parameters (55) of this model, it was clear that the number of parameters that was allowed to be varied during the data fitting process, had to be reduced. Figure 3.2.8, shows diagrammatically the process followed, to filter which parameters could be assigned a fixed value and which ones needed to be estimated, using the reduced model and experimental time course data shown in Figure 3.2.5. The parameters that were reported in the literature and no calculations or assumptions were needed, were assigned a fixed value. From the remaining set, the ones that a parameter in the literature, that embeds the specific rate constant, was available (e.g. dissociation constants), were separated from the ones that no literature information was found. From the first group, the parameters that exhibited the highest sensitivity (Methods 2.2), were selected for the fitting group, while from the unknown group, all but the ones that their relative sensitivity was low, were placed in the fitting set. After this filtering process, four parameters were chosen to be fitted based on the data. The $T_v$, which is the lumped parameter that includes LacZ transcription, translation and mRNA degradation rate constants, $T_o$ that includes the corresponding constants for Cph8, $c_o$ which is the phosphorylation catalysis rate constant of OmpR, from the HK domain of Cph8 and finally $v$, which is the non OmpR-dependent leaky expression rate. The exception from this process, was the non dimensional parameter $z$ (photon adsorbing efficiency, in function $L$ in equation (3.31)), that was fitted separately on the data, of the transfer function shown in Figure 3.2.4 (b). This was done in order to approximate the correct saturating light intensity bounds, prior to the fitting of the time course experiment. These fits for both $z$ and the time course data fitted parameters are shown in Figure 3.2.9(a) and in Appendix A.
Figure 3.2.8. Diagrammatic overview of the parameter fitting group selection. Apart from the known parameters from the literature the estimated and unknown parameters were split according to their relative sensitivity values with respect to LacZ output shown in barchart (b), into the fixed and fitted parameter set. This process aimed to reduce as much as possible the number of parameters that are allowed to vary because of the limited experimental information available for most of them. It is obvious that, given a large parameter set, the model can reach multiple solutions for the values that satisfy the observed output dynamics.
Figure 3.2.9. (a) Data fits, estimated value boxplots and residuals plot, for expression rate parameters $T_r$, $T_{lz}$, basal LacZ transcription $v$ and phosphorylation catalysis $c$. (b) Model and experimental data comparison for LacZ steady state levels for fitted parameter values.
The parameter estimation based on the INLMFIT method (SimBiology implementation), estimated the minimum error for $T$, $T'$, $c$, and $v$ at 0.813, 204.10, 0.104 and 0.0255, respectively. Using these values, the model was able to describe the observed data obtained from the pl-Cph8FD plasmid, from which the fitting dataset was obtained. In fact, from the estimated lumped parameters, individual kinetic rate constants were able to be calculated, in order to run stochastic simulations of the full mass-action model, shown in Figure 3.2.5. The stochastically simulated traces, did not show any qualitatively different behaviour between the deterministic and stochastic regime. The variability observed, was in fact very close to the measurements's error.

In order to perform an additional evaluation, against a non training dataset, the steady state response data of the dual plasmid (pCph8 + pL-PCB) were tested against the fitted model. This system consists a fair perturbation test for the model, since it uses the standard version of CoIE1 replicon (50 copies) as opposed to >200 for the pUC that pL-Cph8FD is based on. More importantly, the biggest different is that this dual plasmid system, does not have the additional PomP:eCFP construct, which means the OmpR-p sink, assumed to be causing a retroactivity effect, is eliminated. Adjusting the initial conditions of the model for this system (Table 3.2.1), and running the simulations, showed some interesting results (Figure 3.2.9 (b)). The model succeeded on predicting the increase of output when the total OmpR-p binding sites are reduced, even if the Cph8 copy number was also reduced by 5-fold. On the down site, the model overestimated by ~800 MU this increase suggesting that further calibration of the parameter set is needed in order to obtain a predictive quantitative model. Furthermore, the model captured closely the active state, when PCB is absent from the system. The above although not a validation, are indicators that this model can be used as a framework for further work.

3.2.5. The Multichromatic System Characterisation - CcaS / Cph8

Similarly, the steady state response of the CcaS green light receptor was obtained. This is another two component system that shares the same network structure with Cph8. Hirose et al, (2008) proposed that CcaS receptor also binds covalently PCB on Cys-141 of the GAF domain through a thioether bond rendering it photosensitive to green 535 nm (Pg form) and red at 632 nm (Pr form). Unlike OmpR, the phosphorylation rate of CcaR is increased on the excited, high energy, state of CcaS when illuminated with green light and falls to basal levels when the receptor is left to return to ground state under no light or when illuminated with red light.

The CcaR system was tested using a combination of a triple plasmid system that incorporates both Cph8 and CcaS in a multichromatic control setup proposed by Tabor et al (2011). Figure 3.2.10 shows the plasmid maps of the pL-PCB(s), pJT122 that harbours cph8, ccaS along with the ccaR regulator, which is not native in E. coli and binds on the PcpcG2:LacZ casette. Finally the pJT106b harbours the PompC promoter, which expresses the CI repressor from lambda phage, which in turn regulates a second LacZ gene under the $\lambda P_L$ promoter. This will cause the inverse expression pattern for LacZ regulated by red light, so that when red light is ON CI levels fall to basal and release repression so that LacZ is ON.
Figure 3.2.10. (a) Plasmid maps for the multichromatic control system. pL-PCB was shown previously in Figure 3.2.4 (b) Steady state levels for green light induction of the multichromatic control system.
Another important aspect of the multichromatic system was that Cph8 exhibits a non negligible sensitivity to green light since the Cph8:Pr excitation spectrum (Figure 3.1.2) is maximum at approximately 630-640 nm but at 520 nm the excitation falls approximately by 80%. In addition the fact that the full proposed experimental system, comprising all three plasmids, is using LacZ for reporting both the Cph8 and CcaS activity, makes the measurement of each promoter’s contribution to the output more challenging. The same authors (Tabor et al, 2011) using the plasmid pJT118, that bears only CcaS, instead of pJT122 have shown that their output is not additive because of undetermined potential interactions, when Cph8 and CcaS are present in the system.

Combinations of the plasmids were assayed for their steady states which are shown in Figure 3.2.10. All three plasmids render a functional multichromatic system however in the absence of pJT106 the LacZ levels are only sensitive to green light. On the other hand, some of the observed fold-differences of the light sensitive strains, was higher than the published values. In addition when the Cph8 system was present, the observed LacZ output was increased by 2-fold, even if no red light was used. These results along with additional information regarding the red light induction of the multichromatic system found published in Tabor et al, (2011) are summarised in Table 3.2.2. Due to the limited literature on parameter values on the detailed mechanism of the system, but also the fact that this is an already characterised system within a synthetic biological context (Tabor et al, 2011), and because no modifications were carried out on it during this work, no further characterisation was performed.

<table>
<thead>
<tr>
<th>Light Condition</th>
<th>Cph8</th>
<th>CcaS</th>
<th>OmpR-p</th>
<th>CcaRp</th>
<th>Cl</th>
<th>LacZ&lt;sup&gt;+&lt;/sup&gt;</th>
<th>LacZ&lt;sup&gt;-&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>Pfr</td>
<td>Pg</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>-</td>
</tr>
<tr>
<td>Green</td>
<td>Pr &lt; Pfr*</td>
<td>Pr</td>
<td>-&lt;sup&gt;+&lt;/sup&gt;</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>Both</td>
<td>Pfr</td>
<td>Pr</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>Dark</td>
<td>Pr</td>
<td>Pg</td>
<td>++</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* A moderate sensitivity of Cph8 to green light was reported

Table 3.2.2. Qualitative activity for the multichromatic sensor for different light conditions

---

Imperial College London, September 2014

Marios Tomazou PhD Thesis 113
The next step, in realising large synthetic networks is to find an efficient way to couple different devices or modules, using the light input systems. The coupling mechanism is crucial, in terms of not altering the behaviour of a device as it was observed in isolation. It is very common for the behaviour of each device, part of a larger network, to deviate, mainly due to crosstalk with other devices within the same host and also availability of resources. In these chapter, I present the implementation of a proposed efficient coupling solution (at the transcriptional control level), of the light input modules as presented in Chapter 3, with other networks. In particular, I designed and build two hybrid promoters that can bind and be regulated by the LacI repressor and OmpR or CcaR, allowing the user to achieve control over the output by both the use of a classic chemical inducer, like IPTG and light.
Chapter 4

Light and Chemical Input Promoters

4.1. Background

Engineered genetic devices commonly rely on the use of regulated promoters, so they can emulate logic functions, like a logic gate in electronics can give an ON or OFF (binary) output, depending on one or multiple inputs. These promoters, can selectively bind transcriptional regulators that can either block, or increase the affinity for the RNA polymerase. However, it is understood that a strict digital behaviour is not really achievable in living organisms, due to the nature of gene expression (Elowitz et al. 2002) and the other confounding factors mentioned in Chapter 1.2 e.g. crosstalk. In fact, orthogonal regulators that can interact strictly and only with the intended promoters, is not the case with most commonly used systems (Voigt et al, 2006). Biological signalling pathways can be interpreted as a noisy analog signal, propagating through multiple interconnected and branched channels. Hence, coupling two synthetic circuits, sometimes need to be by the simplest and 'shortest' way, in terms of intermediated components.

In order to couple the characterised light input machinery, with other existing networks, a way to transduce the elevated OmpR-p levels 'signal', to the expression of another regulator, part of a downstream device is needed. This connection, must be short in terms of intermediate components and as orthogonal as possible, targeting the specific promoter of the specific downstream device that needs to be connected to.

By considering different ways that two networks can be connected and interact, three possibilities were shortlisted based on their complexity and limitations, in terms of the number of available regulators that can be used. The three identified ways that two synthetic devices can be connected, with a minimal use of additional genes are demonstrated in Figure 4.1.2. If device A, is one of the light Input devices and network B any other previously proposed network e.g. toggle switch or oscillator, these options are:

1) The output of A can be a regulator (e.g. LacI ), that is already part of the downstream network B

2) The output of A can be a TF that is not part of B, but can bind on one of the promoters found in the downstream network and co-regulate it with another factor which is part of B (Lutz & Bujard, 1997).

3) An engineered promoter that can directly bind OmpR-p and co-oregulate B with another TF.

From these options 1) was the one with the most limitations, since adding an additional copy of a TF that is already a part of the downstream system, will almost inevitably disrupt the balance of the network, by either changing its original characterised properties or even disrupt its intended function. The reasons, are implicit to the gene expression mechanism in bacteria. Basal expression levels, inducer transfer function shifts or even possible plasmid recombination for repeated DNA sequences, are some of the reasons to avoid this type of connection. With option 2), the crosstalk effect with the device of interest can be eliminated, but an
**Figure 4.1.1.** Repression mechanism of LacI. LacI tetramers bind strongly the lacO1 operators where depending on their position and occupancy state can repress by blocking mRNA elongation, RNA Polymerase or cyclic AMP / CRP binding by either blocking directly the binding sites or by DNA loop formation since a LacI tetramer can bind two operators through its anti-symmetrical binding sites.

**Figure 4.1.2.** Examples of Signal transduction strategies between individually characterised genetic networks. Option 3 provides the solution with the least cost in resources and response time. Details are discussed in the text.
additional gene of a different TF is still needed to be expressed. The response time will be dictated by how fast this intermediate TF can be expressed and regulate network B.

Based on the above, 3) can be the solution with the least cost on resources, limits potential crosstalks between A and B and yield the fastest possible response time, since OmpR-p or CcaR-p will regulate directly a promoter from network B. In addition, such promoters have been demonstrated previously to have a well defined combined regulation (e.g. Plac/ara-1) [Lutz & Bujard, 1997]. The downside, is that complex regulation can be unpredictable with respect to the effect of the two regulators on the output of the promoter. For example the binding of one of the two TF can affect the affinity or block entirely the binding of the other regulator.

Proceeding with the latter, requires apart from OmpR an additional coupling regulator that can be targeted on the same promoter. Considering the most well characterised, reliable and widely used transcriptional regulators that have appeared previously as parts in synthetic circuits over the past few years, it becomes clear, that the list is very short. This is limited to the point where most of the genetic devices build during the last decade, share the exact same components. Table 4.1.1 shows a summary for the most ’popular’ regulators, along with their native and examples of synthetic promoters and finally in which synthetic genetic devices have been used in.

LacI is the most widely used TF in other synthetic circuits. It is well characterised and can be induced by IPTG, a relatively economic galactose analogue, that the cell cannot metabolise, hence a stable concentration of it across the time frame of a typical time course experiment, can be assumed. In addition, the large availability of promoter variants able to bind this regulator, along with multiple types of operator sites and its versatile repression mechanism (Figure 4.1.1), allows for more options in the effectiveness of the repression.

**Table 4.1.1.** Most commonly used regulators in bacterial synthetic genetic circuits

<table>
<thead>
<tr>
<th>Regulator</th>
<th>Native Promoter</th>
<th>Synthetic Promoters</th>
<th>Used in</th>
</tr>
</thead>
<tbody>
<tr>
<td>LacI</td>
<td>Plac (lac operon)</td>
<td>Ptrl, PlacO-1, Ptac, Ptic, Plac/ara-1</td>
<td>DFO, RPRO, TS, BPF, PPF</td>
</tr>
<tr>
<td>TetR</td>
<td>Ptet</td>
<td>PLtetO-1, PTTetR/P22 Mnt</td>
<td>RPRO, TS, PPF</td>
</tr>
<tr>
<td>CI</td>
<td>λPL, λPR</td>
<td>Plac/CI</td>
<td>RPRO, TS, PPF</td>
</tr>
<tr>
<td>AraC</td>
<td>ParaB (araBAD operon)</td>
<td>Plac/ara-1</td>
<td>DFO</td>
</tr>
<tr>
<td>LuxR</td>
<td>Pluxl</td>
<td>Plac/lux</td>
<td>QSO, PPF</td>
</tr>
</tbody>
</table>

4.2. The Light-Chemical induced promoter design and model

4.2.1. Design

Promoters that are controlled by more than one kind of regulator are very common in nature, in fact, more common than a single input, orthogonally regulated promoter. Examples like the native promoters found upstream the lac, araBAD, xylAB, galE and deoCABD operons (Ecocyc; Keseler et al, 2009) are all subject to negative regulation by their corresponding repressor (LacI, AraC, XylR, GalR, DeoR, respectively) but also global regulators like cAMP/CRP or additional regulators (e.g. AraC represses xylAB or CytR represses deoCABD). The above, suggests that native pathways are coupled in a significant degree and become robust, through this complex regulatory scheme at the expense of 'orthogonal' regulation. At the same time as discussed above, this type of coupling is expected to be the most cost effective, in terms of resources, but also faster than using additional TFs. Therefore, it is reasonable to build such promoters and apply this type of coupling on synthetic genetic devices.

A well known and used synthetic double input promoter in E. coli is the Plac/ara-1 (Lutz & Bujard 1997). This promoter comprises an O1 site downstream the transcription initiation site, a synthetic Os operator between the -10 and -35 box and another O1 site at -448. All these three sites, can bind LacI and tightly repress the expression from it, by either blocking the RNA polymerase open complex formation, mRNA elongation or by forming a loop between one of the two downstream operators and the upstream O1 site. In addition, the CRP/cAMP binding sites have been deleted and replaced by the I1 and I2 Arac binding sites, from the PBad promoter (Zhang et al, 1996), found in the araBAD operon. In the presence of arabinose, AraC has increased affinity to these sites enhancing the RNA polymerase binding on the promoter. Therefore, the promoter exhibits maximum activity, when both arabinose and IPTG are present, while the transcription rate falls to basal when both inducers are absent.

Based on the same architecture as Plac/ara-1, a design where the AraC activating component is replaced by the light inducible regulators OmpR-p or CcaR-p, is proposed (Figure 4.2.1.). For the Plac/ompR promoter LacI binding operators can be placed up and downstream the PompC promoter keeping its -10, -35 sequences intact, along with the transcription initiation site. The rational behind the latter, is that there is not sufficient information in the literature on the exact interactions between the OmpR-p bound on the DNA and the RNA polymerase, so by replacing the -10 and -35 sequences with those from the Plac/ara-1, there is a risk of disabling completely the positive regulatory activity from OmpR-p.

On the other hand, the literature around the Cpc-G2 promoter is very limited and the exact binding sequences for CcaRp are not known. To overcome this problem, the entire so called 'energetic region' of PpcpcG2, that includes the estimated region of the CcaR operators the -10, -35 and transcription start site, was placed upstream the -35 box of Plac/ara-1 replacing the AraC binding sites. This suggests that transcription can potentially start from both sites, but overall will result in a positive regulation by CcaR-p and negative by LacI.
Chapter 4. The Light Receptor Characterisation

Figure 4.2.1. Chemical and light inducible promoter design. Starting from an existing dual regulated synthetic promoter (Plac/ara-1), a similar architecture was used for building the red light and IPTG inducible promoter Plac/ompR and Plac/ccaR. Red dashed lines show the sequence fragments that were used in Plac/ompR and green dashed lines the corresponding sequence used for Plac/CcaR.
4.2.2. The Compact TCS Model and Extension for LacI Binding

Modelling the double input promoter activity requires the integration of the light receptor model, derived in Chapter 3, with the additional regulatory function of LacI. However, the multiple operator sites for OmpR if modelled explicitly (Chapter 3, section 3.1) result in a large and complex model. Any attempt to include multiple operator sites, for LacI, is expected to be a significantly more challenging task. For simplification reasons, an abstract or 'compact' model for the Cph8/OmpR and CcaS/CcaR systems, that capture the basic behaviour in terms of expression levels and dynamics, was derived in order to integrate both regulators.

The 'compact' model is an abstract representation of the mechanistic model shown in Chapter 3. It retains the receptor and regulator expression under a single reaction, their complex formation and phosphorylation, the light mediated state transition, and finally the auxiliary phosphorylation.

The kinetic scheme used has as follows:

\[ \begin{align*}
\text{Cph8/OmpR regulation} & : \\
\emptyset & \xrightarrow{a} \text{OmpR} \\
\emptyset & \xrightarrow{a} \text{Cph8}^{pr} \\
\text{Cph8}^{pr} + \text{OmpR} & \xrightarrow{c} \text{Cph8}^{pr} + \text{OmpR} \\
\text{Cph8}^{pr} & \xrightarrow{f_{(LR)}} \text{Cph8}^{pr} \\
\text{OmpR} & \xrightarrow{a_1} \text{OmpRp} \\
\text{OmpRp} & \xrightarrow{d} \emptyset \\
\text{Cph8}^{pr} & \xrightarrow{d} \emptyset
\end{align*} \]

\[ \begin{align*}
\text{CcaR/CcaS regulation} & : \\
\emptyset & \xrightarrow{a} \text{CcaR} \\
\emptyset & \xrightarrow{a} \text{CcaS}^{pg} \\
\text{CcaS}^{pr} + \text{CcaR} & \xrightarrow{c} \text{CcaS}^{pr} + \text{CcaRp} \\
\text{CcaS}^{pg} & \xrightarrow{f_{(LR)}} \text{CcaS}^{pr} \\
\text{CcaRp} & \xrightarrow{d} \emptyset \\
\text{CcaS}^{pr} & \xrightarrow{d} \emptyset
\end{align*} \]

For fitting the basic parameters for the light-regulated-only parts of the model, the same reporter system as in Chapter 3 was used, but collapsing the explicit promoter-regulator binding reactions into a single step:

\[ \begin{align*}
\text{PompC regulation} & : \\
R_{0,0} + 6 \text{ OmpRp} & \xrightarrow{p} P_2 \\
R_{3,3} & \xrightarrow{m} R_{3,3} + mLacZ \\
R_{0,0} & \xrightarrow{b} R_{0,0} + mLacZ \\
R_{3,3} & \xrightarrow{d_3} R_{0,0}
\end{align*} \]

\[ \begin{align*}
\text{Pcpc-G2 regulation} & : \\
P_0 + 2 \text{ CcaRp} & \xrightarrow{p} P_2 \\
P_2 & \xrightarrow{r} P_2 + mLacZ \\
P_0 & \xrightarrow{b} P_0 + mLacZ \\
P_2 & \xrightarrow{d} P_0
\end{align*} \]

\[ \begin{align*}
\text{LacZ expression} & : \\
mLacZ & \xrightarrow{d} mLacZ + LacZ \\
LacZ & \xrightarrow{d} \emptyset \\
mLacZ & \xrightarrow{d} \emptyset
\end{align*} \]
Although this model was not as detailed, it was shown to be sufficient to reproduce the basic behaviour of the fitted, reduced model shown in chapter 3. The minimum required properties that were needed to be reproducible by the compact model are:

- The steady state levels for the active promoter and basal expression
- The transfer function of light to the promoter activity
- ON/OFF or OFF/ON transition dynamics deviation limit to be less than 30 minutes

Among others, the main characteristics of the system that cannot be accounted for, using this compact version are:

- PCB levels are not modelled
- Activity from the apo-phytochrome (receptors without PCB bound on them) is not taken in account
- Variable gene copy numbers for receptors and regulators are now included into their corresponding combined expression reaction
- Hierarchical binding on the PompC is replaced with a single OmpR-PompC binding reaction

Modelling explicitly transcription, translation and degradation for the receptors (Cph8/CcaS) and regulators (OmpR/CcaR) was neglected, based on the fact that they are constitutively expressed. Instead, a constant amount that is distributed between the active/phosphorylated and inactive/de-phosphorylated state for these quantities, was considered. PCB production and binding, was also not modelled. Hence, the total amount of receptor in the system was considered to be a functional phytochrome. Any deviation, due to apo-phytochrome activity was now embedded into the auxiliary phosphorylation rate constants $a_1/a_g$. Phosphorylation reactions, were modelled as a single step transition from de-phosphorylated to phosphorylated state, without modelling the complex formation intermediate. Finally, the promoter binding mechanism shown in chapter 3 is now replaced with a single reaction that binds six OmpR-p or two CcaR-p molecules on the POMPc and Ppc-G2 promoters respectively.

The above simplifications were necessary, in order to keep the combined LacI and light regulated promoter model, at a reasonable degree of complexity while keeping the basic functionality of light induction. Naturally, this model will deviate in terms of dynamics and expression levels since not all reactions are taken in account. A partial error compensation, can be reached by refitting the lumped parameters ($B/a/c_o$ for OmpR, and $P/a/g/c_p$ for CcaR) on the available data. For all the other parameters (e.g. $k_l$, $k_f$) that appear in reactions that are not modified, were used with their corresponding values from the original model fitted in Chapter 3.

Figure 4.2.2 (a) shows numerical simulations of the compact and original Cph8 model after fitting it with the experimental data. The two models, reproduce approximately the same steady states, although some deviation in the dynamics is observed. This fitting was performed purely to confirm that the compact model structure can describe the TCS system behaviour and obtain the light receptor related lumped parameter values to be used as priors, for further fitting of the full double input promoter system.
Figure 4.2.2. (a) Re-fitted compact model on the time course data presented in Chapter 3 (b)
Model derived dynamic range of the promoter activity for varying amounts of light and IPTG.
Once a simpler model for an OmpR/CcaR activated system was available, modelling the LacI regulatory effect was allowed.

Double input promoters like the Plac/ara-1, have been previously modelled (Stricker et al. 2008) for the dual feedback oscillator (discussed in chapter 6). Following the same principle, the AraC binding with OmpR/CcaR can be replaced with the reactions associated with the PompC and Pcpc-g2 promoters (labelled R and P, respectively), within the following kinetic scheme, that describes the dual input promoters.

For OmpR

\[
P_{0,j} + 6 \text{OmpRp} \xrightarrow{\frac{b}{b_{-1}}} P_{1,j}
\]

\[
P_{i,0} + \text{LacI}_4 \xrightarrow{\frac{2k_r}{k_{-1}}} P_{i,1}
\]

\[
P_{i,1} + \text{LacI}_4 \xrightarrow{\frac{k_r}{2k_{-1}}} P_{i,2}
\]

\[
P_{i,1} \xrightarrow{k_r} P_{L,1} + i \cdot 6 \text{OmpRp}
\]

\[
P_{i,2} \xrightarrow{k_r} P_{L,1} + \text{LacI}_4 + i \cdot 6 \text{OmpI}
\]

\[
P_{L,0} \xrightarrow{k_r} P_{0,0}
\]

\[j \in \{0,1,2\}, i \in \{0,1\}\]

\[
k_r = k_{r,0} \left( c_r^{\text{max}} - c_r^{\text{min}} \right) \frac{1}{1 + \left( \frac{[\text{IPTG}]}{k_{r,1}} \right)}
\]

LacI expression and degradation

\[
\varnothing \xrightarrow{\frac{e_v}{d_v}} \text{LacI}_4
\]

\[
\text{GFP} \xrightarrow{d_v} \varnothing
\]

\[
m_{\text{gfp}} \xrightarrow{d_v} \varnothing
\]

\[
P_{0,j} \xrightarrow{d_v} P_{0,j}
\]

\[
P_{i,0} \xrightarrow{d_v} P_{i,0}
\]

\[
P_{i,1} \xrightarrow{d_v} P_{i,1}
\]

\[
P_{i,2} \xrightarrow{d_v} P_{i,2}
\]

\[
P_{L,0} \xrightarrow{d_v} P_{L,0}
\]

Transcription/Translation

\[
P_{0,0} \xrightarrow{\beta_{O}} P_{0,0} + m_{\text{gfp}}
\]

\[
P_{1,0} \xrightarrow{\beta_{gfp}} P_{1,0} + m_{\text{gfp}}
\]

\[
m_{\text{gfp}} \xrightarrow{r_{e}} m_{\text{gfp}} + \text{GFP}
\]

Each promoter has a basal expression rate when no regulator is bound on them. Once OmpR-p or CcaR-p binds, the promoter activity is maximised. However, if a LacI tetramer is bound, transcription is turned OFF. For one or two bound LacI tetrameres, the promoter can form a loop (\(P_i\) form). However, the loop formation mechanism was modified from the model proposed by Stricker et al. 2008 (Details on the modifications are discussed in chapter 6 section 6.2). Looped promoters, are not active and are not allowed to bind other regulators. The rate \(k_r\) with which LacI binds, was determined by a hill type function used in Stricker et al. (2008). Figure 4.2.2 (b), shows deterministic simulations, when the double input promoter reactions replace the PompC and Pcpc-G2 components of the model, prior to fitting experimental data. The box plots, cover the full range of light input against a varying IPTG concentration.
4.3. Experimental Implementation

4.3.1. Dual Input Promoter Experimental Assembly

For the Plac/ompR design, as shown in Figure 4.3.1, the downstream region of the promoter was synthesised as a single 177 bp fragment in order to avoid the required multiple steps of cloning, for inserting correctly the LacI operators in between the -10 and -35 box and downstream the transcription start site. In addition, the multiple OmpR binding operators, would probably have complicated the custom assembly of the specific fragment, from smaller oligonucleotides. The sequence, was restriction digested with EcoRI/SpeI and ligated upstream an EcoRI/XbaI digested RBS:GFPmut3b:terminator construct, obtained from the Registry of Biological Parts. The -520 up to -200 fragment of the promoter, was amplified by PCR from the pL-PCB (A) plasmid, harbouring the Plac/ara-1 promoter and cloned by inFusion®, upstream of the synthesised part. Finally, the complete Plac/ompR:GFP construct was amplified by PCR, with the reverse oligo including in its overhang the constitutive promoter J23110. This, was cloned upstream of the RBS of the PCB production cassette (Ho1 and PcyA) in pL-PCB as shown in Figure 4.3.1. This yielded the plasmid pL-PloPCB.

For the Plac/ccaR promoter, the energetic region of Ppcp-G2, was amplified by PCR using the plasmid pJT118 as template and cloned by In-Fusion® into the Plac/ara-1 promoter in pL-PCB, replacing the araI1 and araI2 sites. Next, the RBS:GFP:terminator construct, was amplified by PCR and cloned by inFusion® in between the Plac/ccaR and PCB cassette, introducing the J23106 promoter on the downstream part. The final resultant plasmid, is pL-PlcPCB. The cloning process and corresponding plasmid maps are shown in Figure 4.3.2.

The cloning strategy for both promoters, resulted to a plasmid that expresses constitutively PCB and GFP regulated by LacI and OmpR or CcaR, accordingly. The constitutive promoter for the PCB cassette used, was the same as the one used into pL-Cph8FD, J23110.

Next, the pL-PloPCB and pL-PlcPCB were co-transformed with pCph8 and pJT118 respectively, into a JT2 strain, so that each is expressed with the receptor that phosphorylates the corresponding regulators. Alternatively, both plasmids can be transformed with the pJT122 plasmid that harbours both light receptors.
Figure 4.3.1. Construction of the pL-PloPCB plasmid which features the Plac/ompR promoter upstream of a GFPmut3b gene. In addition the PCB production genes (Ho1 and PcyA) are now being expressed from the constitutive promoter J23110 (Anderson promoter library in the Registry of Biological parts). Initially the downstream 177 bp of the Plac/ompR that were chemically synthesised were cloned by restriction digestion and ligation upstream of an RBS:GFPmut3b:Double Terminator construct. The upstream part of the promoter that is identical to the Plac/ara-1 was PCR amplified from the pL-PCB plasmid cloned by In-Fusion® (red arrow) with the construct obtained after the first step. The full Plac/ompR-GFP construct was PCR amplified and cloned by In-Fusion® upstream of the PCB expressing cassette on pL-PCB, replacing at the same time the Plac/ara-1 promoter with J23110.
Figure 4.3.2. Construction of the pL-PicPCB plasmid harbouring the Plac/ccaR promoter upstream of a GFPmut3b gene. Similarly with PloPCB the PCB production cassette constitutively expressed. The ‘energetic region’ upstream the cpc-G2 gene (Pcpc-G2) was PCR amplified from the pJT118 plasmid and cloned by In-Fusion® in the Plac/ara-1 promoter in the pL-PCB plasmid. The position of the cloning was such that the Pcpc-G2 fragment replaces the araI1-2 sites on the Plac/ara-1. Next the resulting Plac/ccaR promoter was cloned by In-Fusion® into the pL-PloPCB plasmid in order to replace the Plac/ompR promoter. Finally the plasmid was co-transformed with pJT118 into a JT2 strain.
4.3.2. Experimental Characterization

The two promoters, both cloned into a constitutively expressing PCB plasmid are co-transformed with their corresponding light receptors (Cph8 for Plac/ompR and CcaS for Plac/ccaR) into a JT2 strain (ΔenvZ). They were assayed for the steady state and dynamic behaviour with respect to the effect of the two inducers, light and IPTG.

Similarly to the experimental methods followed in Chapter 3, the time resolution used was limited to 30 min intervals between measurements and in some cases 1 hour. This is due to the fact that no suitable equipment was available that can automate the light induction and take OD and fluorescence measurements at the same time.

More specifically individual colonies of freshly transformed cells were picked from plates and grown overnight at 37 °C with appropriate antibiotics. The following day, the samples were diluted 1:100 and left to grow for two hours. After, an additional 1:10 dilution, into fresh media with antibiotics and IPTG where appropriate, the samples were split in duplicates, in either 14 ml tubes or microplates for light/dark induction.

Steady State Performance

Initially, for characterising the promoters, a preliminary assay was performed in order to decipher whether the Plac/OmpR promoter was responsive to IPTG. The pL-PloPCB plasmid was transformed into a DH5α strain, without the pCph8. The samples were grown in a 96-well plate format for four hours in the dark, for different concentrations of IPTG ranging from 0 to 100 mM and OD\textsubscript{600} / fluorescence measurements were determined using the plate reader (Synergy). The barchart, in Figure 4.3.3, shows that GFP levels rose linearly from 70 AU up to 250 for 0 to 1 mM IPTG. For higher concentrations of IPTG the fluorescence remained at the same levels, suggesting that 1 mM is the saturation point for IPTG induction for the particular strain. More importantly, the promoter is shown to be responsive to IPTG even if LacI is not overexpressed, since the repression in this case was only due to the amount of LacI expressed, from the native lac operon gene.

For the Plac/ccaR case, the corresponding preliminary assay in Dh5α, showed no IPTG sensitivity, and as opposed to Plac/ompR, the overall expression rates stayed approximately 10-fold lower across all IPTG concentrations. This most likely occurred, because of the deletion of the AraC I1/I2 binding site upstream the -35 site of the Plac/ara-1 promoter. Even without LacI repressing the promoter, the affinity of the RNA polymerase to the specific promoter can be significantly lower, since Plac/ara-1 also lacks the native cAMP/CRP site (Lutz & Bujard, 1997). The only active polymerase binding site, was now the -10 -35 box of the Ppcp-G2 promoter and can be active, only when phosphorylated CcaR-p is present. In the case of Plac/ompR, the promoter was active since the strain that the specific system was assayed, had the native EnvZ receptor present, which under high osmolarity conditions can phosphorylate the native OmpR. The low expression can also be an indication that the Plac/ccaR promoter, is significantly weaker than Plac/ompR, something observed in the case of the native light-only promoters, Ppcp-G2 and PompC.
Figure 4.3.3. Plac/ompR IPTG induction in DH5α (a) and steady State performance after 3 hours for varying IPTG concentration and light conditions for (b) Plac/ompR and (c) Plac/ccaR. For each promoter the histograms obtained from the flow cytometer is given along with the corresponding bar chart of the exported statistics.
In order to demonstrate the combined effect of chemical and light induction, both the double input promoters and light receptor harbouring plasmids, were transformed in JT2. The samples were grown in three different concentrations of IPTG (0, 1, 10 mM), under dark or light (red for Plac/ompR and green for Plac/ccaR). Figure 4.3.3 shows the fluorescence levels after four hours of growth. It is clear, that in both cases the effect of light had a significant contribution. In the Plac/OmpR case, the fluorescence was 10-fold higher for samples kept in the dark, compared with the red illuminated ones. However, using IPTG in excess, seemed to increase the activity of the promoter, even if it was kept under red light, to reach levels that are only 2-fold lower than the dark counterpart.

In the Plac/CcaS case, a similar trend was observed with the effect of green light, which caused a 5 to 6-fold increase of fluorescence, for either concentration of IPTG. However, excessive amounts of IPTG did not seem to have a strong effect, such that they can 'bypass' the effect of light, as shown from samples that were kept in the dark and retain their fluorescence to values lower than 2 AU. This is consistent with the DH5α assay where IPTG only, failed to activate the promoter.

Finally, a 96-well plate assay aimed at characterising the IPTG/light input, at a higher resolution of concentrations and intensity. The surface plots in Figure 4.3.4, show these experimental results, along with the corresponding simulations, after fitting the model to the IPTG or light response curves (section 4.3.3). For the Plac/ompR case, the fluorescence decreased as a function of light and increased as a function of IPTG, while the maximum levels were observed in the absence of light and 1 mM IPTG. Although this result agrees with the preliminary assay discussed above, the model prediction, that in the absence of IPTG the output should remain at low to basal levels, was not observed. The latter suggests insufficient repression by LacI or very strong activating effect from OmpR-p.

The Plac/CcaR results, showed better agreement with the model, exhibiting low expression levels for low IPTG concentration regardless of the light intensity. However, the fact the activity rose significantly when IPTG was varied in the absence of light, was another contradicting result to what was observed in the 14 ml tube based assays. Apart from other physiological differences, when cells are grown in tubes or microplates, the most likely cause, was identified to be the difference in the gain used on the FACS machine in order to obtain the data. Due to the fact, that the preliminary assays showed very low levels of expression, from Plac/CcaR the gain was increased, so that better resolution was obtained. With this, it seems that IPTG alone is sufficient to activate the promoter, to some extent, and reach a third of its maximum potential. This effect was probably masked previously by the low overall expression levels. Even if the AraC I1/I2 binding sites are absent, the RNA polymerase is still weakly recruited, either on the downstream part of the Plac/ara-1 or as a result of the leakiness of the PccG-2 promoter, found further upstream.
Dynamic Performance

For determining the dynamic performance for the two double input promoters, the same protocol described in the steady state performance assays was followed, although this time, samples were collected at regular time intervals while the induction point was set at $t=3\, \text{h}$ for Plac/OmpR and $t=4\, \text{h}$ for Plac/ccaR. In addition, the samples were diluted at $t=2, 4, 5, 6\, \text{h}$. In both cases the samples were incubated with 1 mM IPTG. 200 μL of samples were collected at regular time intervals, and measured for green fluorescence using the FACScan machine (BD).

Figure 4.3.5 shows the results for the particular assay. In (a) the Plac/ompR for the first half of the experiment, grown in the dark, showed as expected an approximately four-fold higher fluorescence, compared with the samples starting under red light. After the switching point ($t=3\, \text{h}$), the fluorescence rose for approximately 2 hours and settled at 250 AU after another 2 hours, while the levels of the samples started in the dark, showed an almost linear decay of their fluorescence. The duration of the experiment, was not enough to reach the point where the fluorescence, of the switching OFF sample, settles to the initial basal levels, of the sample that started OFF.

Similarly in the Plac/ccaR case (b) a significant separation (7-fold higher) for samples starting under green light and samples grown in the dark, was observed. After the switching point $t= 4\, \text{h}$, the GFP levels for the samples started in the dark, rapidly increased and reached maximum expression levels, in just over an hour. For the ON/OFF case, the cells were again unable to reach the low basal fluorescence levels within the duration of the experiment.

The slow degradation compared to the system expressing the single light input promoters (Chapter 3), can be attributed to a number of reasons. The most likely cause, is that the reporter steady state levels in the case of the double input promoters, that are expressed from a p15A plasmid is higher, than that of the LacZ expressed from a single genomic copy of PompC::LacZ. This suggests that under same degradation/dilution rates, longer time is needed for the system that started from a higher concentration, to drop down to the basal levels. In addition, GFP variants are known to be very stable proteins in bacteria. Their half life, of over 24 h (Andersen et al, 1998) and mRNA half life estimated at 6 min (Megerle et al, 2008), is larger compared to the LacZ mRNA half life at 3 min (Dong & Kurland, 1995), of the larger LacZ reporter which was used to assess the light receptor systems in Chapter 3. Thus, at least the GFP mRNA (if not the protein itself) is removed from the system slower than LacZ.
Figure 4.3.4. Model and experimental data comparison for the GFP output of the system for varying amounts of light and IPTG. While Plac/ompR (a) shows a linear increase for both IPTG and decreasing amounts of light, Plac/ccaR (b) seems to be half activated, solely by the use of IPTG. Light, in the presence of IPTG, yields the maximum expression levels while in the absence, the activity falls to basal levels. The model, semi-quantitative, due to uncertainty in the parameter values, showed better qualitative agreement for Plac/ccaR rather than Plac/ompR.
Figure 4.3.5. Experimental Time course runs for Plac/ompR (a) and Plac/ccaR (b) for switching light conditions, in the presence of IPTG (1 mM). In both cases similar dynamics were observed with that shown when only the natural PompC and Pcpc-G2 were used (Chapter 3). The only main difference observed was in the ON/OFF scenario. While the system started high, after 3-4 hours of the switching point, the GFP levels dropped almost linearly but did not reach a steady basal level. In the OFF/ON experiment in both cases the GFP levels are rising, but the Plac/ccaR seems to be approaching the maximum expression level significantly faster (just over 60 min), compared to Plac/ompR which needs almost 2 hours.
4.3.3. Model Refinement and Data Fitting

Finally using part of the experimental data obtained the model was fitted as to represent more closely the observed characteristics. To enable this, the scaling parameter \( d \) was introduced in order to map the predicted molecule numbers of the model, to the fluorescence arbitrary units. This scaling is only valid under the assumption that fluorescence and GFP molecule abundance are linearly correlated, within the range of the experimental assay. In addition \( n \) factor was used to normalise the different gains used on the FACScan machine for the Plac/ompR and Plac/ccaR data. While the latter was calculated directly from the gain value at \( n = 0.75 \), \( d \) was treated as a fitting parameter within the simbiology non-linear regression routine.

The fitting dataset was the timecourse data shown in Figure 4.3.5. The fitting parameters, were the transcription rates for basal and activated expression (\( ba, baa \) - \( bg, bag \) for Plac/ompR and Plac/ccaR respectively) and \( d \). The latter was estimated at 0.34 and this value was used for both promoters. Figure 4.3.6 (a) shows the fitness of the model to the time course data.

The steady state values shown in Figure 4.3.4, were not suitable to be used for fitting, since they generated extremely large error for the estimated values. In addition, when the model was fitted with respect to the steady state values, it exhibited higher error when projected against the dynamic data. This, can be due to the relatively strong leakiness, observed at low levels of induction in the particular static characterisation assay, especially for the Plac/ompR promoter. Hence a larger training set is needed for further fitting of this model. Yet, for the further semi-quantitative modelling work of this project, the model was used based on the fact that the observed dynamics could be described sufficiently.
Figure 4.3.6. Fitted model on the time course data for both promoters. The fluorescence values for Plac/ccaR are normalised by the gain so that they are comparable with the Plac/ompR data, while the model output is scaled in order to assume mapping between the fluorescence and the predicted molecule numbers.
In this chapter, I use the Plac/ompR promoter, to achieve control over the genetic toggle switch (Gardner et al, 2000). Starting with the design and followed by a more detailed ODE model, I show how the building of a prototype behaved unpredictably, while the re-design and implementation of the system recovered the expected behaviour. Although more work needs to be done, in order to characterise the proposed circuit, the data have indicated that the specific light tuneable toggle switch, is bistable in the dark, able to switch between states by the use of IPTG. When light was used the system switched to the low state.
Chapter 5

A Light Tuneable Toggle Switch

5.1. Background

The synthetic genetic toggle switch (TS) was one of the first and most extensively discussed synthetic biological devices, proposed by Jim Collin's group approximately 14 years ago (Gardner et al. 2000). It is a device comprised of two mutually repressing regulators able to exhibit bistable behaviour, reaching two possible steady states depending on the initial condition of the system. One stable steady state has one of the two repressors at high concentration and the other one at low, while the second stable steady state is exactly the opposite. It is predicted mathematically that a third unstable steady state exists when the effective concentration of both receptors is perfectly equal; however, the system is quickly driven away from this state once a slight perturbation tips the balance of the repression to one side or the other. The circuit's diagram is shown in Figure 5.1.1. (a).

Other versions of the toggle switch have been proposed over time, like the one from Egbert & Klavins, (2012), which is different by including an additional mCherry reporter for the LacI side of the network, in addition to SsrA degradation tags so that the system has a faster turnaround time. Kramer & Fussenengger (2005), developed a mammalian toggle-switch, based on the same principle of two mutually repressing components. Genetic toggle switches based on alternative mechanisms have also been proposed. Ham et al. (2006 & 2008) demonstrated the use of an invertase in order to flip the direction from which a promoter initiates transcription. Dueber et al. (2003) showed how allosteric regulation on proteins can result in a genetic switch behaviour, while Atkinson et al. (2003) built a network that can both exhibit toggle switch or oscillatory behaviour.

A phenomenological model (Figure 5.1.1. (b)) describing the principal properties of the TS was proposed by its authors, Gardner et al. (2000). Plotting the phase plane shown in Figure 5.1.1 (c), and bifurcation analysis 5.1.1 (d) using this two dimensional model showed, respectively, the possible steady states along with the trajectories of the two repressors, but also the limit points when the expression rate parameter of one of the two repressors is varied. It was also shown how the system can be driven from a bistable to a monostable regime when the expression rate of one side is significantly stronger. Other similar phenomenological model versions like the one used in the next section (5.2.1), that include the effect of inducers, show similar shifts on the nullclines (A specie's function for another variable at equilibrium), by pushing them to a point where only one possible stable steady-state is allowed.
The genetic toggle switch model by Garder et al. (2000).

(a) Genetic toggle switch diagram. (b) Toggle switch model comprised of two time variables LacI and TetR and four parameters, $a_1, a_2$ as production rates and cooperativity degree $\beta$ and $\gamma$ for LacI and TetR respectively. (c) Phase plane analysis. For a bistable toggle switch $a_1 = a_2 = 9$ and $\beta = \gamma = 2$ the nullclines intersect at three points that correspond to the three possible equilibrium solutions. The monostable diagram occurs when one of the two repressors is over-expressed compared to the other (in this case $a_2 = 4$, $a_1 = 9$) and the system can have only one possible steady state solution. (d) The latter is shown also in the bifurcation diagram for $a_2$ when its varied between 0 to 30. The region that a bistable switch can exist is bound by the two limit points. These limits points when extended to an additional dimension, in this case $a_2$, as shown in the graph to the right reveal that the bistable region collapses at the cusp bifurcation point for values lower than 1. As the cooperativity degree decreases, the bistable region is shrinks.
Chapter 5. A Light Tuneable Toggle Switch

Since the TS is one of the most simple and well studied synthetic circuits with a number of potential applications, it is a fairly good target for modification in order to achieve control over it through the TCS light receptors. The objective was to have an additional input through light, apart from the IPTG and aTc, which are the 'default' external stimuli that can be used to flip the toggle switch to either side. As the system uses LacI that represses the Ptrc-2 promoter driving the expression of TetR, makes the double input promoters discussed in chapter 4, compatible to be used with this device. The transcription rate of TetR can then be controlled with the use of IPTG or light. This can potentially eliminate the need for inducers for pushing the system between the stable equilibrium points or vary the amount needed for this transition to happen.

Such a device can show interesting properties when the device is used in conjunction with larger metabolic pathways at an industrial level. This can replace the need to use large amounts of inducers, which is a far more costly solution compared to light, for switching between states. It was also previously described in yeast how the toggle switch can be turned into a timer or time-switches (Ellis et al, 2009) for flocculation. This can be an interesting alternative device where the amount of light alters the timing of the transition.

5.2. Light Tuneable Toggle Switch Model

5.2.1. Two-State Phenomenological Model of the Modified Light Toggle Switch (LTS)

The theoretical work at the design stage, started by introducing simple modifications to the published TS model shown in Figure 5.2.1 (b). The objective of this task was to determine how the phase planes and bifurcation diagrams are affected once the expression rate for the TetR sde is controlled by the double input promoter discussed in Chapter 4. The model (5.1 & 5.2) is derived by the composition of the two Hill-type functions (Gjuvsland et al, 2007) from the original TS work and the light transfer function from Chapter 3.

\[
\begin{align*}
LacI' &= \frac{e_l}{1+\frac{TetR\beta \cdot K_{aTc}^\gamma}{K_{aTc}^\gamma + aTc^\gamma}} + b_l - d_l LacI & (5.1) \\
TetR' &= \left(\frac{1}{m + a}\right) \left(\frac{a}{1+\frac{RL^\nu}{c^\nu}} + m\right) \left(\frac{e_l}{1+\frac{LacI^\nu \cdot K_{IPTG}^h}{K_{IPTG}^h + IPTG^h}}\right) + b_t - d_t TetR & (5.2)
\end{align*}
\]

Where \(\alpha\) and \(\alpha_t\) are the maximum expression rates for LacI and TetR, respectively, \(b_l\) and \(b_t\) the basal expression rates while \(d_l\) and \(d_t\) the degradation rate, constants. The light input related parameters correspond to the leakiness that cannot be suppressed by light, while \(c\) is the Hill constant for the specific function. The Hill exponents, \(\beta, \gamma, \nu, h \) and \(j\) were set to 2. Leaky TCS activation \(m\) was set to 200, \(a\) to 950 and \(c\) to 0.005. Basal expression rates, \(b_l\) and \(b_t\) were set to 0.035. Degradation rates, \(d_l\) and \(d_t\) were set to 0.014, while \(e_l\) and \(e_t\), 2.2 and 3, respectively. Red Light (RL) was variable. Finally, \(K_{IPTG}\) and \(K_{aTc}\) were set equal to 2. IPTG and aTc were set at 0 except for obtaining the Figure 5.2.1 (c).
Figure 5.2.1. The red light tuneable toggle switch (LTS). (a) LTS diagram where the Plac ompR has replace Ptc-2, and the repressors are now fused with the reporter proteins (b) Phase plane showing how the nullclines shift for different light intensities driving the system from bistable to monostable. (c) Numerical simulations showing the shift of the IPTG concentration that is needed to switch the system based on the red light intensity present. (d) Bifurcation diagram for TetR concentration versus the LacI production rate, $e_l$, at different light conditions. The two limit points collapse at 0.1 W/m$^2$ resulting in a monostable system. (e) Limit point continuation for varying light intensity and $e_l$ at different leaky expression rates. For low leakiness, the Cusp Bifurcation point is lost, meaning that light induction is not strong enough to drive the system out of bistability, while high leakiness reduces the bistable region.
Chapter 5. A Light Tuneable Toggle Switch

Model Analysis

Performing a basic analysis of the phenomenological model reveals how red light can affect the operation of the system. Starting with the phase plane analysis and the example shown in Figure 5.2.1 (b) shows that in a hypothetical bistable switch parameter set, when red light intensity is increased (driving the TetR expression rate lower) the nullclines shift such that they can only intersect at one point. This point is the only possible stable steady state solution where LacI is found at high concentrations and TetR expression is driven down to basal levels.

Red light tuning of the TetR expression rate affected the chemical inducer levels needed for switching between states. Figure 5.2.1 (c) shows numerical simulations of the TetR and LacI steady state when IPTG is varied at four different light conditions. In these simulations, the system starts with high concentrations of LacI and repressed TetR expression. The IPTG concentration needed for switching under high red light intensity was approximately 6-fold higher compared to the concentration needed in the dark. Similar behaviour, but from the opposite direction is observed for varying the aTc concentration when the system starts from the initial condition where TetR high and LacI is low (not shown)

Analysing the system with MatCont (Dhooge et al, 2003), bifurcation diagrams were obtained for varying the maximum expression rate ($e_l, e_t$) of either side. Figure 5.2.1 (d) shows that the bistable region for $e_l$ is reduced as red light intensity increases, where the two limit points bounding the region collapse into a monotonic curve, suggesting that the system is monostable to the LacI side when red light intensity is maximised. It is worth noting that the expression rate values that predict a bistable system are increased in both magnitude and range as red light decreases. The shift in the limit points when both red light and $e_l$ are varied was analysed with respect to other parameters. The most interesting realisation, shown in Figure 5.2.1 (e), was the fact that when the leakiness of the promoters (or basal expression rate when fully repressed) is high, the bistable region is significantly reduced. Very low leakiness drives the system to a point where red light induction by itself cannot drive the system out of bistability (given the specific parameter set).

This phenomenological model was used to assess qualitative features of a light tuneable genetic toggle switch. The parameter values used do not necessarily reflect an accurate representation of the physical system. However, parameters like expression rates and degradation rates, as discussed in chapter 1 (section 1.2) are control dials that can be tuned further if needed to bring the physical system within the parameter range that yields similar qualitative characteristics. In addition, the fact that an additional chemical input for each side can be used (IPTG for LacI and aTc for TetR) means that the system can be driven in and out of the bistable regime, under multiple different combinations of chemical and light induction.

In the next sections of this chapter, I discuss how a mechanistic model was derived in an attempt to gain more quantitative insights into the proposed system and the experimental implementation and assessment of a light tuneable switch.
5.2.2. Mechanistic Model Derivation

For building a more detailed mechanistic model, the modelling framework applied in Chapter 4 can be applied for modelling the Plac/ompR promoter, which controls the activity of the TetR side of the TS. In addition, the LacI expression under the PtetO-1 had to be included. Below the kinetic scheme used is shown:

For OmpR and Cph8 expression the same block of reactions discussed in Chapter 4 is used:

\[
\begin{align*}
\emptyset & \overset{a}{\underset{d_\alpha}{\rightarrow}} \text{OmpR} \\
\emptyset & \overset{e}{\underset{d_\epsilon}{\rightarrow}} \text{Cph}8^{pr} \\
\text{Cph}8^{pr} + \text{OmpR} & \overset{c}{\rightarrow} \text{Cph}8^{pr} + \text{OmpRp} \\
\text{Cph}8^{pr} & \overset{k_y}{\leftrightarrow} \text{Cph}8^{pr} \\
\text{OmpR} & \overset{a_1}{\underset{d_{OmpR}}{\rightarrow}} \text{OmpRp} \\
\text{OmpRp} & \overset{d_h}{\rightarrow} \emptyset \\
\text{Cph}8^{pr} & \overset{d_i}{\rightarrow} \emptyset
\end{align*}
\]

The promoter dynamics are modelled following the same framework as in Chapter 4, which is based on Stricker et al. (2008) where each promoter is modelled as a set of occupancy states. Protein expression, folding, dimerisation, and tetramerisation are also modelled, while degradation due to cell dilution is assumed for all free floating or DNA bound species.

**Promoters**

\[
\begin{align*}
P_{0,j}^{'} + 6 \text{ OmpRp} & \overset{b}{\underset{R_{1}}{\rightarrow}} P_{1,j}^{'} \\
P_{i,0}^{'} + \text{LacI}_4 & \overset{2k}{\underset{k_{2}}{\rightarrow}} P_{i,1}^{'} \\
P_{i,1}^{'} + \text{LacI}_4 & \overset{k}{\underset{2k_{2}}{\rightarrow}} P_{i,2}^{'} \\
P_{i,2}^{'} & \overset{k}{\underset{k_{2}}{\rightarrow}} P_{i,3}^{'} + i \cdot 6 \text{ OmpRp} \\
2P_{i,2}^{'} & \overset{k}{\underset{2k_{2}}{\rightarrow}} P_{i,3}^{'} + i \cdot 6 \text{ OmpRp} + \text{LacI}_4 \\
P_{L,0}^{'} & \overset{k}{\underset{k_{2}}{\rightarrow}} P_{0,0}^{'} \\
P_{0,0}^{'} & \overset{5}{\underset{n_1}{\rightarrow}} P_{0,0}^{'} + mt + mg \\
P_{1,0}^{'} & \overset{4}{\underset{n_2}{\rightarrow}} P_{1,0}^{'} + mt + mg \\
P_{0}^{'} & \overset{1}{\underset{n_3}{\rightarrow}} P_{0}^{'} + ml \\
mt & \overset{1}{\underset{h_1}{\rightarrow}} mt + \text{TetR}_v \\
\text{TetR}_v & \overset{k_y}{\rightarrow} \text{TetR} \\
\text{TetR} + \text{TetR}_v & \overset{k_h}{\rightarrow} \text{TetR}_v \\
\text{mt} & \overset{1}{\underset{h_{uf1}}{\rightarrow}} \text{ml} + \text{LacI}_v \\
\text{LacI}_v & \overset{k_y}{\rightarrow} \text{LacI} \\
\text{LacI} + \text{LacI}_v & \overset{k_y}{\rightarrow} \text{LacI}_v \\
\text{LacI}_v & \overset{k_y}{\rightarrow} \text{LacI}_v \\
\text{mg} & \overset{1}{\underset{h_{GFP1}}{\rightarrow}} \text{mg} + \text{GFP}_v \\
\text{GFP}_v & \overset{k_h}{\rightarrow} \text{GFP}
\end{align*}
\]
For modelling the effect of the two chemical inducers, IPTG and aTc, the following Hill functions were used for assigning the rate of association of LacI ($k_r$) and TetR ($k_a$), respectively:

$$k_r = k_{r,r} \left( \frac{1}{1 + \left( \frac{\text{IPTG}}{k_{r,1}} \right)^b} + c_{r}^{\text{min}} \right)$$

$$k_a = k_{a,a} \left( \frac{1}{1 + \left( \frac{\text{aTc}}{k_{a,1}} \right)^b} + c_{a}^{\text{min}} \right)$$

For reasons discussed below, the prototype implementation of the dual reporter LTS, shown in Figure 5.2.1 (a), did not perform as expected. Therefore, the system was re-designed as a single reporter LTS version which does not use SsrA tags as shown in Figure 5.3.1. The model presented above is based on the single reporter LTS.

The simulations shown in Figure 5.2.2 indicate that the transfer function of IPTG to the output of the system is very similar to the behaviour shown with the phenomenological model (Figure 5.2.1). Light is expected to shift the curve to higher concentrations, but in this model, the shift was shorter. In addition, the fact that JT2 that was used for the experimental implementation of the system, had a native LacI repressor being expressed, has some interesting implications. The transfer function of GFP (Figure 5.2.2 (c)) as IPTG is increased, appears to be very steep, in accordance to bistability and hysteresis features, shown in the bifurcation diagrams in section 5.2. However, shortly after the steep transition from low to high, the curve switches to a mode of a slower, gradually increasing function for higher IPTG concentrations. For this effect, caused by the native lacI gene, is evident in the data presented below. The time course simulations showed that red light inhibits GFP from reaching its maximum potential levels, as observed in the simulations for dark condition. What is more interesting, is the fact that the traces are all increasing from the same time point in the simulation, instead of giving a timer behaviour (Ellis et al, 2009). This response was assumed to be masked, again by the presence of the native lacI.
Figure 5.2.2. Model simulations for the LTS. The specific model is based on the single reporter LTS, however it models a hypothetical second reporter (mCherry) for gaining further insights. (a) Time course simulation for variable IPTG under red light illumination or (b) dark. (c) The end points of these traces are projected as a function of reporter versus IPTG concentration. The bi-phasic mode of the function is due to the modelled native, unregulated, lacI gene in JT2.
Chapter 5. A Light Tuneable Toggle Switch

5.3. Experimental Implementation

5.3.1. Assembly of the Light Tuneable Genetic Toggle Switch

The original Collins toggle switch comprises tetR transcribed under a LacI regulated promoter (Ptrc-2) and a lacI gene under the control of Ptet-O1 promoter. In a bicistronic setup, the GFP gene is placed downstream of the tetR gene so that higher green fluorescences indicates that the TetR side is high and LacI is low. For clarity reasons this state is referred as 'high state' while the opposite where LacI is high and TetR and GFP are low is called the 'low state'.

However, the toggle switch on which the proposed modified version was based initially, was a dual reporter-repressor fused version with SsrA tags for fast degradation, kindly provided by Dr James Arpino (unpublished work, personal communication). Shown in Figure 5.3.2, this TS still uses the Ptrc-2 promoter, with one LacO1 binding site, to control TetR expression while LacI is expressed from a PTet-O1 promoter. A GFP is placed bicistronically, downstream of the LacI stop codon, while the mCherry gene is assembled downstream of tetR. Both repressor:reporter cassettes are isolated by a double terminator from the origin of replication and antibiotic resistance cassette. The plasmid map is shown in Figure 5.3.2 (a). The light tuneable version of the toggle was built by PCR amplification of the Plac/ompR from the pL-PloPCB plasmid with primer overhangs with the AvrII and AscI restriction enzyme recognition sites. Both the amplicon and the pRG021 plasmid were digested with AscI and AvrII enzymes and ligated by T4 ligase. This step replaced the Ptrc-2 promoter with Plac/ompR. The resulting plasmid was then digested with PciI and ligated into the corresponding restriction sites in the pL-Cph8ec (Cph8 constitutive expression). Both the resultant pRG021-C8 plasmid and pL-PCB (constitutive) were co-transformed into a JT2 (ΔenvZ) strain and assayed under different light and chemical conditions discussed in the following section. The above cloning process, is shown schematically along with the corresponding plasmid maps in Figure 5.3.2.

The data obtained from this particular system, as discussed in the next section, showed that it did not exhibit the predicted behaviour. Although evidence of light sensitivity were observed and resulted in shifting the switching point of the chemical inducer concentrations, the cells, suffered from growth defects and abnormal expression patterns, making the obtained data non-comparable and non-conclusive across the different conditions. The problem was pinpointed to burden and degradation machinery overloading.

Hence, it was decided that the best course of action was to rebuild the light sensitive toggle switch (LTS) based on the original TS constructs that use a single reporter and no degradation tags. Similarly, this cloning process shown in Figure 5.3.3, started by amplifying the Plac/ompR promoter, using SphI and AgeI restriction sites, in order to ligate the amplicon upstream the RBS-E and PtetO-1 promoter, in pIKE107. For expressing the Cph8 TCS, the pL-Cph8FD casette, which comprises both the chromophore production and receptor genes was used. Several attempts to clone this construct into the pIKE107PlO plasmid were unsuccessful. The reasons for the latter, were found into non-mapped restriction sites, not reported in the acquired plasmid maps of pL-PCB(A) and pL-PCB(S), unspecific products from PCRs possibly due to the large constructs with repeated
sequences, like terminator palindromes, but highly possible deleterious effects when all genes are expressed from a high copy number plasmid. Working around this problem, a two plasmid system was chosen as the best choice under the circumstances. For this, the pL-Cph8FD was digested with EcoRI and PstI so that the entire cassette is extracted and ligated into the pSB3CS vector backbone (p15A origin), obtained from the Registry™ of biological parts. This backbone, consists of a p15A origin and Chloramphenicol resistance cassette. The co-transformation, of pIKE107Plo and resulting pL-Cph8FD(w), was carried out in a JT2 strain (ΔenvZ) and EBP238 (ΔenvZ and ΔlacI). The main difference is the fact that JT2 expresses two copies of the LacI repressor, the native constitutive gene and the TetR regulated one on pIKE107, while EBP238 lacks the native gene.

In addition for improving the fold difference between the ON and OFF states of Cph8, the J23109 upstream cph8 was replaced with the stronger J23110 promoter while the J23108 upstream ho1 was replaced with J23105. As a control, single transformations of the pIKE107Plo or pCPh8FD(w) were carried out. In addition, the tetR gene on pIKE107Plo was replaced by a λCI gene, amplified from the pTAK117 plasmid. This replacement, aimed at obtaining a similar sized plasmid, but remove the repression effect on the LacI site.

![Network diagram of the single reporter Light tuneable Toggle Switch (LTS)](image)
Chapter 5. A Light Tuneable Toggle Switch

Cloning Workflow for the double reporter version of the LTS

Figure 5.3.2. Construction of the dual reporter version of the light inducible toggle. The Plac/ompR promoter replaced the Ptrc-2 promoter upstream of the tetR gene in one cloning step. A second cloning step aimed at combining the constitutively expressed Cph8 receptor with the light inducible toggle circuit. The pRG021-C8 plasmid was co-transformed with a PCB production plasmid (pL-PCBc) under constitutive expression into the available ΔenvZ strains.
Cloning Workflow for the single reporter version of the LTS

**Figure 5.3.3.** Construction of the single reporter version of the light inducible toggle. Similar to the first version the Plac/ompR promoter replaced the Ptrc-2 promoter upstream of the tetR gene in pIKE107. Step 2 was carried out in order to replace the origin of replication and antibiotic resistance cassette of the pL-Cph8FD plasmid discussed in Chapter 3 with p15A origin and Chloramphenicol resistance of the pSB3C5. A complementary step aimed at creating a control where the LacI side cannot be repressed, while in steps 2 and 3, the constitutive promoter strength driving PCB production was increased while the Pom::eCFP cassette was removed in order to increase the OmpR-p mediated induction on Plac/ompR.
5.3.2. Characterisation of the pRG021-C8 Dual Reporter Toggle Switch

The pRG021-C8 plasmid that was obtained after replacing the Ptrc-2 promoter with the Plac/ompR was co-transformed with pCph8 into JT2 (ΔenvZ), EPB238 (ΔenvZ, ΔlacI) and RU1012 (ΔenvZ, PompC::lacz). The phenotype of the colonies grown on Petri dishes (in the dark) was surprisingly diverse, even though the genetic constructs were verified by sequencing prior the co-transformation. Red, green and yellow colonies were observed, indicating GFP, mCherry or simultaneous expression of both. The above could occur if the plasmids are unstable resulting in different copy numbers per cell, possible recombination between the multiple terminator sequences or random mutations at sites other than the loci were the cloning was performed. The fact that this diverse phenotype occurred with approximately the same frequency in all three of the host strains suggests that the observed instability of the system was not strain specific.

For selecting colonies that were more likely to exhibit the predicted behaviour, a preliminary screening was performed. Single colonies were picked and transferred into a 96-well plate based on the layout shown in Figure 5.3.4. and grown for 3 hours with chloramphenicol and ampicilin at 37 °C. At the end of the incubation time two new plates were prepared with 200 μl of LB media in each well, appropriate antibiotics with one plate supplemented with 5 mM IPTG and the other with 500 ng/μl of aTc. The chemical inducer concentration used was presumed sufficient to drive the expression towards either side of the switch under growth conditions in the dark. Colonies that appeared green on the aTc plate and red in their corresponding well in the IPTG induced plate were selected for further assays.

Figure 5.3.4 (b) shows that most of the selected colonies did not switch from green to red or the opposite. The fluorescence was either too weak to be observed by eye or the genetic circuit did not work as expected. A number of strongly fluorescent wells instead of switching from green to red or vice versa, when induced towards the opposite side, simply decreased their fluorescence levels. Nevertheless, six wells appeared to be responding to the chemical inducers as expected. The highest fold difference between red and green for both sides was observed in the EPB238 transformants. A JT2 clone (G3/6) exhibited the brightest green fluorescence when induced with aTc, while the brightest red fluorescence was observed in an RU1012 strain (R1/11) grown with IPTG. While it is expected for a JT2 strain to be shifted further into the green region because of the native constitutive amounts of LacI (as opposed to the ΔlacI EPB238) it is contradicting the fact that the same trend was not observed in the RU1012 host which is also expressing the native lacI gene. Even more so, RU1012 has an additional OmpR-p 'sink' through the genome integrated PompC::LacZ cassette. No conclusive explanation could be drawn at this point, since the observed inconsistencies in the expression behaviour and the low number of colonies able to respond to the chemical inducers as expected pointed towards a highly unsustainable system.

The next step was to assess the selected clones for light sensitivity. In a 96-well plate format, a gradient of chemical inducers was applied as shown in Figure 5.3.5. More specifically IPTG concentrations ranging from 50 mM to 0 mM and aTc concentration from 0 to 500 ng/μl was applied across each row of the two replica plates. Each row was inoculated with one of the clones selected during the preliminary screening step and
Figure 5.3.4. Preliminary screening of the strains transformed with the pRG021C8 and pL-PC8c plasmid. Individual colonies from petri dishes (a) of either green [G], red [R] or yellow [Y] colour were picked, grown and split into two replica 96-well plates (b) with one containing IPTG and the other aTc. The colonies that exhibited sufficient fluorescence for both sides (green under aTc and red under IPTG) were selected for further characterisation. The selected wells are marked with light blue circles. (c) Fluorescence from selected colonies as measured on the plate reader.
the plates were left to grow for 4 hours, keeping one in the dark where the other one was irradiated with saturating red light. From the results of this assay (Figure 5.3.5) was shown that most of the strains were sensitive to the light conditions applied and shifted the chemical gradient point where red fluorescence turns to green. Clones G1/3, R3/8, G3/6 and R1/11 shifted their switching point towards lower aTc values. For G3/6, R3/8 and R1/11 this chemical switching point is found within the aTc region, while for no inducers or IPTG the system rested on the TetR:mCherry site high. (indicating monostability to TetR). The latter, suggests that the expression (depending on the promoter and the RBS strength) of the TetR side is stronger than the LacI:GFP. For these clones the shift was approximately a 10-fold lower aTc concentration, from 50 to 5 ng/μl of aTc. Interestingly, G1/3 exhibited a much larger shift from a dark switching point at 0.05 ng/μl aTc to 0.0005 mM IPTG under red light. In addition it was stable to either state of the toggle switch at the ‘no inducer’ point. This indicates that the specific clone can be switched from the TetR side to the LacI side and vice versa without the use of any chemicals, but only with light. G1/3 is an EPB238 based strain, while the other R3/8, G3/6 are JT2 and R1/11 is an RU1012 based strain. The observed shift was towards the expected direction, as red light, negatively regulates the expression of TetR, hence expression from PtetO-1 is less repressed resulting in higher production of LacI::GFP and further repression for TetR::mCherry.

It is worth noting that for all clones, very high concentrations of aTc inhibited growth, something that is clearly visible in the right most column of the induction plates. On the other hand very high IPTG concentrations resulted in slightly lower mCherry expression, even if the cell growth rate was not severely affected. The exception to the latter was R3/8, which exhibited high red fluorescence at no inducer to low IPTG concentration while for IPTG concentration higher than 5 μM the red fluorescence intensity was decreased to almost basal levels. An interesting fact, was that despite the very low levels of red fluorescence that should correspond to low levels of TetR in the system, no increased lacI/green fluorescence was observed, indicating that this was the result of low overall protein production within the cell. This can be the result of toxic or burdensome conditions on the cells.

The overall fluorescence levels from either side of the TS for the EBP238 based clones, were significantly lower compared to the other two strains. In particular, the red fluorescence is at levels not visible by eye under the transilluminator box, but was only detectable on the plate reader, placing them at about 20% of the corresponding JT2 and RU1012 signals. This, at first sight, was conflicting with the fact that EPB238 is ΔlacI, so one would expect the red side (TerR) to be more dominant. However, there are a number of reasons that could partially explain the observed phenotype. The ompR gene in this specific strain is a fusion of OmpR and YFP used in Batchelor & Goulian (2006) to image the localisation of ompR to EnvZ. This fusion could be disruptive to the binding or the OmpR-p dimerisation on the PompC operators used in the Plac/ompR driving down the overall expression rate from the double input promoter. The latter hypothesis, fails to the fact that some EPB238 gave high red fluorescence in the screening assay (Figure 5.3.4 - R1-3 rows). In addition the authors of the strain reported the use of it for regulating PompC (Batchelor & Goulian, 2006). A last possible explanation can be that the difference in the genotype, can result in different tolerance levels for expressing such a large genetic circuit. The measured growth curves revealed that the lag phase for EPB238 was indeed
Figure 5.3.5. Induction plate assay. (a) Images of the two replica plates grown in the dark or light. Each row was inoculated with a different clone and the media used in each column was supplemented with different amounts of chemical inducers in order to apply a gradient of IPTG and aTc. From the images, the only clone that gave a strong visual outcome was G3/6 were the switching point of the illuminated cells was shifted by an order of magnitude towards lower aTc concentrations. Similar shifts, can be seen for G1/3 and R3/8 and R1/11 although the green visual output was significantly weaker. It is also worth noting that in the presence of high concentrations of aTc (500 ng/μl) most of the clones did not grow or grew significantly slower compared to the other conditions. (b) Endpoint measurements of the green and red fluorescence across the chemical inducer gradient for cells grown in the light or dark. For G3/6 and R3/8, the switching point between red and green was observed at 50 ng/μl of aTc in the dark and 5 ng/μl when the grown in the presence of red light. Unexpectedly red fluorescence for R3/8 at high IPTG concentrations (>0.05 mM) drops almost to basal levels but this drop does not seem to be able to switch the circuit to the LacI/Green side. Interestingly, G1/3 switches around the 'no inducer' point as the dark incubated cells switched at 0.05 ng/μl aTc but the red illuminated ones at 0.0005 mM IPTG. Thus, this strain can potentially be driven to either of the two states solely by using light, without any chemical supplements. On the other hand, the results seem to be significantly noisy, making it hard to draw any conclusions. The reasons for the latter as discussed in the text were attributed to the abnormal growth behaviour of the specific system. The data shown were normalised for the maximum expression level for each sample (Max fluorescence = 1). This was done to smooth out the noisy data and in order to highlight the switching points at the expense of missing the information on the relative fluorescence level difference between different samples.
Chapter 5. A Light Tuneable Toggle Switch

significantly extended compared to JT2 and RU1012.

Attempts to characterise further the system revealed problematic dynamics more likely due to the unstable genetic constructs and burden effects that inhibited growth for increasing amounts of inducers. More specifically time course assays were performed where the fluorescence and cell density over time was recorded. The OD and doubling time results shown in Figure 5.3.6 (a-b) suggested that for an increasing amount of aTc the cells are growing slower with extended lag phases, up to the point of complete inhibition of growth at 500 ng/μl. This caused the normalisation method to yield counter intuitive results, since the time evolution of the fluorescence could not be followed across different conditions in a comparable way. More specifically, cells that exhibited no or very little growth their OD remained close to the lower limit of the sensitivity of the microplate reader. When the fluorescence reading is normalised by an OD which lies at the boundaries or below the linear range of detection it yields an overestimated value. This problem is avoided in cells growing normally, since the normalised values can be corrected by subtracting the autofluorescence of the medium and controls. However, when in the problematic cells that do not grow as fast as the control cells this baseline correction is not valid. An additional problem was caused by the pre-pierced film designed to seal the plate to prevent evaporation but allows for better gas exchange. The time needed to transfer the samples growing at 37 °C from the incubator to the plate reader is enough for vapour condensation to be formed on the film, something that distorts the OD measurements. To correct for the latter, the method was adjusted so that the film is removed under sterile conditions right before the measurement, followed by the replacement with a new film, leading to clearer measurements at the expense of longer time intervals outside the incubator and risk of contamination across adjacent wells.

A fluorescence microscope revealed the nature of the growth problems observed in cells grown in the presence of aTc. Figure 5.3.6 (c) shows brightfield images of cells grown on agar pads containing either IPTG or aTc. While in the IPTG cells grew normally and started forming regular bacterial monolayers, in the presence of aTc they appear to be forming long elongated cells that fail to divide. Similar images have been published before (Camberg et al, 2011) demonstrating that disrupting or eliminating the ClpXP activity within the host, is the primary reason for this phenotype. In fact, this is in complete agreement with this case because the toggle switch version used had an SsrA degradation tag on the LacI:GFP fusion while the TetR:RFP not and is only removed from the system by dilution due to cell division. This finding strongly suggests that when the system is induced towards the LacI-GFP-SsrA site the amount of degradation tags expressed in the system, congests the proteosome ClpXP which is responsible for recognition and cleavage of the tagged proteins. This queuing effect was more recently discussed and modelled by Cookson et al (2011). Therefore it is strongly suggested that this is the primary cause for the growth inhibition something that was further confirmed by Dr. James Arpino (unpublished data) where the same toggle switch constructs (not a light tuneable version) with lower RBS strengths on the LacI side recovered their full growth velocity. Another contributing factor for the abnormal growth could be the expression of a fusion protein that may result in misfolded proteins and aggregation, something that can activate the stress response pathway.
Figure 5.3.6. Growth profile of the strains transformed with the dual reporter LTS plasmids (pRG021-C8 and pLPCB). (a) Growth curves as OD$_{600}$ over time for clone G2/3 at different chemical inducer conditions in the dark. The plot reveals the large impact that the high chemical inducer concentration has on the growth rate. More specifically, very high concentration of IPTG (50 mM) seem to extend the lag phase by 200 min, while high concentrations of aTc (500 ng/μl) seem to be inhibiting growth. The fastest growing conditions for the specific host are moderate amounts of IPTG. (b) Doubling times for all assessed clones under different chemical inducer conditions in the dark. The growth inhibition when aTc is present was observed across all assessed clones while the IPTG toxicity effects were not significant in R3/9 and G3/6. For the other clones IPTG resulted into a 10-20 % increase in doubling time. R3/8 and R1/3 were the ones with the overall longest doubling time while G1/3, was the one with the fastest. (c) Microscopy image of G2/3 growing on an agar pad with IPTG or aTc. In the presence of IPTG the bacteria exhibited a normal growing phenotype forming monolayers of microcolonies. In the presence of aTc instead, the bacteria appear extremely elongated and unable to divide.
5.3.3. Characterisation of the pIKEPlo Single Reporter Light Toggle Switch

At this point it was decided that the best course of action was to redesign and re-implement the experimental system based solely on the original Gardner et al. (2000) TS plasmids that were kindly provided by Prof Jim Collins. This comprises the same regulators but have no SsrA tag and instead of being fused with reporters, a single GFP is used in a bicistronic setup (Figure 5.3.3). This theoretically should prevent any growth defects due to ClpXP overloading and reduce the burden caused by the expression of an additional reporter. Finally, it remains unclear if the LacI:GFP and TetR:RFP fusions were responsible for misfolding and protein aggregation of the reporters, something that can trigger the stress response pathway (Villaverde & Carrió, 2003). Any such possibility was eliminated by using the original bicistronic cassette of TetR and GFP downstream the same promoter.

The growth experiment shown in Figure 5.3.7 (a) confirmed that the cells grew at rates similar to the non transformed host strain JT2. For JT2 harbouring only pL-Cph8FD(w), a ~10 min increase in doubling time was observed when grown in the presence of aTc, more likely due to the toxicity of the EtOH in which aTc was diluted. The same was observed in JT2 transformed with both single reporter LTS plasmids in aTc. However, this was increased by an additional 20 min delay on the doubling time when IPTG was used. In addition, a ~10 min longer time for the cells to exit the lag phase was observed when they were transformed with the full LTS. Microscopy imaging (Figure 5.3.7 (b)) showed cells without any observable phenotypic defects and able to form regular *E. coli* microcolonies on agar pads under both inducers at saturating concentration. Overall, the above suggested that there is still some effect on the growth rate when the full LTS is expressed within the same host but, it is significantly smaller than the dual reporter LTS case. More importantly, the cell division inhibitory effect was eliminated allowing for comparable OD measurements during any further assays.

From the same growth assay both green and cyan fluorescence was recorded after growing the microcolonies in the dark for 3 h. Qualitatively, the microscopy imaging showed less green fluorescence when the agar pads were supplemented with 5ng/μl aTc compared with the ones supplemented with 1mM IPTG. In the aTc case, some basal green fluorescence was observed most likely due to promoter leakiness. However, a small population of cells was abnormally bright (red circle 1 in Figure 5.3.7 (b)). The second observation was that the GFPmut3b expression was relatively uniformly distributed while the eCFP levels appear to be highly variable. This can be attributed to a number of reasons. First eCFP is expressed from a lower copy number plasmid (p15A origin) compared to the high copy GFP from the CoE1 origin, something that can contribute to higher cell to cell variability due to stochastic gene expression. What was interesting is that the apparent variability in eCFP levels, which would reasonably suggest similar high variability in OmpR-p levels, contrasts with the fact that GFP expression which is also activated by the same regulator is more uniform. This can be an indication that the stochastic effects, due to the sequestering effects on OmpR-p shared between the two promoter pools, can be stronger on the low copy number promoter. However, the exact opposite result was found in stochastic simulations of the full mass-action model of the TCS (Appendix A 9.1.2), where it showed that under OmpR-p limitation it is the high copy reporter that exhibited high variability. Another possible
Figure 5.3.7. Growth assays for the single reporter LTS. (a) Growth curves for JT2 transformed with pL-Cph8FD(w) (blue) showed similar growth behaviour with the controls (black), doubling time ~63 min with 1 mM IPTG, while addition of 5 ng/μl aTc increased the doubling time by approximately 10 min. Larger growth cost was observed on strains harbouring also pIKE107Plo (red) where IPTG increased the doubling time by 20 min and aTc by 10 min. For either inducer condition these strains had an elongated lag phase by ~20 min. (b) Microcolonies of JT2 harbouring both plasmids, formed on agar pads. Unlike the dual reporter LTS case, the cells formed regular microcolonies without any observable phenotypic defects. Using suitable fluorescence excitation-emission filters allowed to decipher that both reporter, GFPmut3b from pIKEPlo and eCFP from pL-Cph8FD(w) are expressed at detectable levels, but more importantly in a distinguishable way. Red circles 1,2 and 3 showed single cells that were unusually bright to one of the wavelengths. 1 had strong GFP expression and eCFP comparable to the neighbouring cells , while 2 and 3 stronger eCFP with 3 showing a slight increase and 2 showing no increased of GFP compared to their adjacent cells. These suggest that fluorescence bleed-through is not prohibitive for further assays. Finally, as shown in boxes 4 and 5, GFP expression in the population was uniformly distributed while cyan fluorescence levels appeared to be highly variable amongst the colony.
reason can be that the pl-L-Cph8FD(w) is not stable under these conditions, resulting in recombination or its ejection from some of the cells. Adjacent cells may be degrading the antibiotic from the agar pad fast enough such that cells can reject the plasmid and still survive. The plasmid instability hypothesis can also be supported by the fact that overexpression of \textit{ho1} and \textit{pcyA} can be cytotoxic (Tabor \textit{et al}, 2011). In fact all experiments shown throughout this report had use freshly transformed cells older than 1 or 2 days, otherwise the light sensitivity of the samples was rapidly decreased. Finally, this experiment was not performed under strictly controlled light conditions (the microscope was not set up as a dark room) and this could have resulted in some ambient light affecting the Cph8 activity.

The final information that could be extracted from this preliminary assay, was whether the two reporters can be distinguished under an experimental setup. When two fluorescent reporters are expressed, there is a possibility of 'bleed-through' between the two channels, due to overlaps of their excitation/emission spectra. Using the specific bandwidth for green (540 ± 20 nm) and cyan (460 ± 25 nm) seemed to be sufficient to attenuate such effects. Circles 1, 2 and 3 in Figure 5.3.7 (b) show individual cells that were abnormally bright compared to their adjacent cells but to only one of the two channels. Although, the fact that eCFP levels are positively correlated to GFPmut3b levels, as both are up-regulated by OmpR-p, does not allow for safe conclusions, these results suggest that the machine can differentiate between the two fluorescent proteins using appropriate bandwidths. These are shown in Methods section 2.1.2.

As discussed in the cloning section, 5.3.1 the Plac/OmpR promoter replaced the pTrC-2 upstream tetR and the plasmid was co-tranfromed with pl-L-Cph8FD(w). In this case the RBS strength used was much lower, so that the bacteria had to cope with less burden compared with the dual reporter LTS. This however, did not allowed visual (naked eye) observation of the fluorescence, hence all the following data shown were readings collected using the microplate reader according to the protocols discussed in Chapter 2 (section 2.1.2).

\textbf{Steady State Characterisation}

\textit{Chemical Induction Response}: The steady state transfer function of GFP versus IPTG levels under dark or light conditions was tested using a 96-well microplate setup. Based on layout C shown in Methods (Chapter 2, section 2.1.2) two replica microplates were prepared. An IPTG gradient was applied across different rows. One plate was kept in the dark (wrapped in aluminium foil) and the second was grown under the red LED device (0.5 W/m²). As shown in Figure 5.3.8(a), a ~3-fold increase in green fluorescence was observed for the LTS samples grown with IPTG concentration higher than 10 μM for both dark or illuminated cells. This suggested that the system is still responsive to chemical induction however the switching point from low to high state was now found in the IPTG region as opposed to the dual reporter LTS. An additional 3-fold (from 150 AU to 450 AU) increase in green fluorescence was observed between samples grown in the dark compared to the illuminated samples while the cyan fluorescence showed a 3.5 to 4-fold difference regardless of the IPTG concentration. However the average cyan fluorescence across the IPTG gradient showed a relatively higher variability compared with the control strain JT2 that harbours only the light reception plasmid (pl-L-Cph8FD(w)). The latter maintained low green fluorescence value since it does not have the \textit{gfp} gene (with the
Figure 5.3.8. (a) LTS response on chemical inducers under light or dark conditions. (b) Hill fits for IPTG induction and comparison between LTS and dLTS constructs. The results are discussed in the text.
exception of the sample grown with 0.1 IPTG). The latter was most likely due to cross contamination across the microplate wells. The control samples, JT2 transformed only with pIKEPlo, showed no light sensitivity in the green fluorescence channel but interestingly were responsive to IPTG induction. This suggested that the promoter still has some activity even in the absence of the EnvZ HK domain. This can be attributed to the fact that OmpR-p can still be phosphorylated by non cognate to OmpR kinases like CpxA (Siriyaporn & Goulian, 2008; Groban et al, 2009) or the promoter can weakly recruit RNA polymerase without OmpR-p bound. It is worth noting that the maximum green fluorescence point was slightly shifted between the 0.1 to 1 mM IPTG for the LTS and control strain, respectively. The cyan fluorescence of the control remained low, but it was still detectable at levels higher than the no-GFP control. In also showed some light sensitivity although the standard deviation bounds for light and dark were overlapping. This can be partially the result of cross contamination. Separate experimental runs (data not shown) indicated that some LB autofluorescence is quenched after long light exposure (especially under blue light) however this drop should be embedded in the correction factor measured from JT2 strains with no plasmids. Other reasons for this drop can be attributed to photo-bleaching of GFP (Bogdanov et al, 2009) by small molecules that can act as electron acceptors in the media. This effect could have been masked in the data for samples expressing eCFP at high rates.

\textit{Hill Function Fitting:} A similar assay, shown in Figure 5.3.8 (b), was run using layout B (Section 2.1.2) where the IPTG response curve of the full LTS was compared against the dLTS variant where tetR was replaced with the CI repressor. This yielded a similar sized plasmid but deactivated repression on the lacI:gfp side. The IPTG response curve showed a slightly increased 4.5-fold difference across low and high IPTG concentration compared with the LTS sample shown in Figure 5.3.8(a). The 3-fold difference under light was the same. The corresponding values for the dLTS were 5-fold induction for IPTG concentration, but only 2.6-fold for light induction. Fitting of the data to a Hill function of the form shown in Figure 5.3.8 (b) using the cftool, yielded a number of interesting results. First, the Hill fits with the lowest error are plotted and their corresponding parameter values are given in the table of Figure 5.3.8 (b). Parameters $a$ and $b$, as expected, rested to the values determined by the maximum and basal expression levels for each sample. $a$ equalled the difference between the maximum expression level and basal $b$. The Hill constant $k$, which is determined by the dissociation constant or the concentration of inducer (IPTG in this case) needed such as half the LacI molecules are induced (Ang et al, 2013). This parameter was estimated at 0.09 - 0.11 for the LTS but increased at 0.15-0.19 for the dLTS samples, indicating that for the latter higher amount of inducer is generally needed for reaching its maximum expression level. This is more likely because LacI abundance is higher in the dLTS case since no TetR is present to repress the LacI expression. Finally, the parameter $n$, which represents normally the cooperativity degree, and an indicator of the steepness of the curve, was approximately 5 units higher for the LTS samples. Values of $n$, around 2 are fairly reasonable considering the mechanism of LacI repression, discussed earlier in Chapter 4. However the increased values do not seem to have any physiological meaning. The reason for this increase is can be interpreted by the bifurcation diagrams Figure 5.2.1 and in Gardner et al. (2000). Starting from the low state, as LacI expression rate decreases (due to the increase of IPTG), a system is expected to reach a limit point after which a step transition from OFF to ON occurs. This means that
Figure 5.3.9. Surface plot of two dimensional induction by light and IPTG. The surface plot shows the LTS activity in terms of GFP, for a range of IPTG concentration $10^{-6}$ up to $10$ mM and red light intensity between $0$ to $0.5$ W/m$^2$. [IPTG] caused a 3-fold rise in green fluorescence for concentrations greater than $0.01$ mM in the absence of light. Illumination with $0.05$ W/m$^2$ causes a minor drop in fluorescence, but once this value approaches $0.1$ W/m$^2$, a ~3-fold drop in fluorescence was observed.

**Fluorescence for Variable IPTG and Light Conditions**

![Surface plot of two dimensional induction by light and IPTG](image)

**Time Course Assay with Red Light Pulse**

Figure 5.3.10. Light pulse assay. The figure shows samples growing in the dark (in the plate reader), followed by a 2 hour light pulse (red curves) and then, continuation of growth in the dark. Replicates of samples were covered so that they are not illuminated (black curves). The effect of light, that resulted in decreasing rates of fluorescence rise, was observable in less than an hour of the initiation time of the pulse. The pulsed cells showed a small rise in fluorescence when they were placed back to the incubator, before reaching the stationary phase.

![Time course assay with red light pulse](image)
Hill fits for bistable systems cannot have a physiological meaning for $n$ since for a step function it should be infinite. In other words, fitting a Hill function to a step response system like the toggle switch, is not justified for bistable switches. However, in this work was done simply to have a comparable steepness and switch point metric with the dLTS. It is worth noting that even if this data indicate some trends in behaviour, they were very limited, hence, the statistical uncertainty was very high, to the point that no safe conclusions can be drawn. Measurements in a more controlled environment and in higher resolution of induction conditions has to be performed. The reason that this was not possible at the time of this characterisation is discussed further in Chapter 7 (Conclusions and Discussion).

Two Dimensional Response Curve: Finally, for the static characterisation, a dual induction plate was set up based on Layout A, where IPTG concentration was varied using a similar gradient with as the assays above, but also by using neutral density (ND) filters, a gradient of light intensity was applied. Starting from the left-hand, columns 1-2 have no light let through (aluminium foil light shield), columns 3-4 with 25% filter, 5-6 50%, 7-8 75%, 9-10 90% filter and ending with the two right-hand columns 11-12 being completely exposed. On each row, a different concentration of chemical inducers was applied. Rows A-E had 10-0 mM IPTG. The results (Figure 5.3.9) showed again, a 3-fold induction between 0.01 and 0.1 mM of IPTG concentration was observed and an additional 3-fold difference between the samples covered in aluminium foil and the illuminated once with the maximum intensity (0.5 W/m$^2$). What was interesting from this assay is the shift in the shape of the plot towards a more strict AND gate-like response compared to the corresponding surface plot during the Plac/ompR characterisation in Figure 4.3.4 (Chapter 4). In this case, the basal expression is lower and sharply rises once the high state conditions are reached, as opposed to the more linear-like response observed in the Plac/ompR case in Chapter 4. These differences can be attributed to three main reasons. First, that the regulatory network difference here is a toggle switch circuit that as discussed above can generate steeper response curves than a one way repressor device. The second reason is the fact that pL-Cph8FD(w) is used in this case where it is expressed from a lower copy number plasmid, reducing any basal kinase activity, making the available OmpR-p levels lower compared to the high copy ColE1 replicon based pCph8 used for the Plac/ompR characterisation. Finally, the fact that LacI in this case is over expressed from a high copy vector ColE1 results in a more strict repression.

Dynamic Characterisation

The light Pulse Assay: In order to investigate the response time between the application of light and the change in fluorescence a time course 96-well plate based assay was set up according to the layout B (Methods 2.1.2) by preparing a gradient of chemical inducers between rows A-D in LB and inoculating with live cells. Half the contents of the well was transferred to rows E-H so that the plate has two identical halves. The entire microplate was grown at 37 °C with shaking in the Gemini EM (Molecular devices) microplate reader and the fluorescence was recorded at 525 nm and 620 nm in 5 minute intervals. During this time the entire microplate was grown in the dark. After two hours of incubation, aluminium foil was used to cover the upper half of the microplate (rows A-D) and it was placed in a shaking incubator at 37 °C for three hours under saturating red
Figure 5.3.11. Time course and stability assay. (a) Stability assay for samples starting in either high or low state and grown under different conditions for 18 hours. Measurements were taken in 1 hour increments for the first 5 hours and a final measurement was taken at 18 hours after overnight incubation. The data shown in the four slices are not sequential, as this was a staggered-start time course reconstruction. The four slices A-D were run in parallel, starting at different induction conditions and connected at the points after the overnight incubation. These points were assumed identical to the end of the initial overnight incubation prior to the execution of this experiment. The model simulation overlays are discussed further in the text. The time course assay with the full spectrum of [IPTG] used is shown in (b), where the thick lines are the highest 10 mM [IPTG] (continuous line) and the lowest at 0 mM [IPTG] (dashed lines).
light. During this period the fluorescence was recorded every 1 hour. Subsequently, the microplate was placed back into the microplate reader and left to grow overnight under dark conditions while fluorescence was recorded every 10 minutes overnight. The time course traces are shown in Figure 5.3.10. During the first part of the experiment (0 to 180 minutes) all the wells showed an increase in fluorescence at very similar rates. The highest rate was observed as expected in the samples with the highest concentration of IPTG (black and red thick continuous lines) while the samples without any IPTG exhibit the lowest rate (black and red dashed lines). Thinner gray and red lines indicate intermediate concentrations of IPTG used in this experiment. The light blue shading indicate all intermediate IPTG gradient concentrations. From the assay it was shown that light has a fast effect on the GFP expression rate which was estimated to be a 2-fold decrease, from ~ 80 to 40 AU/h for 10 mM IPTG.

Time course- stability assay: 96-well plate based assays were carried out to determine the time evolution and stable states of the system. In this case the samples were grown overnight in conditions such that the assay was initiated at either the high or low state. For clarity reasons ‘High State’ was defined as the state when TetR and GFP expression rate is high. Similarly to the static assays, a gradient of IPTG concentration was used according to the layout D shown in Methods 2.1.2. One of the replica plates was covered with aluminium foil while the other was grown under 0.5 W/m² red light. Time point measurements were taken in 1 hour increments for 5 hours and a last measurement after overnight incubation. Figure 5.3.11 summarises these results. In (a) a ‘staggered start’ (Olson et al, 2014) adaptation was used to reconstitute a long time course from 4 assays that were actually run in parallel shown in (b). These individual samples were assumed to have finished at an endpoint that is the identical starting condition of a different assay, e.g. the samples that were induced to the high state during the assay will end with the same high GFP levels as the samples induced to the high state during the overnight incubation prior to the start of the assay. With this assumption, the pseudo-continuous experiment was tested against the model discussed below.

Cells initiated in the low state showed an increase in fluorescence only when IPTG was present while in the absence of chemical inducer or presence of aTc the GFP production is maintained at basal levels. The above applies for both illuminated and non-illuminated samples with the difference that when GFP expression is switched ON in the presence of IPTG the fluorescence was higher for samples kept in the dark. A qualitatively different behaviour was observed for the samples initiated at the high state. The samples grown without IPTG show GFP expression levels that almost match the basal expression levels (below 200 AU) during the first 3 hours of growth. From these, the ones kept in the dark, recovered during the overnight step to ~250 AU while the illuminated ones maintained low levels (~100 AU). The IPTG induced samples showed a similar initial drop in fluorescence, however, they retained approximately 100 AU and 200 AU higher fluorescence, for the light and dark conditions, respectively, compared to the lowest point of the drop for the samples without IPTG. During the overnight step, the fluorescence settled at 220 AU for the illuminated samples, but for the dark condition rose by 2-fold to ~550 AU. From the above results a number of interesting observations can be noted. The initial drop in fluorescence (during the first 3 hours) was more likely to be the result of fast dilution. During the incubation steps prior to the assay the culture will have reached stationary phase, where the cells
are growing slow and GFP accumulates because of the lack of dilution from cell division. When diluted into fresh medium for the actual assay, cells begin to grow again and re-enter the exponential phase, resulting in faster dilution of GFP. The second key observation is the fact that cells starting in the high state settled at higher levels when the inducer was removed. To help interpret this result, three fluorescence intensity bands were identified as shown in the colour coding in Figure 5.3.11 (a), a low (0-200 AU, Blue) a medium (200-400 AU, Yellow) and a high one (> 400 AU, Green). The samples starting high and then grown in the dark ended in the medium band, but when light was used the fluorescence settled in the low band.

The first hypothesis that can explain the data is that the LTS is bistable in the dark between the medium and low levels and what causes the drop from the high band to medium is the fact that JT2 is not a ΔlacI strain, so some LacI will be present even if TetR is dominant over the LTS LacI. This can explain why when IPTG was used the samples reached high fluorescent levels but when it was removed they settled at medium values, a drop that is attributed to the unregulated expression of native LacI which cannot be repressed by the high levels of TetR. Instead when IPTG was removed the native LacI downregulates TetR expression, but at levels not low enough to release the repression of the regulated LacI from the PtetO1 promoter and flick the LTS to the low TetR / high LacI state. The same behaviour was observed not only for the sample without IPTG but also samples with concentrations lower than 0.1 mM (cluster of points at ~ 220 AU in Figure 5.3.11 (b), plot D).

A contradicting hypothesis can be that the LTS is monostable to the low TetR and high LacI state and the observed recovery of the fluorescence of the dark/-IPTG samples that started high, was simply because the cells reached the stationary phase before the existing GFP / TetR pool was diluted to levels similar to the samples starting at the low state. Once the stationary phase was reached, dilution was lower and allowed basally expressed GFP to be accumulated on top of the remnants of the pre-existing pool. However the drop in fluorescence, of the samples starting in the high state, during the exponential phase was down to 106.5 AU (lowest point), which was only 19 AU higher than the lowest point (87.7 AU) of the samples that started in the low state. If the recovery was solely based on accumulation of basally expressed GFP when LacI has switched OFF the Plac/ompR promoter, then a similar rise (from their lowest points) should have been observed in all samples grown in the dark regardless of where they started. Instead, a 2.17-fold and a 1.34-fold rise was observed for samples starting in the high and low states, respectively.

The model was used in order to test these hypotheses further, by introducing a reaction for the transcription of native unregulated LacI.

$$\emptyset \xrightarrow{E} ml$$

This model structure, although not quantitative, was able to explain the qualitative behaviour of the system in terms of the native LacI gene effect. The simulations shown in Figure 5.3.11 (a), show how this toggle switch mechanism with native LacI crosstalk can qualitatively reproduce the observed behaviour. However, the dynamics for the first 5 hours, especially for the samples starting in the high state are different between model and experiments. This can be explained by the fact that the model does not simulate a variable dilution...
rate due to differences in growth rate. The most important caveat of the modelling work though, is the fact that the model is semi-quantitative. It was fitted with known parameters for the light input module and the IPTG induction function, however a mapping of simulated molecules to fluorescence arbitrary units was not available. Hence, the translation rates were arbitrarily fitted as a 1:10 ratio of AUs to GFP molecules. For this reason the model cannot be used for a conclusive answer, but points only to some qualitative results that support the initial hypothesis.

A summary of the LTS qualitatively interpretation is shown in Table 5.3.1. Following the colour coding in Figure 5.3.11 (a) the ON state is assumed to be above 200 AU (Green) while the OFF state below 200 AU. When cells were induced to the high level expression band and then the inducers were removed, GFP expression settled in the medium band in the dark. When light was present GFP levels were reset to the low state.

<table>
<thead>
<tr>
<th>Starting GFP Levels</th>
<th>Light Condition</th>
<th>IPTG</th>
<th>No-Inducer</th>
<th>aTc</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIGH</td>
<td>DARK</td>
<td>548.5</td>
<td>231.3</td>
<td>&lt;100</td>
</tr>
<tr>
<td></td>
<td>LIGHT</td>
<td>248.2</td>
<td>121.2</td>
<td>&lt;100</td>
</tr>
<tr>
<td>LOW</td>
<td>DARK</td>
<td>471.7</td>
<td>120.4</td>
<td>&lt;100</td>
</tr>
<tr>
<td></td>
<td>LIGHT</td>
<td>233</td>
<td>58.9</td>
<td>&lt;100</td>
</tr>
</tbody>
</table>

Measurements in a chemostatic environment is needed for concluding on the stability of the system. The above experiments were all time finite and any observed behaviour can be potentially different for fast growing cells for long periods of time. While using the model, for investigating the above hypothesis, it was clear from the simulations shown in Figure 5.3.12, that if the specific LTS is successfully cloned into a suitable ΔlacI strain, given that the model parameters fall within relevant biological range, it can potentially recover a bistable behaviour with characteristics similar to the template TS (Gardner et al, 2000). Yet, light can still exert its effect for shifting the IPTG response curve.
Figure 5.3.12. Time course comparison between (a) +lacI and (b) ΔlacI strains. What is worth noting, is that for the latter, none of the IPTG concentrations of cells starting at the high state and kept in the dark (black curves), converged to the intermediated band. Its only two possible states, are high and low. Similarly, for all light conditions for (b) there are three clusters that can be distinguished, at low, medium and high values, where the trajectories converge, while for (a) the clusters, are visible, but an obvious wider spread distribution of the trajectories over time, is observed.
Chapter 6
Theoretical design - A light Tuneable Oscillator

6.1 Background

6.2.1 Model Implementation and Preliminary System Design
6.2.2 Genetic Constructs and Proposed Plasmids Maps

6.2 Model Refinement and Modifications

6.2.1 LacI Mediated Loop Formation
6.2.2 Degradation Mechanism

6.3 System Assessment and Re-Design

6.3.1 Sensitivity Assessment
6.3.2 System Re-Design for Independent Tuning
6.3.3 On the Robustness of Oscillations - Stochastic Model

This Chapter is a theoretical study on how an inherently unstable system such as an oscillator can be coupled with the developed light input components in order to realise fine tuning of its characteristics. A short introduction of existing oscillators is given, which leads into more detailed description of the most robust synthetic genetic oscillator implemented to date in bacteria by Stricker et al, (2008). This dual feedback oscillator is a network that can accommodate the TCS responsive to light through the double input promoters discussed earlier in Chapter 4. The preliminary design that is initially proposed, along with the corresponding plasmid maps and genetic constructs, was put through a first assessment cycle, which showed that light control components can shift the period and the amplitude independently. However, further refinement and modifications on the model to represent the biological system more precisely revealed that such independent tuning is not possible with the specific connectivity of the network. A more detailed assessment cycle, including a thorough sensitivity analysis on the characteristics of the amplitude and period shifts for red and green light revealed potential ways for system re-design that can yield a more orthogonal waveform tuning scheme. Finally, using the stochastic counterpart of the model, a period-amplitude variability profile was extracted, pinpointing the network components that contribute more to the noise in the oscillation.
6.1. Background

Engineering genetic oscillators has been a focus of synthetic biology (Elowitz & Leibler, 2000; Stricker et al. 2008). Oscillatory behaviour is encountered in some of the most crucial biological processes such as the cell cycle, developmental processes, energy metabolism, circadian cycles, neuronal and cardiac cycles (Fung et al., 2005)) and in large metabolic pathways were the precise temporal expression of genes is important for saving resources and maximizing the efficiency of a particular biological function. A number of synthetic oscillatory networks have been proposed and implemented during the past decade, although designs have been considered theoretically as far back as 1963 (Goodwin, 1963). The Goodwin oscillator (Figure 6.1.1 (a)) is the simplest device that can exhibit oscillations, consisting of only a single regulator repressing itself in a classic negative feedback loop. The oscillations arise from the time delay between the transcription of the repressor and its translation, multimerisation and final operator binding and repression. Time delay in a negative feedback loop is the core principle motif, on where all natural and more recent synthetic oscillators are based on. Another well known example of a synthetic genetic oscillator, also known as the “repressilator” (Elowitz & Leibler, 2000), is a circuit of three repressors each regulating transcription of the next in a cyclic fashion (Figure 6.1.1 (b)). Although, as first published this network exhibited some clear oscillatory behaviour, it was not robust, and unable to generate a truly sinusoidal sustainable function, as the output (of the GFP reporter at least) would accumulate over time and skew the mean level of the oscillating genes. In addition, it was reported that only 40% of the cells exhibited oscillations (Elowitz & Leibler, 2000).

More recently, Stricker et al. (2008) demonstrated a different kind of network that, to a large extent, exhibited the characteristics that the repressilator was lacking. The network consists of one activator (AraC) and one repressor (LacI), each acting on the other but also having feedback action on their own expression (Figure 6.1.1 (c)). The design is based on the fact that there is a time delay between the effect of the two regulators, that drives the oscillation. In a hypothetical cell where at time zero both regulators are absent, a cycle starts with the basal expression of both. However, AraC acts as a dimer and LacI as a tetramer, meaning that the effect of the activator will occur faster than that of the repressor. Promoter bound AraC results in an increased rate of expression (relative to the basal rate), thereby generating a quick burst of expression before the levels of LacI tetramer are high enough to effectively shut down the expression from both promoters. Levels of both regulators then begin to decay due to dilution (from cell division) but mainly due to active degradation from the ClpXP proteosome. The two regulators are expressed from the same promoter (Plac/ara-1) as shown in Figure 6.1.1 (c). When the levels of LacI drop sufficiently, expression from the promoters is released and another burst occurs starting the next cycle. This system was shown to be robust as nearly 98% of the cells exhibited oscillations with remarkably consistent periods. More importantly, the period was tuneable from...
**Figure 6.1.1.** Types of synthetic oscillator designs. (a) The Goodwin oscillator is the simplest form consisting only of one negative feedback loop. (b) The repressilator, which is a ring of three repressors each acting on its successor. (c) The dual feedback oscillator where two genes act on each other and themselves through a positive and negative feedback loop. For each repressor a network diagram and time course simulation is given, along with experimental data as published by Elowitz & Leibler (2000), for the repressilator.
15 to 60 min by using different concentrations of inducers, IPTG for LacI and arabinose for AraC. Some of the limitations were the fact that there is no way to independently tune the amplitude, the limited range of frequency tuning, and the fact that the system was not independent of the reporter (discussed below).

Here, I explore the possibility of building a light tuneable oscillator were we can utilise one light wavelength for tuning the amplitude and a different wavelength for tuning the period, given a fixed concentration of chemical inducer present. From all the proposed synthetic genetic oscillators up to date, the dual feedback oscillator is considered as the one that is most robust, evident by its tuning capabilities and in the analysis of its published detailed mass-action model (Stricker et al, 2008), its ability to function in different organisms like Salmonella typhimurium (Prindle et al, 2012) or even tumor-targeting bacteria (Danino et al, 2012). Hence, the proposed light tuneable design in this work, was based on the specific oscillator. The principle that the system is based on makes it robust enough to allow the introduction of modifications with lower risk of driving the system outside its oscillatory state. On the downside, the period is correlated with the peak levels of regulators (or reporters) since the longer the period the more time these proteins have to accumulate, or vice versa, the higher the levels of these proteins the more time is needed for the degradation mechanism to effectively reduce them and restart the cycle. This chapter addresses to what extent (under biologically acceptable conditions) it is possible to engineer an oscillator that can be tuned independently in terms of frequency and amplitude.

### 6.1.1. Model Implementation and Preliminary System Design

As a first step, the published model for the specific oscillator shown below was used to analyse and gain insights on the mechanics of the oscillation, specifically in terms of how expression rates change with respect to AraC and LacI abundance over one cycle. Figure 6.1.2 shows a break down of five identified phases during a cycle. Phase A, is the lag phase where ClpXP is not saturated and degrades rapidly the newly formed protein, while in phase B the increased expression rate from the unoccupied promoters overloads ClpXP resulting in slow degradation. This allows for AraC dimers to accumulate and activate the promoters which will result in a rapid expression burst, leading into phase C where the LacI tetramer level will ‘catchup’ and decrease the overall expression rate by repressing Plac/ara-1. Finally, at phase D the expression rate falls below the degradation rate, since most of the active promoters are repressed in a looped state, and the regulators levels fall slowly through phase E until the promoters return to their original unoccupied state where they can initiate the next cycle. Table 1, in Appendix C shows the duration of these phases both for the original and redesigned system under different conditions.

Although this is a high-order non-linear system, making an accurate interpretation of the dynamics is challenging, the effect of chemical inducers can be mapped on these phases. Arabinose is the natural inducer of AraC that increases its affinity to the DNA operator site. AraC dimers bound on the Plac/ara-1 promoter increase its expression rate. Hence when present, arabinose can lead to a faster occurring B phase since the positive feedback loop from the basal AraC levels will be stronger. It can also increase the slope of the bursts leading to higher levels of regulators during phase B when LacI has not reached an effective repressing concentration.
Model Implementation and Preliminary System Design


**Promoter - regulator binding**

\[
\begin{align*}
P_{0,i,j}^{a/r} &+ a_2 \xrightarrow{k_{a}} P_{0,j}^{a/r} \\
P_{i,0}^{a/r} &+ r_4 \xrightarrow{k_{r}} P_{i,1}^{a/r} \\
P_{i,1}^{a/r} &+ r_4 \xrightarrow{2k_{r}} P_{i,2}^{a/r} \\
P_{1,2}^{a/r} &\xrightarrow{k_{L}} P_{1,2}^{a/r} + a_2 \\
P_{0,2}^{a/r} &\xrightarrow{k_{L}} P_{1,2}^{a/r} \\
P_{0,0}^{a/r} &\xrightarrow{k_{L}} P_{0,0}^{a/r}
\end{align*}
\]

\[j \in \{0,1,2\}, i \in \{0,1\}\]

**Transcription, translation, protein folding and multimerisation**

\[
\begin{align*}
P_{0,0}^{a/r} &\xrightarrow{k_{b}} P_{0,0}^{a/r} + m_{a/r} \\
P_{1,0}^{a/r} &\xrightarrow{ab_{b}} P_{1,0}^{a/r} + m_{a/r} \\
m_{a/r} &\xrightarrow{k_{b}} m_{a/r} + a_{b/r} \\
a_{b/r} &\xrightarrow{f_{b}} a / r \\
a + a &\xrightarrow{k_{m}} a_2 \\
r + r &\xrightarrow{k_{m}} r_2 \\
r_2 &\xrightarrow{k_{L}} r_4
\end{align*}
\]

Where \(P_0\) is the Plac/ara-1 promoter expressing AraC and \(P\) the promoter expressing LacI, while the index \(i\) shows the number of AraC dimers and \(j\) LacI tetramers bound on the specific promoter. Index \(L\) represents the looped state of the promoter. \(m\) represents mRNAs, \(a\) AraC molecules, \(r\) LacI molecules, with subscript numbers indicating their multimerisation order (2 for dimers and 4 for tetramers) while index \(uf\) indicates unfolded proteins.

**Degradation**

\[
\begin{align*}
P_{1,j}^{a/r} &\xrightarrow{f(X)} P_{0,j}^{a/r} \\
P_{j,0}^{a/r} &\xrightarrow{f(X)} P_{1,0}^{a/r} \\
P_{1,1}^{a/r} &\xrightarrow{2f(X)} P_{1,1}^{a/r} \\
P_{1,2}^{a/r} &\xrightarrow{2f(X)} P_{1,2}^{a/r} \\
P_{1,0}^{a/r} &\xrightarrow{2f(X)} P_{1,0}^{a/r} \\
P_{0,0}^{a/r} &\xrightarrow{2f(X)} P_{0,0}^{a/r}
\end{align*}
\]

\[f(X) = \frac{\gamma}{c_e + X}\]

Where \(f(x)\) is the function that determines the variable degradation rate due to the ClpXP overload. This function was derived by standard Michaelis-Menten kinetics with \(\gamma\) representing the \(V_{\text{max}}\) of the reaction and \(C_e\) the Michaelis-Menten constant \((k_m)\). \(X\) is the substrate and equals the sum of all SsrA tagged species. That includes all LacI and AraC found in monomers, dimers, tetramers, free or bound on promoters.

**LacI-DNA binding constant**

\[
k_r = k_r \left( C_r^{\text{max}} - C_r^{\text{min}} \right) \frac{1}{1 + \left( \frac{\text{IPTG}}{k_{r}} \right)} + C_r^{\text{min}}
\]

**AraC-DNA binding constant**

\[
k_a = k_{a-} \left( C_a^{\text{max}} - C_a^{\text{min}} \right) \frac{\text{[ara]}^{\text{el}}}{k_{a-e} + \text{[ara]}^{\text{el}}} \cdot \frac{1}{1 + \left( \frac{\text{IPTG}}{k_{e}} \right)^2} + C_a^{\text{min}}
\]
Chapter 6. A Light Tuneable Oscillator

**Figure 6.1.2.** Numerical simulation of the dual feedback oscillator. The figure shows the AraC dimer (blue line) LacI tetramer (red line) and the degradation rate function over time. In the lower left corner the state of the Pa promoter is shown. The blue line represents the free promoter and the red line the promoter that is found in a looped state and unable to initiate transcription. The rest of the lines represent intermediate states with the first index denoting AraC dimers bound on the promoter and the second, the number of LacI tetramers. The letter L shows a Pa promoter that is looped. An oscillation cycle starts with basal expression of the regulators from the unoccupied promoters (A). Here, the degradation rate is too large to sustain any detectable levels of regulators and the promoters found in a looped form from the previous cycle are returned to their relaxed state. The regulator levels remain low through this phase until their expression overloads the ClpXP machinery represented by the degradation function, $f(x)$. Once $f(x)$ becomes small enough the expressed regulators start folding and accumulating to form dimers and tetramers. However, the AraC dimers are formed faster than LacI tetramers, both because of the multimerisation order and the gene/protein size difference. (B) This results in the occupation and activation of the free promoters by AraC ($Pa00 \rightarrow Pa10$ transition in the model), which leads to a rapid burst of expression for both regulators. (C) As the LacI tetramers accumulate, with a delay relative to AraC dimers, they bind and repress the promoters ($Pa10 \rightarrow Pa11 \rightarrow Pa12 \rightarrow PaL2$ transitions). This, gradually decreases the number of active promoters and decelerates the initial expression burst. (D) Eventually, nearly all of the active promoters will transit to a looped state. The expression rate falls below the degradation rate and the overall protein levels decline. Phase D also shows an approximately 3 min delay between the point that LacI tetramers reach their peak level compared to AraC dimers. (E) Finally the regulators are steadily degraded by ClpXP until they reach a point (after ~25 min) where LacI repression is lifted and the looped promoters can return to their initial relaxed state ($PaL0 \rightarrow Pa00$ transition). This marks the initiation of the next cycle.
Higher peak protein levels lead to a larger period, as the degradation phase E is extended. However, as arabinose concentrations increase they are expected to have a counter effect, since the increased rates of expression will lead to a faster accumulation of LacI and hence a faster-occurring C phase. This, limits the extent that the peak level of the oscillation can be increased by arabinose.

In the case of IPTG, similar effects can be observed. IPTG binds to LacI and decreases its affinity to Plac/ara-1, weakening the negative feedback loop strength from a given pool of tetramers. For an effective repression and burst termination, the LacI tetramer concentration needs to increase even more. In this case phases B and C will be prolonged, again generating longer periods and higher protein peak levels. On the other hand, large enough IPTG concentration can have the same counter effect as arabinose, since the repression is less effective but the repressor is produced at its maximum possible rate. The published data and model for the specific oscillator (Stricker et al, 2008), showed an initial increase in the period in the range 0-2 mM of IPTG followed by faster cycles, for concentrations higher than 2 mM. This drop however, turned out not to be inherent to the oscillator mechanics, but due to the fact that IPTG was reportedly interfering with arabinose binding on AraC (Sung et al, 2007 in Stricker et al 2008). This suggests that as IPTG is increased above 2 mM the positive feedback loop is weakened leading to faster and lower amplitude cycles.

Based on the above, a system re-design solution aimed at controlling the length of phase C as a key point for tuning the frequency was explored. Light induction can be introduced into this system via the proposed double input promoters discussed in Chapter 4 and control the length of phase C. The simplest strategy found determined from simulating preliminary models (not shown due to large size), was to express an additional LacI gene from the Plac/ompR or Plac/ccaR promoter so that light can be used to alter the strength and consequently the delay of the negative feedback loop. This additional promoter’s activity was able to oscillate in phase with the core oscillator (Figure 6.1.3) since it is repressible by LacI and additionally, when not repressed it could exhibit a variable expression rate, subject to light conditions. This will allow the user to change the levels of the additional LacI from the second copy of the gene. In addition, to increase the range of this tuning, the additional LacI gene chosen was a mutant version which is unable to bind IPTG (LacI\textsuperscript{d}) (Hasan & Szybalski, 1995). This means that when IPTG is present, increasing Plac/ompR expression rate will enhance the negative feedback strength, not just by increased LacI abundance but also due to increased affinity of LacI\textsuperscript{d} (resulting from its IPTG insensitivity). Furthermore the additional gene impacts the ClpXP degradation mechanism. Increased SsrA tagged TF abundance from the additional lacI gene will shorten phase A, as a result of congesting the free ClpXP pool at faster rates.

For an amplitude tuneable design, the model had to be extended to include the actual output of the experimental system. That is an additional Plac/ara-1 promoter expressing a GFP, that was replaced in the this work’s model with the Plac/ccaR. Similarly to Plac/ompR, the oscillating LacI levels drive the oscillatory expression behaviour, while the presence of green light increased this rate, leading to higher peak levels for the GFP output.
The exact kinetic scheme used to model the proposed design, is discussed next. The original model was expanded in order to include the reporter gene and an additional promoter. In addition, each promoter's occupancy format is expanded so that it can interact with more than one variant of the repressor (LacI and LacI d). These variants can be the same molecule with different binding affinities to their chemical inducers or non-tagged versions. The expanded kinetic scheme which is summarised graphically in Figure 6.1.3 is:

Promoter - regulator binding

\[
\begin{align*}
P^{a/r}_{0,jk} + a_2 & \xrightarrow{k_{-1}} P^{a/r}_{1,jk} \\
P^{g}_{0,jk} + 2 \alpha_p / c_p & \xrightarrow{k_{-1}} P^{g/l}_{1,jk} \\
P^{a/r/g/l}_{1,00} + r_4 & \xrightarrow{2k_{-2}} P^{a/r/g/l}_{2,10} \\
P^{a/r/g/l}_{1,01} + r_4 & \xrightarrow{2k_{-2}} P^{a/r/g/l}_{2,01} \\
P^{a/r/g/l}_{1,10} + r_4 & \xrightarrow{k_{-2}} P^{a/r/g/l}_{2,10} \\
P^{a/r/g/l}_{1,01} + r_4 & \xrightarrow{k_{-2}} P^{a/r/g/l}_{2,01} \\
P^{a/r/g/l}_{1,11} + r_4 & \xrightarrow{k_{-2}} P^{a/r/g/l}_{2,11} \\
P^{a/r/g/l}_{1,12} + r_4 & \xrightarrow{k_{-2}} P^{a/r/g/l}_{2,12} \\
P^{a/r/g/l}_{1,02} + r_4 & \xrightarrow{k_{-2}} P^{a/r/g/l}_{2,02} \\
\end{align*}
\]

\[\quad \quad \quad \quad j \in \{0,1,2\}, k \in \{0,1,2\}, i \in \{0,1\}\]

\[\quad \quad \quad \quad 0 \leq j + k \leq 2\]

Where the promoters \( P \) with indexes \( g \) and \( l \) are the Plac/ccaR and Plac/ompR promoters controlling the GFP and mutant LacI d (\( r_q \)) expression, respectively. While the number of available operator sites remains unchanged, the index \( j \) represents the number of wild type LacI tetramers (\( r_q \)) bound and \( k \) is the number of mutant LacI d tetramers (\( r_q \)). A maximum of 2 tetramers of either type (or combination) can bind the promoters at any given point. \( Op \) and \( Cp \) represent OmpR-p and CcaR-p, respectively, where the former can only bind and enhance the transcription rate of the \( P^i \) and the latter of the \( P^0 \) promoter.
Transcription, translation, protein folding and multimerisation

\[ P_{0,00}^{a/r} \xrightarrow{b_{s}} P_{0,00}^{a/r} + m_{a/r} \]
\[ P_{1,00}^{a/r} \xrightarrow{a_{s}} P_{1,00}^{a/r} + m_{a/r} \]
\[ P_{0,00}^{g/l} \xrightarrow{b_{s}} P_{0,00}^{g/l} + m_{g/l} \]
\[ P_{1,00}^{g/l} \xrightarrow{a_{s}} P_{1,00}^{g/l} + m_{g/l} \]
\[ m_{a/r}^{g/i} \xrightarrow{e_{s}} m_{a/r}^{g/i} + a_{of} / r_{of} / r_{af}^{q} / g_{of} \]
\[ a_{of} / r_{of} / r_{af}^{q} / g_{of} \xrightarrow{g_{s}/r_{s}/e_{s}} a / r / r_{af} / gfp \]
\[ a + a \xrightarrow{k_{a}} a_{2} \]
\[ r + r \xrightarrow{k_{r}} r_{2} \]
\[ r_{2} + r_{2} \xrightarrow{k_{r}} r_{4} \]

Where \( m_{g} \) and \( g_{of} \) represent the mRNA and unfolded GFP, respectively.

Light receptor dynamics

\[ \text{Cph8}^{pr} + O \xrightarrow{k_{a}} \text{Cph8}^{c} \]
\[ \text{Cph8}^{c} \xrightarrow{C} O_{p} + \text{Cph8}^{pr} \]
\[ \text{Cph8}^{pr} \xleftarrow{f_{(LR)}} k_{y} \text{Cph8}^{pg} \]
\[ \text{Cph8}^{pg} \xrightarrow{f_{(LR)}} k_{y} + f_{(LR)} \text{Cph8}^{pr} \]
\[ \text{CcaS}^{pr} + c \xrightarrow{k_{a}} \text{CcaS}^{c} \]
\[ \text{CcaS}^{c} \xrightarrow{C} c_{p} + \text{CcaS}^{pr} \]
\[ \text{CcaS}^{pg} \xrightarrow{C} \text{CcaS}^{pg} \]
\[ \text{CcaS}^{pr} \xrightarrow{c_{p}} \text{CcaS}^{pg} \]
\[ \text{CcaS}^{pg} \xrightarrow{c_{p}} \text{CcaS}^{pg} \]

For modelling the TCS the 'compact' version of the light receptors was used as presented in Chapter 4.

Finally, the global enzymatic degradation function \( f(x) \) was edited in order to include the substrate \( X \), all the new variants of occupied promoters and combinations of tagged and non-tagged regulators that can be bound on them (The reactions are shown and discussed in section 6.2.2). The fully implemented computational model is shown in Appendix C. Although multidimensional, with a large number of parameters, the original oscillator model with its proposed parameter set seems to be in a good agreement with the experimental results published from Stricker et al, (2008). Hence, an initial assumption for the proposed modified model, was that since the same components and regulators are used, it is justified to use the same parameter values for the oscillators shared components, combined with the parameter set for the light receptors used in Chapters 3 and 4.
Figure 6.1.3. Diagram of the expanded model that includes the two double input promoters (Plac/ompR and Plac/ccaR) providing an additional negative feedback loop (dashed red line) and a light tuneable GFP transcription rate. Preliminary numerical simulations reveal that when green light is ON (t=200 min and t=600 min) the amplitude (peak GFP levels per cycle - green curve) gains a 10-fold increase while when red light is OFF the period of the oscillation drops from approximately 40 min to 15 min.
Summarising, the light tuneable design model shown in Figure 6.1.3 maintains the original core network as the oscillation generator but an additional negative feedback loop is introduced under the Plac/OmpR promoter. This loop is mediated by a mutant version of LacI (LacId) which cannot bind IPTG. The LacI operators on the Plac/OmpR promoter, will result in an oscillatory expression pattern driven by the core oscillator with the amount of LacId expression tuned using a variable intensity of red light. Furthermore the reporter output promoter is replaced with Plac/ccaR. Similarly, LacI expression will force oscillatory expression with a rate tuned by green light.

Shown in the simulations in Figure 6.1.4, when red light was used, the period was predicted to vary between 15-45 min with a significant effect on the amplitude. The faster the oscillation the lower the amplitude, since for fast cycles the GFP had limited time to accumulate. When green light is used, an increase in the amplitude of the oscillation from 500 molecules to 4000 occurred, with almost no effect on the period. In total, the proposed design has four user inputs that can be tuned during operation: red light, green light, IPTG and arabinose. The effect of light wavelength intensity on the amplitude and period is demonstrated in Figure 6.1.4 under fixed chemical conditions, while the effect of the chemical inducers on both the oscillation and the operating range for light induction is demonstrated in Figure 6.1.5. More specifically, as shown in Figure 6.1.5 the amplitude increased for increasing concentrations of arabinose reaching a plateau for values higher than 1%, with the period following that increase from 15 to just over 40 minutes. Green light induction could drive the peak GFP expression from approximately 3000 up to 45000 molecules with no impact on the period. IPTG caused a sharp increase in the period and amplitude for concentrations of 0-3 mM, followed by a decrease in both, for any concentrations higher than 4 mM. Green light, contolred the amplitude, however the corresponding charts for red light induction, suggested that both the amplitude and period are highly sensitive to the intensity of red light. More specifically, in the presence of saturating green light, the peak GFP levels dropped by more than 50% for some chemical inducer concentrations while the period, was also varied by a similar percentage. In addition, the tuneable range of the period seemed to be proportionally related to the amplitude of the oscillation for each chemical inducer combination. The higher the fold difference between the peak GFP levels for 100% and 0% green light, the higher the tuneable range for the period was.

On the other hand, for very high chemical inducer concentrations (i.e. > 2% arabinose and 20 mM IPTG) the oscillations collapsed into damped oscillations with the GFP reaching a stable steady state (grey shaded area in Figure 6.1.5. It was not possible to pinpoint a definite cause for this, however, some potential explanations can be suggested. For example, for high concentrations of arabinose the high expression rate from all promoters, can be such that the time delay between LacI repression and AraC activation is reduced. Alternatively, very high concentrations of IPTG are likely to eliminate entirely the negative feedback loop, driving the system to a continuous expression mode. In both cases, the inducer concentrations that enhance or lead to a continuous transcription activity seemed to be breaking the oscillatory pattern, while low concentrations, had the opposite effect. The only other non-oscillating point is for 0 mM IPTG, where LacI repression is strong enough to prevent the initial burst needed to start a cycle.
Figure 6.1.4. Modelled effect of green and red light on the amplitude and period of the oscillation of the modified model, under fixed chemical conditions (IPTG 2 mM, Arabinose 1%). Simulations for varying green light intensity predict that the amplitude (peak GFP level) can be tuned from 500 up to 4000 molecules without any effect on the period of the oscillation which is constant at approximately 40 min. On the other hand, the effect of red light induction is shown to be able to drive the system from a large amplitude, slow oscillation (40 min period) down to very fast oscillations (15 min period) at the expense of the amplitude levels.
Figure 6.1.5. Modelled effect of Arabinose, (a) and (c), and IPTG, (b) and (d), on the amplitude and period of the oscillation under varying light conditions for the originally proposed loop formation model (Stricker et al. 2008). The range of amplitudes that can be achieved for varying green light, (a) and (b), is shown to increase for arabinose concentrations up to approximately 2% while for any values higher than 2.2% the system falls into dammed oscillations (Grey shaded area). For (b) an increase from 0 to 2 mM IPTG is followed by a linear decrease up to 24 mM. For IPTG values higher than 24 mM the oscillations are not sustainable. In both (a) and (b) the period (blue markers) is not altered by the green light and follows the same trend as the maximum amplitude indicated by the top edge of the box plots. For (c) and (d) a similar trend on the amplitude is observed.
6.1.2. Genetic Construct and Plasmid Maps Design

The light tuneable oscillator discussed above requires a large number of genes and genetic parts to be cloned and co-expressed in the same host. The two plasmid system that the original oscillator (Stricker et al, 2008) was build on, can be edited and extended to accommodate the TCS along with the other necessary genes for their operation. An additional plasmid with a pSC101 origin of replication and a chloramphenicol resistance cassette can harbour the additional lacId under the PLac/OmpR. A three plasmid system was shown possible in Tabor et al. (2011) since the CoLE1, p15A and pSC101 origins belong to different incompatibility groups.

Figure 6.1.6 shows the proposed plasmid architecture. This specific setup was chosen such that the system could be realised experimentally with the lowest possible cloning steps by taking into account which constructs, synthetic operons and cassettes are already available from previous studies (Stricker et al 2008, Tabor et al 2011) and the current project. However, no suitable existing unique endonuclease recognition sites could be identified for this purpose, due to the fact that the plasmids are relatively large in size or carry repeated parts. Hence all the cloning steps proposed below must be carried out using an alternative cloning method like Gibson assembly (Gibson et al, 2009). These methods and recent advances in combinatorial assembly strategies (Casini et al, 2013) enable the assembly of each plasmid in only one reaction, if suitable overhang combinations are used for the PCR amplification of the required sequences.

The three plasmid system chosen is using ORIs within different incompatibility groups (CoLE1, p15A and pSC101) shown previously to be functional when the TCSs are expressed within the same host (Tabor et al, 2011). The first plasmid required is the pJS167-G1, which is based on the CoLE1 based kanamycin resistance, pJS167 (Stricker et al 2008), but with the Plac/ccaR promoter having replaced the Plac/ara-1 driving yemGFP expression. In two more cloning steps, the CcaR and CcaS expression sequences have to be amplified from pJT122 and cloned to form this plasmid with a final size of 11.5 Kbp. It is important that the amplification of ccaS starts from the terminator locus downstream of cph8 in pJT122 so that its ORF is isolated from the Pccpc-G2 green light depended promoter (the particular promoter is bidirectional). Another plasmid needed, is the pSB4C5 (pSC101 and chloramphenicol resistance) from the Registry of Biological Parts, harbouring an LAA tag as an insert (BBa_K365006). This will allow the cloning of the lacId sequence (excluding the c-terminus sequence that needs to be replaced by the LAA tag) upstream the specific sequence, followed by insertion of Plac/ompR upstream the ORF. Finally, the plasmid pJS167-TCS harbours the PLac/ara-1 controlled wild-type lacI found in the original oscillator study, (Stricker et al, 2008) as well as the Cph8-PCB cassette found in pL-Cph8FD (Chapter 3).

Another important criterion for the genetic construct design was to keep the parameters as close as possible to those, where the individual TCS were initially characterised (Chapter 3, section 3.2). This will allow the reuse in the Light Tuneable Oscillator (LTO) model parameters shown to adequately describe existing experimental data. More specifically the sequences were designed to be cloned into plasmids so that the copy number remains similar to that in the initially characterised system. araC and yemGFP, ccaS and ccaR were demonstrated to work in a high copy CoLE1 vector, while lacI, ho1 and pcyA from a low copy p15A.
Figure 6.1.6. The proposed plasmid maps for realising the light tuneable oscillator. Plasmid pJS169-G1 can be build in three steps. One to replace the existing Plac/ara-1 promoter on pJS169 driving the yemGFP expression with Plac/CcaR and the next two include the cloning of the CcaS cassette from pJT122 upstream the AmpR of pJS167 and insertion of the CcaR ORF upstream the Ppcp-G2 promoter. pSB4C-LacI\textsuperscript{a} can be build by initially cloning the lacI\textsuperscript{a} gene upstream the LAA (SsrA tag) sequence in pSB4C5-k36500 obtainable from the MIT's registry of biological parts. Next the Plac/OmpR promoter needs to be cloned upstream the lacI\textsuperscript{a} ORF. Finally the plasmid pJS167-TCS can be build in one step by inserting the Cph8 cassette the single plasmid Cph8 system from pl-Cph8FD (which contains both cph8 and PCB production genes). The first step for all proposed plasmids and the second step can be carried out in parallel.
Chapter 6. A Light Tuneable Oscillator

cph8 was successfully used in a p15A vector backbone in Chapter 5. One exception is Plac/ompR which was characterised only in p15A vector, while in this case it needs to be expressed from a pSC101. However, the fact that PompC was shown to be functional in such a vector (Tabor et al., 2011) suggests that its lacI repressible counterpart will be as well. Apart from plasmid copy number, an important factor is the potential shift in RBS strength when Plac/ompR has a lacI ORF instead of gfp. The downstream sequence interactions may result in different RBS strengths, so it is necessary when designing a suitable reverse oligo overhang (for amplifying Plac/ompR) to use an RBS calculator (Salis et al., 2011) in order to obtain a new RBS sequence of the same targeted RBS strength.

Finally, the proposed genetic system is designed in a modular way such that plasmids can be removed or replaced giving amplitude only, period only, or tunable by both along with controls for the any experimental assays. More specifically Table 6.1.1. shows that all three pJS167-G1, pJS169-TCS and pSB4C-LacI<sub>d</sub> plasmids, can express the dual wavelength responsive oscillator as described above. However, if the unmodified pJS167 is used, this will remove the green light-dependent amplitude tuning while maintaining functional the red light-dependent expression of LacI<sub>d</sub> that can tune the period. Alternatively, if the pSB4C-LacI<sub>d</sub> plasmid is removed then the period tuning actuator is eliminated but the green light dependent output is still functional. The pJS169-TCS is necessary in either case because it carries the chromophore production genes that are needed for both TCS.

The main limitations of these designs are the large total size of the constructs, that amounts up to 21.4 kb. Networks of these sizes are more likely to introduce burden in the host strain, however, it is not possible to know to what extent, without testing the physical system. In addition, the strain needed for this system to be expressed into, has to be a ΔrecA strain to avoid or limit recombination between the multiple copies of promoters and terminators (T0,T1,T2) within the same plasmid. The strain must also be ΔenvZ, ΔaraC and ΔlacI in order to eliminate any crosstalk of these native genes with the synthetic network components. Such a strain can be constructed with a single gene knock out step for envZ on the JS006 (MG1655 ΔaraC ΔlacI).

<table>
<thead>
<tr>
<th>Plasmid combination</th>
<th>Resistance Markers</th>
<th>Light Sensitivity</th>
<th>Tuneable by Light</th>
</tr>
</thead>
<tbody>
<tr>
<td>pJS167-G1, pJS169-TCS, pSB4C-LacI&lt;sub&gt;d&lt;/sub&gt;</td>
<td>AmpR, KanR, CmR</td>
<td>Red and Green</td>
<td>Period and Amplitude</td>
</tr>
<tr>
<td>pJS167, pJS169-TCS, pSB4C-LacI&lt;sub&gt;d&lt;/sub&gt;</td>
<td>AmpR, KanR, CmR</td>
<td>Red</td>
<td>Period</td>
</tr>
<tr>
<td>pJS167-G1, pJS169-TCS</td>
<td>AmpR, KanR</td>
<td>Green*</td>
<td>Amplitude</td>
</tr>
</tbody>
</table>

*The red light sensitivity of CcaS was not explore in the model at this point
6.2. Model Refinement and Modifications

6.2.1. DNA Loop Formation by LacI

The original model, on which the proposed design was based, was identified to have some inconsistencies with the biological mechanisms, proposed in literature. A first observation, was on the way that the LacI loop formation is modelled. The initial model (Stricker et al., 2008) shows that two LacI tetramers are needed to be bound on the promoter, to form the loop. Structural and kinetic studies (Rutkauskas et al., 2009), however have shown that only one LacI tetramer is responsible for DNA looping, where it can bind one DNA operator with each half of the tetramer as shown in Figure 6.2.1. To model the correct looping mechanism, a modified biochemical reaction set was used:

\[
\begin{align*}
P_{i,10}^{air} & \xrightarrow{k_i} P_{i,10}^{air} + A \\
P_{i,01}^{air} & \xrightarrow{k_i} P_{i,01}^{air} + A \\
P_{i,20}^{air} & \xrightarrow{2k_i} P_{i,20}^{air} + r_4 + A \\
P_{i,20}^{air} & \xrightarrow{2k_i} P_{i,20}^{air} + r_4 + A \\
P_{i,02}^{air} & \xrightarrow{2k_i} P_{i,02}^{air} + r_4 + A \\
P_{i,11}^{air} & \xrightarrow{k_i} P_{i,11}^{air} + r_4 + A \\
P_{i,11}^{air} & \xrightarrow{k_i} P_{i,11}^{air} + r_4 + A \\
\end{align*}
\]

The model was then simulated, for the loop-corrected mechanism and scanned its behaviour across different concentrations of chemical inducers and both light inputs. Unlike the graphs given in Figure 6.1.5, for the corrected model in Figure 6.2.2, the oscillations are sustained for high values of chemical inducers (>2% for arabinose and 25 mM of IPTG). The reactions, allowing additional loop formation events increased the overall rate that this occurs, given that the proposed parameter values are assumed correct and not modified. This fact, is in agreement with what Stricker et al. (2008) suggested for the loop formation rate when tested for the negative only feedback oscillator. For higher values of the specific rate constant the system seems to be generating stable oscillations across a wider range of IPTG concentrations, but at the expense of narrowing its tuneable range. It is worth noting, that the models discussed here, do not include cell growth and stress mechanisms, which in the physical system are expected to limit the range of chemical inducers that can be used. Modelling the correct stoichiometry and loop formation, can have significant implications on the tunability and the robustness of the proposed design when additional components are included.
Figure 6.2.1. Crystal structure of one LacI tetramer bound on two operators in loop formation. Red indicates DNA binding domains, green the core protein and blue the tetramerisation domain. Each dimer is indicated by light and darker shades. In addition the figure shows possible topoisomers of LacI-DNA looped complex (A1,A2,P1,P2,P1E). It is clear from the structure that only one tetramer is responsible for the loop formation. This figure was copied from Rutkauskas et al, (2009)
Figure 6.2.2. Simulated effect of the corrected loop formation mechanism. The Arabinose, IPTG and light effect on the amplitude and period is demonstrated. (a) Green light can vary the amplitude of the oscillation almost 10-fold, for arabinose concentrations higher than 0.5% and IPTG concentrations between 2-8 mM. For higher IPTG concentrations, the amplitude's tuneable range, is settled between 15000 - 1000 molecules per cell. Green light seems to have no effect on the period (indicated by the blue diamonds). On the other hand, red light (b) is shown to have an effect on both the amplitude and the period (red box plots and blue bars respectively). The system, however, unlike the original proposed loop formation scheme shown in Figure 6.2.2 seems to be robust in its oscillatory behaviour across all the concentrations of chemical inducers used here.
6.2.2. Degradation Mechanism and GFP SsrA Tag Feedback Effect

A crucial topic of the original model (Stricker et al., 2008) is the way that the enzymatic degradation mechanism is introduced. While in common biochemical mass-action models, the degradation rate is given by a constant, which is the sum of the dilution and natural decay rate of a particular molecule, in this case, the model uses a variable degradation rate for the SsrA tagged regulators. More specifically Michaelis-Menten kinetics are used to model the proteosome activity like a regular enzyme with tagged proteins as a substrate:

\[ f(x) = \frac{\gamma}{c_e + X} \]  

\[ X = r_{ef} + r_{ag} + a_{ef} + r'_{ag} + a'_{ef} + r + r'' + a + 2 \cdot r_{a2} + 2 \cdot r_{a4} + 2 \cdot a_{a2} + 4 \cdot a_{a4} + 4 \cdot r_{a4} + j \cdot P_{i,j}^{ag} + (i + j) \cdot P_{i,j}^{ag} + c_e + X \]  

where \( \gamma \) is the ‘processing’ capacity of the ClpXP machinery (proteosome), \( c_e \) is the Michaelis-Menten constant and \( X \) is the sum of all SsrA tags present in the system at a given time point. The latter, breaks down to one tag for each unfolded repressor, one per monomer, two per dimer and four for each \( lacI \) tetramer. \( X \) includes the multimers bound on DNA. As this equation shows, the more tagged proteins in the system the slower the degradation rate will be. This, naturally occurs from the fact that regardless of the type of molecule, the tagged population competes for the same number of finite ClpXP catalytic sites. The implications of this congestion or ‘queuing’ effect, was discussed more recently by the same group (Mather et al., 2010) and Cookson et al., (2011)).

Although the Stricker et al. (2008) model of the dual feedback oscillator, used the variable degradation rate form shown in (6.1) it was observed that the particular degradation mechanism in the model can be improved in a number of ways to reflect more precisely the experimental system. The modifications address some inconsistencies, with the assumed biological mechanism. These are:

1) The model neglects the contribution of the SsrA-tagged yemGFP that was used as a reporter, to the congestion effect for ClpXP. Although the specific model was demonstrated to describe the experimental data fairly well, it is imperative that for design purposes and modifications to the oscillator, the model has to account for the contribution of GFP to the degradation rate. This, can be simply achieved by adding one tag per GFP molecule into the calculation of \( X \) in equation (6.2):

\[ X = r_{ef} + r_{ag} + a_{ef} + r'_{ag} + a'_{ef} + r + r'' + a + 2 \cdot r_{a2} + 2 \cdot r_{a4} + 2 \cdot a_{a2} + 4 \cdot a_{a4} + 4 \cdot r_{a4} + j \cdot P_{i,j}^{ag} + (i + j) \cdot P_{i,j}^{ag} + g_{of} + g_{fp} \]  

2) Even with the correction shown in Eq(6.3), the model does not account for the contribution of cell dilution into the total degradation rate of each tagged species. The latter can be corrected by adding the relevant reactions where applicable:

\[
\begin{align*}
\text{a}_{ef} / \text{a} / \text{a}_{ef} / \text{r} / \text{a}_{ef} / r'_{ag} + a'_{ef} / r + r'' + a + 2 \cdot r_{a2} + 2 \cdot r_{a4} + 2 \cdot a_{a2} + 4 \cdot a_{a4} + 4 \cdot r_{a4} + j \cdot P_{i,j}^{ag} + (i + j) \cdot P_{i,j}^{ag} + g_{of} + g_{fp} & \xrightarrow{D=\text{Dilution}} 0 \\
\text{P}_{i,j}^{ag} & \xrightarrow{\text{Dilution}} \text{P}_{(i,j)}^{ag} / \text{P}_{(i,j)}^{ag} / \text{P}_{(i,j)}^{ag} / \text{P}_{(i,j)}^{ag} / \text{P}_{(i,j)}^{ag} / \text{P}_{(i,j)}^{ag} / 0,00 \\
\text{P}_{i,j}^{ag} & \xrightarrow{\text{Dilution}} \text{P}_{(i,j)}^{ag} / \text{P}_{(i,j)}^{ag} / \text{P}_{(i,j)}^{ag} / \text{P}_{(i,j)}^{ag} / \text{P}_{(i,j)}^{ag} / \text{P}_{(i,j)}^{ag} / 0,00 \\
\end{align*}
\]

Where \( dy \) is the dilution rate.
3) The final identified issue, was the validity of the quasi-steady state approximation used to derive Equation (6.1). The available literature was not clear concerning the assumption that forward and reverse binding rates of an SsrA-tagged substrate to ClpXP are large enough compared to the catalysis rate constant (Mather et al., 2010). To overcome the latter, but also allow for the SSA stochastic solver to be used, a final purely mass-action model was derived, including the necessary reactions for the first two corrections. This model simulates explicitly the ClpXP enzyme, ClpXP and SsrA substrate complex formation and enzymatic degradation step. The following general kinetic scheme was used for all free or promoter bound species bearing the tag.

\[ Y + n \cdot D \xrightarrow{df} D Y \xrightarrow{dcat} n \cdot D \]

where \( Y \) is any tagged species and \( n \) is the multimerisation factor, two for dimers and four for tetramers. Each free floating or DNA bound SsrA-tagged species can be removed enzymatically by the proteosome \( D \) (ClpXP) or diluted due to cell division with a rate \( dy \). These corrections in the degradation mechanism had some significant implications. The most important one is the fact that in this model for an increasing amount of GFP the finite amount of available ClpXP becomes increasingly sequestered, so that the global proteosome mediated degradation rate is decreased. This introduces an indirect feedback effect of GFP on the network meaning that the waveform, period and amplitude is no longer independent of the reporter. This is in agreement with Mather et al., (2010). The results obtained by re-simulating the corrected model, suggests that truly independent amplitude tuning is no longer possible with the specific network. More specifically, Figure 6.2.4 shows in (a) that while the competing and non-competing model (before the degradation modification) are initially in phase, when GFP expression rate is increased due to green light induction, this results into a larger period for the competing model (corrected model). (b) Shows how period and amplitude relate between the four extreme conditions of light induction (green or red only, both or none). The slope of the line, connecting the points between green being OFF and green ON, indicates by what
amount the frequency is affected due to the increasing amounts of GFP in the system. Similarly the slope of the red lines (connecting the red light OFF and ON conditions) show by what extent the amplitude is affected as a result of increased LacI\textsuperscript{d} production. Figure 6.2.3 (c) and (d) are showing the modified transfer functions of amplitude and period for green and red light intensities. The results showed that when green light is increased from 0 to saturating values (>0.1 W/m\textsuperscript{2}) a shift of approximately 14 min and 10 min, respectively, towards longer periods occurs. This is more likely due to the fact that when GFP numbers are increased the effective degradation rate from ClpXP for any tagged species is reduced because of the additional load. This allows for the regulators to accumulate at faster rates during the initial burst of transcription (Phase B initiation of each cycle) but take longer to be degraded (Phase E) and initiate the next cycle. The result shown in 6.2.3 (c) however is qualitatively different compared with the insensitive period shown previously in Figure 6.1.4. The second observation is that the peak GFP levels of the noncompeting compared to the competing model are significantly reduced. If the enzymatic degradation is slower, from first sight it can be contradicting the fact that peak levels for GFP (under load conditions due to green light activation) are decreased by more than 50 % as shown in Figure 6.2.3 (a). However this can be explained by the fact that the corrected model accounts in parallel the removal of GFP-SSrA and all tagged regulators by dilution due to cell division which immediately increases the total degradation rate for these species. Another possible reason is the fact that LacI tetramers are now being accumulated at a faster rate (during phase C) and reach concentrations that can effectively repress the protein production from all the LacI regulated promoters at an earlier point during each cycle.

Strangely, for Figure 6.2.3. (b) and (d) where the effect of variable red light was simulated, the amplitude shifts observed, where significantly milder compared to the corresponding simulations of the initial noncompeting model shown in Figure 6.1.4. More specifically shifting the period by varying the red light in the initial model from 21 to 41 minutes resulted into an amplitude shift of 27000 molecules (~64 %) of the green light tuneable range while for the corrected model the corresponding shift from 39 to 59 min was only 5000 molecules (~20 %) of the green light tuneable range. Although this is a very complicated high ordered network and pinpointing the exact cause for the latter is a challenging task, the simplest possible explanation can be attributed to the increased periods observed and the timing of phase C (LacI repression). For a given cycle period, the regulators can accumulate during the initial burst, driven by basal expression and enhanced by the positive regulation from AraC or CcaRp for GFP. The transcription stops, as soon as the LacI tetramers reach levels, such that all copies of the promoters are effectively in a repressed or looped state. The competing model resulted to an overall shift to longer periods, compared to the noncompeting one, for all possible starting conditions which, more likely due to GFP having more time to reach higher levels under ClpXP overloaded conditions, even if lacI\textsuperscript{d} is being produced (in the absence of red light). In fact a short time window exists during each cycle (phase B) where lacI\textsuperscript{d} will overload further ClpXP before it reaches an effective concentration to stop protein production, as the unfolded, monomers, dimers and tetramers are targeted by ClpXP while the repression occurs with a delay until and only if some tetramers have formed. The latter, together with the fact that GFP levels, in the presence of red light, were reduced, as stated earlier in this section, resulted to a much smaller difference of peak GFP for varying red light.
Numerical Simulations of the Corrected Model for the Degradation Mechanism

**Figure 6.2.3.** Numerical simulations demonstrating the effect of red and green light on the amplitude and frequency. (a) Comparison of a time course simulation with the initial kinetic scheme where GFP is not competing for the ClpXP sites (blue continuous line) and the updated model where GFP-SsrA is processed through ClpXP reducing its pool which can be accessed by the other tagged species (red dashed line). For 100% red light and no green light between t=0 and t = 200 both models are in phase but when green light is switched on, the increased amounts of GFP are shifting the oscillation to longer periods (in the competing model) due to the decreased degradation rate. (b) The period plotted against the amplitude , as estimated from the numerical simulations, for varying light conditions. Black dashed lines, show the derivatives that are used as the metric for the sensitivity for amplitude and period, on red and green light, respectively (discussed extensively in 6.3). $\Delta A^R$ and $\Delta A^G$ show the shift in amplitude, caused by red light activation when green light or not is present, respectively. For $\Delta T^R$ and $\Delta T^G$, the the period displacement when green light is switch on, under red light or no red light conditions, respectively, was similar. It is apparent that the higher the slope of the edges, of the tuneable region, the higher unintended sensitivity of period and amplitude to the light wavelengths, occurs. The results suggest, that an increased rate of production of GFP due to the green light activation, shifts the period from approximately 23 min to 37 min when red light is OFF, while under 100% green light, the period will shift from 50 min to 58 min. This is qualitatively different from the simulations shown in Figure 6.1.4 where the period remained unaffected for varying green light intensity. However the relatively large impact of red light on the amplitude observed in the initial model, now is reduced (section 6.3.1). The above, can also be seen in (c) and (d), where the amplitude and Period are plotted versus the green and red light intensity respectively. Continuous and dashed lines represent initial conditions where the opposite light wavelength is either 100% ( > 0.1 W/m²) ON or OFF respectively.
Concluding on the modifications and properties of the corrected model for the degradation it is worth noting that despite the loss of independent tuning of amplitude, the effect of green and red light still appeared to be serving their intended purpose to a satisfactory degree. As shown in Figure 6.2.3 (b) the slope if the lines bounding the tuneable region of the oscillation reveal that when green light intensity is increased, it mainly affects the amplitude and to a lesser extent the frequency. The opposite effect occurs when red light intensity is varied where it primarily affect the period at a much smaller impact on the peak GFP level.

6.3. System Sensitivity Assessment, Re-Design and Variability Profile of the Model

6.3.1. Amplitude and Period Sensitivity on Light

To investigate further the nature and conditions that affect the impact of green and red light onto the amplitude and period the following metric was introduced:

\[
S_{AR} = \frac{R}{A^G} \frac{dA^G}{dR} + \frac{R}{A^D} \frac{dA^D}{dR}
\]

(6.4)

\[
S_{PG} = \frac{G}{T^G} \frac{dT^G}{dG} + \frac{G}{T^D} \frac{dT^D}{dG}
\]

(6.5)

\[
S_N = \frac{1}{2} (S_{AR} + S_{PG})
\]

(6.6)

Where \(S_{AR}\) is the normalised sensitivity of the Amplitude \(A\) on red light, \(S_{PG}\) the normalised sensitivity of the Period \(T\) on green light and \(S_N\) the combined sensitivity. Indices \(G\), \(R\), \(D\) stand for green, red and dark as the condition of the opposite light wavelength when \(S_{AR}\) and \(S_{PG}\) are calculated. Small sensitivity indicates conditions where the unintended effect of green light on \(T\), and red light on \(A\) is minimised. This allowed to perform a series of numerical simulations exploring the space of chemical inducers as complementary tuning dials and identify under which conditions the sensitivity is minimised.

Figure 6.3.1 shows the results for the specific scans where for each combination of chemical inducers either light wavelength is varied between 0 – 1 W/m² (0-100 %) giving the corresponding tuneable region area in (b). The thick blue bound tuneable region indicates the one with the lowest combined sensitivity \(S_N\) (IPTG 1.25 mM and arabinose at 5 %) while the thick green line indicates the one with the largest area. (IPTG 2.5 mM, 5 L-arabinose). Plotting a heat map for the \(S_N\) against the chemical inducer concentration (c) reveals a low sensitivity range IPTG 0-2 mM and arabinose 2-10 %, and medium sensitivity for values outside these bounds except for arabinose concentrations lower than 1% which return a very high sensitivity. The latter is expanded up to 2% when IPTG is used at high concentrations >4 mM.

It is worth clarifying that the upper limits of simulated chemical inducer concentration were chosen arbitrarily such that they include saturating amounts of them. Normally for any experimental implementation the maximum IPTG concentration used without any observable growth rate drop is just over 1 mM for cells that express the LacY transportase (Wilson et al, 1981; Kosinski et al, 1992; Perfeito et al, 2011) and ~1% for arabinose subject to the genotype of the specific strain. It is worth noting that arabinose even in non
Figure 6.3.1. **(a)** Examples of the time courses simulations obtained for GFP when chemical inducers and light conditions are varied. These are analysed into **(b)**, tuneable region under different concentration of arabinose and IPTG. The numerical simulation results were used to scan the chemical inducer space in order to identify the combination of IPTG and arabinose with the minimum possible dependence degree (referred as combined sensitivity $S_N$) of amplitude and period to red and green light respectively. This low $S_N$ tuneable region is marked with thick blue lines (arabinose 5% and IPTG 1.25 mM) while thick green lines represent the combination which bounds the largest possible tuneable area (arabinose at 2% and IPTG at 5 mM). At very low arabinose concentrations below 1% the dependence degree is maximised as the differential effect of the two light wavelengths seems to be collapsing into a mode where both light wavelengths will affect the system in a similar way. **(c)** Heat map of the dependence degree for varying concentrations of arabinose and IPTG. The lowest sensitivity $S_N$ is observed at concentrations of arabinose higher than 3.5% and IPTG in within the range 0.5-1.5 mM.
cytotoxic concentrations can affect the global cell metabolism because cells growing in LB or in minimal media can switch to metabolise arabinose as the primary carbon source. Even if the experimental setup is based on a microfluidic platform, where a chemostatic environment is assumed, so that the effective arabinose concentration interacting with AraC is constant, the cost for producing the required machinery for metabolising arabinose can affect the behaviour of the system.

6.3.2. System Re-Design for Independent Amplitude and Period Tuning

Looking further into alternative ways to operate the oscillator, in a region where the GFP feedback effect is limited or eliminated, a number of alternative design solutions can be investigated.

The most realistic and feasible tuning dial, would be to reduce the RBS strength upstream the GFP-SsrA construct. This, can decrease the translation rate of GFP, so that the overloading effect on ClpXP due to the reporter is minimised, while the transcriptional regulation is unaffected. Numerical simulations with a 10-fold slower translation rate, for GFP-SsrA, shown in Figure 6.3.2, predicts an almost complete recovery of the period loss, due to amplitude tuning by green light. The period sensitivity $S_{PG}$ remained very low at 0.1 - 0.2 across the IPTG and arabinose concentration range, when the weaker translation rate was used as opposed to $S_{PG}$ 0.3 - 0.5 observed with the default value at 9 min^{-1}. A consequence of this, is the overall decrease of the amplitude, so this solution can be of practical use, only if the required output levels of the oscillator are relatively low. A practical limitation can also be the GFP detection limit, of the equipment that can be used to measure green fluorescence for assessing the system experimentally. However, modern inverted fluorescent microscopes can detect fairly well fluorescent molecules that lie in the region of hundreds per cell or even single molecule level (Hammar et al, 2014).

A second solution for the decoupling of period from amplitude tuning, could be to over-express the ClpXP machinery within the host cell. Since the increase of $S_{PG}$ of the competing model was identified to occur because of the limited amount of ClpXP, a potential re-design, can be the overexpression of ClpXP. The numerical simulations in Figure 6.3.3, showed the predicted sensitivities when different amounts of available proteosome (D) is assumed. The number of proteosome molecules was varied between 300 and 6000 molecules and for each case the heat-maps for $S_{PG}$ , $S_{AR}$ , $S_{N}$ were calculated, within the IPTG and arabinose predefined space. The simulation results revealed two opposing trends. The sensitivity of period to green light ($S_{PG}$) 6.3.3 (b) as expected, is decreased as the number of available D (ClpXP molecules) increased. Maximum $S_{PG}$ , approximately 0.6, was observed at D=300 while for D=3000 the value dropped to its minimum at 0.2- 0.3. The opposite trend was observed for the sensitivity of Amplitude to red light ($S_{AR}$), as its minimum value at 0.4, was reached for low ClpXP numbers (D=700) and increased to its maximum at 0.8-1 for D=6000. In either case, when arabinose concentration fell below 1%, the sensitivities climb as the system falls into a different mode of oscillations. The above, yield a combined sensitivity $S_{N}$ which exhibited a minimum (0.4 - 0.5) at intermediate to low values of D (700-1000). Experimental implementation of this solution, would require further investigation into some important limitations that the model does not take in account.
Figure 6.3.2. Numerical simulations showing the relationship of period on green light. The figure, shows the tuneable region box and Period sensitivity heat map, for the default translation rate constant (left) and a 10 times smaller constant (right). It is clear, that in the case of the weak RBS strength, the upper and lower bounds of the box appear to be more horizontal suggesting that by using green light, the system can be driven from 500 to approximately 1500 molecules, without any deviation on the Period, which remains constant at 54 min and 30 min in the presence or absence of red light, respectively. The corresponding heatmaps, that cover the chemical inducer space, also indicate that regardless of the chemicals concentration the sensitivity, for a low translation rate reporter, remains at 0.1. Exception is the concentration of arabinose below 1 % where the value rises at 0.3. On the contrary, a strong RBS returned sensitivity values between 0.3- 0.5, while for low arabinose, that value was increased up to 1.
Further cloning would have been necessary into an already large multiplasmid network, but more alarming should be the fact that the additional ClpXP amounts, can completely eliminate the oscillation. For example if AraC, expressed due to basal expression at the beginning of each cycle is degraded much faster than it can be produced, that would eliminate the initial burst and the system will not oscillate. Almost inevitably, this will affect the tuneable range of the oscillator, so the sensitivity tuning through ClpXP should be done with respect to the required output peak levels and period. Finally, as explained in chapter 5 section 5.2 the ClpXP plays an important role during the cell division process (Camberg et al, 2011). Overloading effects, were demonstrated to inhibit cell division, resulting to elongated bacterial cells. With the current data, there is not enough information to predict how the cell cycle is affected, in the under-loaded regime (when ClpXP is in excess relative to tagged protein ) in the system.

Finally, the most radical and effective solution, would be to attempt to de-couple completely the reporter from the ClpXP machinery. Removing the SsrA tag, will most likely result to a reporter that degrades much slower (at the rate of dilution due to cell division) than the cycle of the oscillation, resulting to an accumulation of the reporter over time. Hence, it is imperative to have a fast degrading reporter through a mechanism other than ClpXP. To date, the literature does not refer to any inherently unstable reporter in bacteria, which would be ideal. The only option is the use of alternative protease activities like Lon (Gur & Sauer, 2008) or the ubiquitin proteosome pathway (UPP) found in euakaryotes (Shang & Taylor, 2011) (for a yeast implementation of the system). The former, is native while the latter is not. In either case, an alternative functional protease activity in E. coli orthogonal to ClpXP, can potentially provide substantial ground for a complete independent tuning of the peak GFP during cycles, provided that its degradation rate, is fast enough to effectively remove the reporter from the system, before the next cycle starts. A system that targets the core oscillator regulators (AraC, LacI and LacI$^c$) to ClpXP, but GFP to an alternative pathway, can eliminate any coupling through overloading effects. Figure 6.3.4, shows the numerical simulations, when such an alternative degradation pathway was assumed in E. coli. These, confirmed numerically, that the combined sensitivity $S_{n}$ dropped to 0.2. $S_{P_{G}}$ is predicted to fall down to zero since the presence of GFP is no longer interacting with the rest of the network but instead is channelled to a separate proteosome for degradation. $S_{A_{r}}$ seems to be returning a similar sensitivity heat-maps with the ones obtained previously in Figures 6.3.2-3 with intermediate values at 0.5 and 0.7 for low arabinose concentration. In fact the impact of red light on amplitude is almost eliminated in the absence of green light (left vertical bound of the tuneable region in Figure 6.3.4 (a)). Although this solution returned the most ‘orthogonal’ tuning prediction, it is worth noticing that it is the solution with the highest risk in terms of experimental implementation due to the lack of characterisation around the proteosome’s activity (e.g. Lon) in a synthetic biological context . Proteosomes are normally active as large heteromeric complexes complemented with regulating co-factors within a given organism, e.g. Lon might be too slow to degrade effectively the GFP in E.coli depending on where its cognate recognition tag was placed, resulting to an accumulation of the reporter. To overcome these further characterisation is required, something that makes this solution the most costly in terms of resources and time.
6.3.3. On the Robustness of Oscillations - Stochastic Model

The final assessment for the proposed oscillator design addressed the robustness and stability of the oscillations. Up to this point all the numerical simulations shown were based on partially or purely mass-action models solved deterministically using OD15s (stiff/NDF) or Sundials solvers in MATLAB. Although deterministic solutions indicated whether a system can be unstable and oscillating or falling into a fading oscillatory behaviour under certain conditions it does not hold any information about the variability in gene expression. Previous studies have demonstrated that this variability or noise can be distinguished between intrinsic and extrinsic noise (Elowitz et al 2002, Ozbuduk et al 2000). The particular model structure cannot account for cell to cell variability as higher order metabolic and cell cycle processes are not included. However using the Simbiology package allows for the final pure mass-action model to be converted automatically and solved with the SSA stochastic solver which is an implementation of the Gillespie algorithm. Solving stochastically can give an indication on how the intrinsic noise arising from discrete particle interactions can affect the system. There are three main reasons for this model to be assessed stochastically before any experimental implementation is attempted:

- It is imperative to have a model informed indication that before allocate resources and time to implement experimentally such a complicated network will yield distinct and measurable oscillations as opposing to a noisy fluctuating signal. Such fluctuations can potentially mask any oscillatory behaviour.

- Stochastic simulations can indicate which components of the system are responsible for the variability. Identifying those can lead to a refined system that is more robust and exhibits more consistent waveform across different conditions and inducers.

- Although not common there are cases where a stochastic model exhibits a qualitatively different behaviour compared to the deterministic counterpart. In the deterministic regime under the same conditions only one state of the network can be reached while in the stochastic regime the variability can be large such that under the same conditions more than one states can exist. This is the case with Stricker et al 2008 where it was demonstrated that the oscillator in the stochastic regime had a bimodal distribution in the period. One high amplitude long period state and a fast low amplitude one. The deterministic model was of-course unable to reproduce both states.

The model used for all stochastic simulations was the final corrected full mass-action model which includes explicit enzymatic reactions for ClpXP degradation, dilution rate effect on all species and loop formation corrected mechanism. The algorithm for each time point converts all the reaction rates to propensities from where the state of species is going to be determined for the next time point. The log decimation used for data recording was 1000 in order to increase the speed of simulations since the stochastic solutions for such a large model are exceptionally heavy to compute. Custom scripts in MATLAB where used to analyse the simulation generated data. Figure 6.3.5 shows time traces from both the deterministically solved model and the stochastic counterpart with altering light conditions over time. Qualitatively at this point the stochastic model behaved similarly to the deterministically solved time trace.
Chapter 6. A Light Tuneable Oscillator

(a) Period versus Amplitude

(b) Period Sensitivity on Green Light $S_{PG}$

(c) Amplitude Sensitivity on Red Light $S_{AR}$

(d) Combined Sensitivity $S_{N}$

\[ D = 300 \]

\[ D = 500 \]

\[ D = 700 \]

\[ D = 1000 \]

\[ D = 1500 \]

\[ D = 2000 \]

\[ D = 3000 \]

\[ D = 6000 \]

Amplitude (molecules)$^{14}$
Figure 6.3.3. Numerical simulations for tuneable region shifting and sensitivity profile for Amplitude (S_{AR}) and Period (S_{PG}) on light, for varying proteosome number D. For each value of D used the (a) tuneable region bounds for different chemical inducer concentration is shown and blue thick lines mark the chemical condition with the lowest combined sensitivity. The corresponding heat-maps of (b) S_{PG}, (c) S_{AR} and combined sensitivity S_{N} (d) is shown. It is clear from the heat-maps that for an increasing number of D the sensitivity of the period to green light tends to decrease from approximately 0.6 to 0.2 while the opposite is observed for the sensitivity of the amplitude on red light which increases from approximately 0.4 to 0.9. The combined optimal sensitivity value is found when D = 700 giving a S_{N} = 0.4. This number of ClpXP is the one that is more likely to provide the conditions for tuning the amplitude by green light with a minimal effect on the period and similarly period tuning by red light with minimal impact on the amplitude.

Figure 6.3.4. Numerical simulations of the light tuneable oscillator with the dual proteosome pathway. In this version of the model the regulators LacI and AraC are targeted to the ClpXP machinery while GFP is tagged so it can be recognised and cleaved by the Lon proteosome. In this case the GFP will not compete and alter the degradation rate of the regulators resulting to a more orthogonal looking tuneable region shown in (a). The period in this case can be tuned between 35 to 55 min (IPTG 1.11 mM, Arabinose 5.56%) and the amplitude from 1000 to 80000 molecules with almost no impact on the period. The only unintended sensitivity observed was an amplitude shift by red light when green light is also present. The corresponding heat maps show that throughout the chemical inducer space S_{PG} is maintained at 0 (d) while S_{AR} (b) between 0.4-0.6, and a combined sensitivity of 0.2 (c) (with the exception of low arabinose concentrations) where again it shows a relatively sharp rise.
Chapter 6. A Light Tuneable Oscillator

However, it is obvious that some cycles fell short in terms of their period and peak GFP levels. In order to analyse a longer and repeated signal for each condition, an ensemble run of 20 simulations from 0 to 1000 minutes was ran. An example of visualised ensemble simulations signal, is shown in Figure 6.3.5. Each run is not in phase with the others hence averaging the signal will yield not meaningful results. Instead for each trace the peak level of GFP was recorded, along with the peak to peak distance (time) and the resulting matrices can be used to plot the histograms showing the amplitude and period distribution across different conditions.

Figures 6.3.6-6.3.8 show these histograms and corresponding box plots, for varying chemical inducers under the four extreme light conditions (dark, green, red and both light wavelengths). It is worth noting that the mean value and standard deviation of periods and amplitudes, is also not a reliable metric, in cases where the distributions, were not normal or even bimodal, something that made the comparison of the results by these parameters non meaningful. Nevertheless, some conclusions can be drawn from the results shown in Figures 6.3.6-6.3.8.

The stochastic model still exhibits clear oscillations that are not masked by the noise. Light conditions are shown to be affecting the system towards the same direction shown in the deterministic regime i.e. green light increased primarily the amplitude while red light increased the period. The magnitude of the amplitude period shifts also seemed to be in good agreement between determinist (red star markers) and stochastic regime (green circles) as shown in the histogram Figures 6.3.6-7. The only cases where significant differences occur is when the distributions are highly abnormal or bimodal (red dashed squares). Conditions that yield bimodal oscillations (e.g. IPTG 2-5 mM, arabinose 5% in the absence of red light) had the deterministic value lying exactly in between the two peaks of the stochastic distribution. Most distributions appear to be upwardly skewed, suggesting that the system under certain conditions is prone to be randomly generating high peak high amplitude cycles. The peak GFP levels exhibited higher variability (relative to the period) especially when GFP expression is enhanced by green light. The condition generating the least consistent periods was the absence of red light. In fact in the absence of red light and presence of high concentration of chemical inducers, highly abnormal distributions were observed. For 2-5 mM IPTG and 5% arabinose with green light only present, bimodal oscillations occurred. Similar behaviour was also reported in the stochastic model of the original dual feedback oscillator (Stricker et al 2008) but due to the fact that the LTO model was modified and included additional components there was no way to compare these observations. Strangely enough in the LTO case the amplitude had a wide distribution under 0% arabinose in the presence of green light (Figure 6.3.7) opposed to the period where low arabinose resulted to a much tighter distribution compared to the abnormal histograms (red dashed squares) occurring at 2 and 5% (Figure 6.3.6).

The fact that the wider distributions occurred in the absence of red light indicated that a possible source for this variability could be tracked to the PLac/OmpR controlled lacI gene. When red light is absent this promoter is activated by OmpR-p and the LacI output enhances the secondary feedback loop that affects all active promoters of the network. Hence if this gene suffers from noisy expression this will generate nosier oscillations especially when green light is present where high amplitude long periods occur.
Figure 6.3.5. Stochastic simulations generated the same qualitative behaviour with the deterministic counterpart. Red and green light shifted the oscillation to higher periods and amplitudes respectively. Ensemble simulations for each condition were used to analyse the system further. Each stripe represents one time course run while the colour represents the GFP level (blue to red, 0 to 5*10^4 molecules respectively). The averaged trajectories show clearly that the cycles for each run are out of phase from an early time point hence obtaining the period and amplitude distributions is a way to analyse the full signal in a more informative way.
Period Distribution For Varying Light and Chemical Inputs

Figure 6.3.6. Period distributions of the stochastic light tuneable oscillator for varying light and chemical inducer condition. Each condition was simulated for 5000 minutes and the trace was analysed using custom MATLAB scripts to calculate the occurring periods and amplitudes. The range of the histograms cover periods from 0 to 100 minutes distributed in 20 bins (5 min intervals). Green thick circles on the x axis mark the most populated bin of the stochastic model while the red star (*) marker shows the value obtained from the deterministic simulations of the corresponding model under each condition. Most of the periods appear to be normally distributed with low standard deviation (SD) values for conditions giving short periods (e.g. Dark/ IPTG 2mM/ Arabinose 0%) and larger SD for longer periods (e.g. red & green /IPTG 5 mM/ Arabinose 0%). However under certain conditions the shape changes to an upwardly skewed distribution suggesting that the system is prone on stochastic effects that generate bursts of high period cycles. The most striking conditions are circled in red dashed lines. In this cases a high degree of variability was observed resulting to highly abnormal or even bimodal distributions. More specifically under dark, the periods are upwardly skewed for high concentration of inducers while when green light was ON a bimodal distribution with peaks at approximately 30 and 65 min was observed (e.g. green/IPTG 5mM/Arabinose 5%). Last it is worth noting that these extreme cases where the ones that the deterministically calculated periods deviated relative to the stochastically determined dominant period.
Figure 6.3.7. Amplitude distributions calculated from stochastic simulations. Each condition was simulated for 500 min and the traced was analysed to calculate periods and amplitudes. The range covered was 0-10000 molecules sorted in bins of 500 for conditions absent green light while in the presence of green light the range covered was 0-70000 molecules in bins of 5000 molecules. Compared to the period the amplitude is distributed much tighter. The deterministic (red star marker) and stochastic dominant values (green circle) under most conditions showed no significant deviation (± 1 bin). The most abnormal upwardly skewed distributions (bound by red dashed line) were again observed in the presence of green light, 0% arabinose conditions generated the largest variability. This is in contrast with the observation the period variability was higher when arabinose was used at high concentrations.
Chapter 6. A Light Tuneable Oscillator

Figure 6.3.8. Stochastic simulations for the light tuneable oscillator model. For each condition a time trace of the GFP levels was obtained for 5000 min. MATLAB custom scripts were used to identify the peaks and valleys of the trace and calculate the period (time between consecutive peaks) and amplitude (Peak GFP minus the valley GFP level). The results are shown as boxplots with the red line representing the median, upper and lower bound of the bar representing the 25% first and 75% third quartile of the distribution and the whiskers mark the maximum and minimum value. Red marks show the outliers. Overall the same trends for varying chemical or light conditions were confirmed with the stochastic model. Green light showed to be able to increase the amplitude while red light shifted primarily the period but with either wavelength subject to the limitations and sensitivities discussed in section 6.3.1. Overall for the scanned range of inducers 0-5 mM for IPTG and 0-5% for Arabinose the period increased subject again to the light conditions. The highest median for period reached was 70 min for 5 mM IPTG and both red and green light on while absence of either chemical or light inducer generates the shorter cycles. The highest amplitudes observed at 50K molecules was in the presence of 5 mM IPTG under both light wavelengths but in the absence of arabinose. The lowest variability was observed in the absence of green light and chemicals. A first sight on this results showed that the variability increases in the absence of red light however more meaningful conclusions can be drawn by the actual distributions shown in Figures 6.3.6-6.3.7.
Previous studies on stochastic gene expression (McAdams & Arkin 1997, Ellowitz et al 2002, Ozbudak et al 2002) demonstrated how molecules found in very low numbers within the cell, suffer the heaviest stochastic effects. In addition they showed how translation efficiency is correlated with phenotypic noise. Based on these, two possible modifications were identified in order to re-design the theoretical system for decreasing the noise and generating more consistent oscillations. The first was the copy number of the lacI gene which was set to 3 because of the low copy number origin of replication (pSC101) that this gene was proposed to be expressed from. The second possible re-design target was to decrease the translational efficiency by decreasing the RBS strength in front of this gene. However this change of RBS by itself would require increase in the promoter strength in order to maintain the required LacI expression rate, something that requires more time and resources to achieve. Instead changing the copy number of the gene by cloning it into a different plasmid and reducing the RBS strength is a much simpler experimental process.

The model was re-simulated for an increased gene copy number for LacI to match that of a p15A vector (approximately 20). The corresponding histograms and boxplots are shown in Figures 6.3.9-11. In these figures only 4 different chemical conditions are shown and were chosen such that they include the ones that generated the abnormal and bimodal distributions shown in Figures 6.3.6-7. The brown shaded area of the figures shows the results obtained from the model where a higher copy number for lacI was assumed while no shaded graphs correspond to the low copy number model. Interestingly the results showed a radical decrease in the variability of periods and amplitudes measured in the absence of red light. The period distributions under no light conditions were still upwardly skewed however the high dominant period frequency indicates more consistent period compared to the low copy number model with the only exception the 0% arabinose condition. In the case of green light only present, the previously observed wide or bimodal distributions were completely reduced to a sharp unimodal distribution ranging from 30 - 50 minutes indicating that this setup is more likely to yield clear and consistent oscillations when implemented experimentally. However a 3-fold decrease in the amplitude was observed in the absence of red light (Figures 6.3.10-11) which can be a contributing factor to more consistent periods. Nevertheless, for similarly lower amplitudes in the first low copy version of the model, the spread remained high suggesting that increasing the copy number was the primary reason for tightening these distributions. A final observation with the higher copy number model was that the sensitivity of period on IPTG when arabinose is present was significantly reduced. This means that red light is the primary dial for tuning the frequency. This is more likely due to the fact that the ratio of LacI to wild type LacI is increased. The former cannot bind IPTG hence this inducer affects a much smaller portion of the repressor pool. In either case the combined effect of the two through the negative feedback loops was still sufficient to drive the oscillation.

With this results the proposed plasmid map can be modified so that a two plasmid system is implemented where the lacI gene under PLac/OmpR is cloned in pJS169-TCS. The overall stochastic simulations increased the possibility that an experimentally implemented system will exhibit oscillatory behaviour while maintaining the light tuneable trends. The large complexity of the model and uncertainty in some parameters on top of confounding factors that are not included are preventing for confident quantitative predictions however any qualitatively working LTO is predicted feasible with reasonable size genetic constructs and gene numbers.
Chapter 6. A Light Tuneable Oscillator

Figure 6.3.9. Period distributions of the stochastic light tuneable oscillator for the low copy number pSC101 vector for expressing lacI² (white area) and the medium copy number origin p15A (brown shaded area). The histograms cover the range of period from 0 to 100 minutes and green circles indicate the bin with the most frequently occurring periods. While under conditions where red light is present the distribution spread remained unaffected, significant differences occurred in the absence of it. Under dark conditions the distributions remained skewed to higher periods but with a sharp peak around the dominant period. With the exception of 0% arabinose / Dark the distributions appeared to be tighter. Even more clear were the results in the presence of green light where in this case the bimodal and wide distributions observed in the low copy number model are collapsing into a distinct unimodal and more gaussian shaped distribution.
Figure 6.3.10. Period distributions of the stochastic light tuneable oscillator for varying light and chemical inducer condition. The range of the histograms cover periods from 0 to 100 minutes. Green thick circles indicate the most populated bin. Similarly to the corresponding period distributions no observable differences are produced from the two versions of the model when red light is ON. However a significant drop in the amplitude (approximately down to 33% of the low copy model) was observed. This can be a contributing factor to the apparent improvement in terms of period and amplitude variability.
Figure 6.3.11. Boxplots showing the comparison between the low (white shaded) and medium copy (yellow shaded) versions of the model regarding the lacI\textsuperscript{a} gene. The boxplots reveal that the median values for the medium copy number for both period and amplitude are shifted downwards compared to the low copy, and in particular in the absence of red light the shift is approximately 3-fold. In addition the medium copy version was shown to be less sensitive to changing IPTG concentrations. e.g. in both green and red/green light conditions with 5% arabinose, when IPTG changes from 0 to 2 mM there is a 10 minute increase in the period and 5000 molecules in the amplitude for the low copy number version while the medium copy shows no (green) or minimal shift (red/green).
In this chapter, I discuss the findings of this work, in a general context, followed by more specific topics for each chapter. Finally, the implications and future directions for the developed circuits and models are discussed.
Chapter 7

Conclusions and Discussion

7.1. Conclusion and Discussion on the Scope of the Project

Overall, this project has demonstrated how starting from the understanding, modelling and characterisation of small circuits, one can scale up to the design and implementation of larger and more complex systems, by rationally coupling components of a predefined function. More specifically, the characterisation and modelling framework developed for the Cph8 and CcaR TCSs, was followed by the implementation of a novel way to connect them to other circuits, demonstrated with the toggle switch and the oscillator. Alongside, the model was used to make informed qualitative predictions of how the individual parts would behave in a more complex implementation.

The double input promoter Plac/ccaR, although weak, was shown to be responsive to both green light and IPTG, while Plac/ompR was shown to be a reliable component for translating IPTG and red light into a single output strong enough to drive the light tuneable toggle (LTS) switch between the two hypothesised stable equilibrium points. It also resulted in a high GFP output mode in the presence of IPTG.

For the LTS, it was shown that overexpressing proteins that are SsrA tagged can result in growth inhibitory effects. The data supported the hypothesis that this mechanism was due to the overloading of ClpXP, which can affect the normal cell division cycle. Although the system was not driven from the low state to the high state solely with the use of light, the latter, without any aTc was enough to switch it from the high state to the low state. This behaviour is analogous to a memory component, as when induced to the high levels, the only way for the system to reset to the low state is by shining red light.

The oscillator was the largest and most complicated network that was proposed. Although it remained at the modelling stage of the forward engineering cycle, this work yielded some important insights into how the period and the amplitude of the dual feedback oscillator can be decoupled to some extent and controlled independently by the light sensory TCSs via the dual input promoters. The model was used to analyse how each wavelength will affect the system, the impact that tuning specific dials (e.g. RBS strength and degradation rates) will have, to propose a specific plasmid architecture and finally to assess through stochastic simulations, factors that affect the robustness of the system. This work has provided the necessary theoretical background for an experimental implementation of the light tuneable oscillator (LTO).

Finally, the overall project succeeded on delivering theoretical and experimental systems while staying within a forward engineering framework. Each chapter was structured as an independent engineering cycle, but building on and contributing to the overall scope of the project with respect to most of the engineering principles discussed in the introduction. 'Decoupling' and 'Modularity' were evident in the way that each chapter can be viewed as standalone synthetic circuit and used as such, or combined to build a larger network.
'Characterisation' and 'Abstraction' was demonstrated by the fact the for the dual input promoters, the only information actually needed from the TCS characterisation was a minimum model capturing the transfer function of light to OmpR-p levels. Similarly, the only information needed to be passed from the dual input promoter level to the LTO or LTS was the combined response function of the reporter to IPTG and light.

Having stated the above, the challenges and factors that affected negatively the completion time and resolution of the data collected were numerous. On the theoretical level, the poor level of understanding of living organisms, due to their complexity, in addition to the limited characterisation and the exact parameter values of the components used, unfortunately left many open questions on the validity of the derived models and their predictive power. The identifiability question (MacDonald et al, 2013), along with the limited data for fitting multidimensional models, allowed for only semi-quantitative or qualitative analysis in many cases. On an experimental level, the cloning process was the most time and resource consuming task. For the characterisation aspects of this work, the fact that suitable equipment (e.g. microfluidics) for live cell experiments under chemostatic conditions, or an automated programmable light induction and measurement platform was not available, has imposed significant limitations on the time that a system could be measured as well as generating large variability in the measurements. Last, unexpected confounding factors like pL-Cph8FD instability, the effect of crosstalk between the LTS and the native LacI and the ClpXP overloading in the prototype implementation of the LTS, were only a few examples of what can affect the rational design and implementation of the circuits attempted in this work for realising synthetic genetic networks that are controlled by light.

7.2. Results Discussion for Each Module

7.2.1. Module 1: Light Receptor Characterisation

Theoretical Work

An in silico network was reconstructed from the current literature information and understanding of two component systems (TCS). More specifically, this network was detailed enough to include the genetic expression of the genes that convert Heme to the photoactive moiety (chromophore) PCB that binds and dictates the phosphorylation activity of the TCS HK domain and the hierarchical binding and regulation of the PompC promoter by OmpR-p. The network was expanded during the modelling process to include auxiliary phosphorylation mechanisms along with the apophytochrome activity when PCB is not present.

From the kinetic scheme of the network’s reactions, a non-linear ODE model was derived from first principles according to the law of Mass-Action (MA). This high dimensional, complex mechanistic model was rigorously reduced from 25 ODEs to 2. The reduction process started with an initial abstraction of the system where three modules, PCB production, TCS phosphorylation activity and PompC promoter regulation, were reduced independently. The impact of some of the assumptions used for each module was tested individually or
collectively when all three where connected. Imposing strict tolerance on how the reduced model could deviate from the corresponding steady state levels of the full mass-action model resulted in rejecting some of these assumptions.

Rejected assumptions: More specifically, a deviation at the final steady state between the reduced model and MA model was caused by assumptions A(3.5) and A(3.14), regarding biliverdin (BV) being at negligible amounts and free PCB pool being too large for the Cph8:PCB complex formation to affect it, respectively. This deviation was larger when the expression rates of Ho1 and PcyA were reduced or Cph8 was increased, imposing a limitation on the range that the system can be redesigned and tuned based on this model. When the assumptions were rejected, the resulting quadratic expressions for BV and PCB were solved for their real roots and their use eliminated the error between the computation of the specific species with the two models.

Similarly, assumption A(3.9) that Cph8:OmpR complex levels will be much smaller than the available OmpR-p pool, limited the range that the model can reproduce the same steady states as the full MA model when Cph8 expression was increased or when large copy number promoters were used. This was solved by the rejection of this assumption and using the real solution of the resulting quadratic steady state expression for the unphosphorylated OmpR.

A(3.11) and A(3.12) imposed a very narrow range of [OmpR-p] in order for them to be valid and neglected degradation of DNA bound OmpR-p, respectively. Although A(3.11) can be supported with experimental data for the native TCS (EnvZ) (Cai & Inouye, 2002), there is no evidence for the same in Cph8 since this is lacking the phosphatase activity (PA). Although A(3.12) is an approximation that can been used for simplifying the reduced models where promoters do not have multiple operators, in the case of the PompC promoter it causes high deviation. This is due to the fact that the 6 OmpR-p binding sites can act as a sink (as discussed in Chapter 3.1), resulting in a falsely large pool of non-decaying OmpR-p molecules bound on them. In fact, DNA bound TFs are diluted exactly like any other intracellular component. The rejection of these assumptions yielded a very large 6th order steady state expression for the PompC promoter. However it was accurate compared to the full MA model.

The initial assumption that OmpR-p phosphorylation was only via the holophytochrome was corrected by the introduction of a free Cph8 (apophytochrome) and OmpR complex formation and phosphorylation reactions. Finally, to correct for the LacZ basal expression, an auxiliary phosphorylation reaction for OmpR was added (Siryporn & Goulian, 2008; Groban et al., 2009) in addition to transcription activity for non-fully occupied states of the promoter at a lower rate than the fully occupied.

Retroactivity: The model provides insights about how the use of multiple copies of a promoter with multiple TF binding sites such as PompC, can be a heavy load on the available pool of free floating OmpR-p. This retroactivity effect successfully captured by the model, imposes limits on the total copy number of the promoter that can be used to connect other downstream circuits controlled by the same regulator. The model was generalised so one can use the same reduced ODE for simulating multiple genes controlled by the same
Chapter 7. Conclusions and Discussion

promoter. An analytical expression, for the optimal value of copy number as a function of TF abundance, was too complicated to be derived. However, the graphically derived results, suggested that the total PompC needs to be approximately at a ratio of 1:6 to ompR-p, in order to yield the maximum output levels. Lower copy numbers reduced the total output simply because the TF pool had the capacity to activate more promoters, but higher copy numbers will reduce the output since the available TF is not sufficient to fully occupy all the promoters on which it is evenly distributed (in a deterministic scenario).

Parameters and Data fitting: The identification of parameters based on information available in the literature was a challenging task. The EnvZ/OmpR system is a relatively well studied TCS, however published parameter values could potentially not apply to the re-engineered receptor where the sensory domain was replaced by Cph1. In addition to the latter, the Cph8 lacked the PA domain of EnvZ. In order to compensate for the uncertainty in some of these parameters, the initial condition for heme concentration was determined experimentally by employing a heme quantification method. The non-dimensional parameter $z$, was fitted on the transfer function of light to LacZ output, followed by the estimation of the LacZ expression rate $(T_{Lz})$, Cph8 expression rate $(T_{r})$ and phosphorylation catalysis $(co)$. By fitting these parameters, the model was able to approximate the time course assay of Cph8, while it was still able to account for the retroactivity that yielded lower LacZ levels for the reassembled version of Cph8.

Multichromatic system model expansion: The modelling framework and reduction strategy was adapted and used for the CcaS/CcaR system, but also the combination of the two TCSs. For this, a set of additional reactions that describe the λCI inverter controlled by PompC and regulating a PA driven LacZ was used. The qualitative behaviour of the model was able to capture the behaviour of the system published by Tabor et al., 2011. However, due to the fact that CcaS was not as well studied as EnvZ, the available information and parameter values in the literature was limited. For this reason, a quantitative model via parameter and data fitting was not possible for the specific system. For the simulations shown, the parameter values used were assumed to be in a similar range to the corresponding ones for the Cph8 system, since the two TCS share many features e.g. same chromophore. Of course the nature of this model, which even in its reduced form carries the parameters from the initial kinetic scheme without losing their physical meaning, allows for its use and evaluation as soon as more information becomes available in the literature.

Experimental Work

The general strategy followed for the characterisation experiments was to freshly transform cells, pick individual colonies that were grown overnight and initiate an assay after appropriate dilutions during the next day. The light induction for all the experiments for the TCSs was carried out by using custom built LED devices as shown in Methods 2.3 and discussed in Chapter 3.2. These devices were shown to follow the inverse square law fairly well. Therefore, the light intensities applied to a given sample were calculated on this basis. The intensity was adjusted by the use of ND filters or changing the LED-sample distance in the incubator. It is worth noting that the propagated error from the light versus distance function and the distance between samples and LED, was not negligible. However, the error during the Miller assays where light was
kept constant was significantly greater as shown in Figure 3.2.5. This is due both to human error since the specific assay requires more steps, compared to a fluorometric assay, and also due to noise in the actual signal transduction pathway.

A general point that needs to be stressed for all light induction assays across chapters 3, 4 and 5, was the fact that the best fold difference between light or dark induced samples was obtained when the samples were tested 1 or 2 days after fresh transformations. The second important precondition was good aeration of the cultures. Hence, when 14 ml tubes were used, the lid was not tightly sealed and when microplate assays were prepared the transparent film used to prevent evaporation or aluminium foil used to provide a light shield was pierced at multiple points to allow good aeration. The reasons for this requirement are assumed to be the shift in metabolism for anaerobic growth is such that it renders the light receptors inactive and also the oxygen requirements for the reduction of heme to BV and PCB and heme homeostasis (Ishida & Hino, 1972).

Characterisation: In this work, it was shown that the Cph8 TCS was reassembled from BioBricks™ into a single plasmid that harbours all the required components for the system to work in *E. coli* *hao1, pcyA* and *cph8* where placed under constitutive expression as opposed to the dual plasmid system where TetR, and LacI/AraC regulated promoters were used (Levskaya *et al*, 2005). The reassembled system pL-Cph8FD exhibited a ~3-fold decrease in eCFP fluorescence of the peak emission wavelength (477 nm) shown in Figure 3.2.3. for red light illuminated samples, while the corresponding decrease in LacZ activity assay, was only 2-Fold.

Compared to the originally published dual plasmid system (Levskaya *et al*, 2005), pL-Cph8FD transformed cells exhibited half the maximum LacZ activity (~600 MU compared to 1250 for the dual plasmid). This, as discussed in detail in chapter 3, was assumed to be more likely the result of the heavier load, *i.e.* retroactivity, that the high copy number of the PompC:eCFP casette on pL-Cph8FD imposed on the available pool of OmpR-p. Sequestering of OmpR-p to the additional DNA binding sites might be the primary cause for the drop in LacZ from the genomic PompC. Another possible reason can be lower expression rates for the light receptor components, if the constitutive promoters used were of significantly lower strength than the ones in the original system. Unfortunately, not enough data were available in the literature to compare the strength between these promoters. Finally, the controls harbouring only the pL-PCB(A) exhibited no LacZ activity while the ones harbouring just the pCph8 were constitutively ON at 900 MU a value lower than the full system. Possible reasons for this were already discussed specifically for this assay in Chapter 3.

The transfer function of light to LacZ expression shown in Figure 3.2.4 was fitted to a Hill type function for repressor (where light intensity was used as such). The ~0.1 W/m² intensity needed to switch OFF lacZ expression approximated the published values (Levskaya *et al*, 2005; Tabor *et al*, 2011; Olson *et al*, 2014).

The dynamic characterisation showed the transient behaviour of the system when it is switched from low to high or the opposite. The calculated rise and decent times were 2 and 2.5 hours. This response approximated the published values (Tabor *et al*, 2011). These results were used further for fitting the relevant model.

Finally, a static assay was performed by using the multichromatic control system to decipher if green light can
affect the specific system. It is worth noting that the plasmids used are the unmodified ones obtained from Tabor et al, 2011. The system responded qualitatively according to the published figures. In particular, green light increased the LacZ activity by 10-fold. This was in fact a higher fold difference than the published value of 1.5-fold. On the other hand, Davidson et al, (2013) and Olson et al, (2014) showed higher fold differences using the same TCS, from 5 to 8-fold when a GFP reporter was used instead of LacZ. Nevertheless, it is possible that the observed high fold difference is the result of error in the measurement of samples incubated in the dark or simply an artefact of the correction applied (where the baseline in LacZ activity in controls is subtracted from the samples). If the baseline was overestimated by error in measurements, this can result in corrected values for the OFF samples close to zero, or even negative (e.g. pL-PCB in 3.2.4 (a)). The other interesting fact that arose from the specific assay, was that when pJT106b was also present (plasmid that expressed the Red light the PompC:CI construct and Pλ:LacZ) and green light was used the LacZ maximum level rose to values 3-fold higher than before. Unfortunately, this is a strong indication that there is crosstalk between the two TCS in terms of induction wavelength overlapping. The control strains performed as expected, being not sensitive to light, when only pJT106b was expressed and exhibiting no LacZ activity when only pL-PCB(S) was present. Although the tested plasmids appeared to be generating an output that is dictated by green light, the observed inconsistencies in the fold difference and the absolute values measured, with the published data from other authors, strongly suggested that these results have to be interpreted with caution. In addition, the available information on parameters regarding this system, or even necessary information like the Pcpc-G2 promoter structure and number of CcaR-p binding sites did not allow for further data fitting to the model leaving it as a semi quantitative representation.

Module 2: Dual Input Promoters

Given the limited numbers of well characterised TFs that have been used in synthetic circuits (shown in Table 4.1.1.) in addition to all the confounding factors discussed in Chapter 1.2 (burden, crosstalk, retroactivity) that they become more evident as the complexity and size of a circuit increases, this work aimed at implementing an efficient way to couple the TCSs with other circuits. The result was a newly proposed ‘device’ (Endy, 2005) that can have a modulated output with respect to both light and IPTG. In this work it was possible to demonstrate experimentally that the combination of OmpR-p and CcaR-p regulated promoters with LacI operators can work towards modulating the transcription rate of a hybrid promoter by IPTG and light. The architecture proposed for the Plac/ara-1 promoter from Lutz & Bujard (1997), was applied to derive the Plac/ompR. In addition, the promoter sequence PcpcG-2 was cloned into the Plac/ara-1 in order to yield a green light and IPTG dual input promoter.

Theoretical Work

Due to the complexity of the model derived in Chapter 3, in order to include the LacI regulation into the PompC regulation, a compact model was derived. This abstract model included only the absolutely necessary biochemical reactions in order to reproduce the observed light response. The variables modelled were
OmpR, OmpR-p, Cph8Pr/Pfr, and the promoter occupancy states. Some parameter values were informed from the model in Chapter 3; however, due to the structural differences, this model had to be refitted to reproduce the time course data for Cph8. When the LacI regulation was introduced, the model was able to approximate qualitatively the features exhibited by the experimental implementation of Plac/ompR and Plac/ccaR. However, since a mapping between GFP molecules and observed fluorescence was not available, the model remained semi-quantitative.

**Experimental Work**

The assembly process produced a plasmid that harbours Plac/ompR or Plac/ccaR controlling the transcription of a GFPmut3b and with the genes for PCB production plasmids under constitutive expression. This plasmid was compatible with the pCph8 plasmid. While Plac/ompR was responsive to IPTG even without the light receptor in DH5α, Plac/ccaR appeared to have lost all activity when CcaS was not expressed, more likely due to the deletion of the AraC and absence of c-AMP/CRP sites. FACscan measurements of the Plac/ompR, showed a ~10-fold induction between samples grown in the light and dark, but as a IPTG concentration increased, the GFP expression in samples treated with red light rose 5-fold compared to lower IPTG concentrations, resting at 2-fold lower than the samples grown in the dark. For the samples in the dark, [IPTG] induced a 1.5-fold increase in fluorescence. While Plac/ccaR was also light sensitive showing 6 to 4-fold induction between samples grown in the dark or in the presence of green light, increasing IPTG concentration showed a similar 1.5-fold rise in fluorescence for illuminated samples and a 2-fold induction for samples grown in the dark. It is worth pointing out that the maximum fluorescence is 10-fold higher for Plac/ompR compared to Plac/ccaR.

A 96-well microplate based assay was used to produce the two dimensional induction profile of the two promoters. The specific assay using the light induction incubator (Dr. Eric Davidson), yielded better IPTG fold induction (5-fold) for both promoters while the light receptor activity contributed an additional 2-2.5-fold increase towards the maximum expression rates. IPTG addition under red light conditions drove the expression levels from basal to approximately 1/3 of the maximum level observed. This suggested that even under non-activating light conditions, both dual input promoters exhibit non-negligible activity. Since the only active RNA polymerase recruitment -10 and -35 sites were the ones originating from PompC and Ppcp-G2, it is reasonable to assume that similar activity is present under non-activating light conditions (red and dark respectively), in the light sensing constructs assayed in Chapter 3. The difference in the case of the dual input promoters is the fact that this activity is decreased when LacI is able to bind the promoter and block transcription (in the absence of IPTG). This can be a possible explanation why the PompC or Ppcp-G2 yield a relatively small fold induction in the previously published data (Levskaya et al, 2005; Tabor et al, 2011) and in some assays shown in Chapter 3. Furthermore, based on the fact that Plac/ccaR transformed in *E. coli* DH5α was not active in the absence of CcaS / CcaR and presence of IPTG it can be hypothesised that the main contributing factor for the non-light induced activity of the Ppcp-G2 promoters is the basal phosphorylation rate of CcaR-p. Similarly for Plac/ompR, the non-dark induced activity can be the effect of basal phosphorylation of OmpR-p. However, OmpR-p independent transcriptional activity cannot be ruled
out as a main contributor at this point, since this construct was not tested in a ΔompR strain. Furthermore, supporting the high basal expression of PompC and consequently Plac/ompR under red light conditions, is the comparison of the surface plot for Plac/ompR (Figure 4.3.4 (a)) with the corresponding assay for the LTS (Chapter 5, Figure 5.3.8). In the LTS case the overexpression of LacI from a high copy number plasmid drives the basal levels even lower giving a relatively higher dark to red light fold difference.

Finally, the transient behaviour of the system showed similar characteristics with the native PompC and Ppcp-G2 promoters (Olson et al., 2014), with a rise time of ~2 hours but significantly slower decent time of over 3 hours compared to the results obtained in Chapter 3 for the PompC dynamics. This can be the result of the reporter being expressed from a higher copy number (p15A replicon) compared to the genomic PompC used to express LacZ. Higher copy should have produced relatively high amounts of GFP, that need longer time to be removed by dilution. An additional reason is the longer half-life for GFP mRNA compared to LacZ mRNA, 6 min (Megerle et al, 2008) and 3 min (Dong & Kurland, 1995), respectively.

7.2.2. Module 3: Light Tuneable Toggle Switch

This module aimed to use the light receptor module Cph8/OmpR to regulate one of the first and most discussed synthetic genetic circuits, the toggle switch. With this the dual input promoter shown in chapter 4 was shown as a fairly effective connector between the TCSs discussed in chapter 3 with existing circuits that involve LacI in their regulatory network.

Theoretical Work

The idea of using light to modulate the behaviour of a bistable genetic toggle switch was initially tested using a phenomenological model. This model comprised of a Hill function to account for the effect of red light combined with the Hill type function of the original model published by Gardner et al., (2000). The phase plane analysis showed that light shifted the phase plane curves, from a bistable (2 stable equilibrium points) state to a monostable state where only one stable equilibrium point exists. Light shifted the transfer function of IPTG to the two regulator levels so that the switching point occurs at lower IPTG concentration, while bifurcation analysis demonstrated how the hysteresis region of a bistable system can be reduced or collapse the system into a monotonic function. Finally, high leaky expression rates for the two regulators, resulted in a more narrow bistable region.

The phenomenological model was followed by another mechanistic model, which shared the same kinetic scheme with the one showed in Chapter 4, but with the addition of tetR under control of the Plac/ompR and placing LacI under the control of a PtetO1. This ODE model generated the same trends as the phenomenological model, i.e. the direction that light was pushing the toggle switch. However, some qualitative differences were observed regarding the steepness of the IPTG transfer function under light or dark conditions.

Due to the lack of relevant data, the parameters of this model were based on the literature and parameters that were obtained from a non-light inducible version of the toggle switch by Dr James Arpino fitted by
Dr Mariano Beguerisse-Diaz (unpublished data, personal communication). In addition, for a qualitative comparison with the experimental data the maximum and minimum observed expression levels were used to tune the transcription rates of the model in order to relate modelled GFP molecules to fluorescence AU at a ratio of 1:10. The latter was an arbitrarily chosen factor, it holds no quantitative information and served only as a semi-quantitative analysis for explaining some of the collected data.

**Experimental Work**

The experimental work succeeded in cloning the Plac/ompR promoter upstream the *tetR* gene in the pRG021 plasmid kindly provided by Dr James Arpino and integrating it with the Cph8 light receptor and PCB production genes in a two plasmid system. However, this particular construct was shown to be unusually unstable, suffering from heavy burden effects that resulted in inhibition of cell division and overall growth of the cells. The dramatic pathological phenotype resulted in normal forming colonies, the cells under the microscope appeared elongated and unable to divide. This phenotype, was found to be very similar to that observed by Camberg *et al.*, (2011) in cells where the ClpXP proteosome activity was disrupted by a series of gene knockouts in *E. coli*, something that was shown by the authors to be inhibitory for the normal cell division. In this version of the toggle switch, this phenotype occurred when aTc induced the expression of LacI that had SsrA tags fused at the C-terminus of the protein with a strong RBS (20 477 AU). Based on this and what was also reported on how ClpXP can be congested by recombinant TFs (Cookson *et al.*, 2011), it is strongly suggested that cell division was disrupted because of the overexpression of SsrA tagged LacI. Another contributing factor can be protein misfolding and aggregation of LacI:GFP fusions previously discussed by (Gordon *et al.*, 1997).

To overcome the above, a second version of the LTS was built, based on the plasmids published by Gardner *et al.*, (2000). Here, no SsrA tag was used and only a single GFP reporter was placed bicistronically downstream of the *tetR* gene. Preliminary imaging of the cells did not show any observable phenotypic defects and cells formed regular microcolonies on agar pads. In addition, the cells transformed with the LTS strains were able to produce both GFP from the pIKE107Plo plasmid and eCFP from the pCph8FD(w). Both of these reporters could be distinguished by using appropriate wavelengths for emission and excitation. This was necessary in order to avoid bleed-through of fluorescence because of overlapping GFPmut3b and eCFP spectra. It is worth pointing out the fact that while in the green fluorescence channels a uniform distribution of GFP was observed, the cyan channel showed greater variability, something that points towards plasmid instability. This can explain why some cells showed no cyan fluorescence at all. On the other hand, what needs further investigation is the fact that while some cells showed no or reduced cyan fluorescence, the green levels appeared to be as highly intense as the rest of the microcolony.

The static characterisation assays showed that the LTS was responsive to both IPTG and red light. Samples lacking the pL-Cph8FD(w) plasmid did not show any light sensitivity, and the overall GFP levels although responsive to IPTG were ~2.5 fold lower than the full LTS. Samples lacking the pIKE107Plo plasmid showed light sensitivity but no GFP expression. Samples that were transformed with both plasmids, increased their fluorescence by 3-fold when IPTG was present. Red light, on the other hand, reduced their fluorescence level
by the same factor. eCFP levels in the full LTS and pL-Cph8FD(w) exhibited similar light sensitivity across all IPTG concentrations used. On the other hand, the apparent variability between different IPTG concentrations for the full LTS was greater compared with the pL-Cph8FD(w) only, while the former recorded approximately 100 AU less cyan fluorescence across all conditions, compared to the latter. A repeat of an assay using the same format for induction, but this time comparing the full LTS against the dLTS was performed. The results obtained from this assay were fitted with a Hill function as the qualitative interpretation and to some extent the physiological meaning of the terms of this function is well documented (Ang et al, 2013). From the best fits to the data (lowest error to the data it appeared that the LTS exhibited a steeper transfer function compared to the dLTS as the $n$ exponent was higher for the former. In addition, the IPTG concentration needed for reaching half the maximum induction levels, indicated by the Hill constant $\kappa$, was higher for dLTS. These tendencies, however, were drawn from a low resolution assay in terms of IPTG concentrations across the transient region (since there are a limited number of concentrations that can be run in a single plate), therefore these results cannot be conclusive by themselves but have to be interpreted with caution under the wider context of all the experiments with the LTS.

A more interesting behaviour was observed when a time course experiment of samples starting with either high or low GFP fluorescence was progressed differently over time. More specifically, the samples starting low remained low in the absence of IPTG and presence of red light, while when in the dark, their fluorescence reached high levels for high IPTG. Samples illuminated with red light were unable to reach high green fluorescence values even for the highest IPTG concentration. The samples starting high exhibited a qualitatively different behaviour. The illuminated ones ended in the low or medium fluorescence band for low or high IPTG concentration, respectively, while the dark incubated samples remained highly fluorescent for high IPTG concentration, but had intermediate fluorescence values when IPTG was not present. This indicated that when the LTS was in the state that TetR is expressed at high rates, TetR will keep repressing the expression of the LacI gene even when IPTG is removed. The decrease in fluorescence from high to intermediate levels when that happens, is mainly due to the fact that the LTS was tested in the $E. coli$ JT2 strain where the native LacI gene is constitutively been expressed. Since this LacI gene is not repressible by TetR, a constant LacI amount is present under all conditions, which is able to repress TetR once IPTG is removed. However, the repression from the native LacI is not sufficient to flick the switch towards the low state, so the LTS is bistable in the dark. The model was able to reproduce and support to some extent the above hypothesis. More specifically, when native LacI expression was included, a similar qualitative behaviour was observed when comparing dark incubated samples where the chemical inducer is not present. The ones starting with high fluorescence were stable at intermediate values while the ones starting in the low state remained at low fluorescence. However the dynamics during the first 5 hours do not match, due to the faster dilution rates as the cells switch from stationary phase to exponential growth and the opposite. It is worth stressing the fact that Figure 5.3.11 (a) does not show a genuine 62 hours long experiment, but a staggered sequential projection of assays that were performed in parallel. The main assumption here is that the end point of each sample coincides with the starting conditions of another, i.e. samples incubated with IPTG in the dark during
the assay were assumed to have reached the same conditions as the samples that were incubated with IPTG in the dark during the overnight prior to the start of the experiment. The reason that such a long experiment was not possible, is the observed instability of the pL-Cph8FD(w) plasmid which can yield non-comparable data between the start and the end of the assay.

An alternative hypothesis is that the observed trends are produced from a LTS which is monostable to the low state due to insufficient dilution of the samples starting in the high state during the exponential phase of growth, followed by accumulation of GFP from leaky expression while LacI concentration was high. The fact that the fluorescence dropped to similarly low levels as the samples which started in the low state, but increased at higher rate during the following hours, suggests that TetR and GFP transcription were proceeding at higher rates, something that indicates that LacI in the system was lower possibly due to TetR which was still dominant over the LTS’s Lac.

The main weakness of the time course experiments carried out, which prevented clearer conclusions, was the fact that they were not performed in a chemostatic environment with constant growth conditions. That would require a microfluidic platform or a bioreactor where a constant media flow can ensure that the cultures are diluted and remain at the same phase of growth for long periods while metabolites, antibiotics and chemical inducers are maintained constant in the growth medium. This would enable monitoring if the system is stable for long periods or if LacI dominates over the switch at a later time point, making the system behave as a timer instead of toggle switch, similar to Ellis et al (2009). Such setups were previously demonstrated to yield high resolution characterisation at single cell level (Bennett & Hasty, 2009; Ferry et al, 2011) and even scaled up for high throughput characterisation of synthetic circuits (Maerkl, 2009).

Finally, testing the system in a ΔenvZ ΔlacI strain was not possible. The only strain with the specific genotype available was E. coli EPB238 (Batchelor & Goulian, 2006). Although this strain showed some fluorescent activity during the testing of the dual reporter implementation LTS, its intensity was significantly lower (~20%) compared to the E. coli JT2 strains. When E. coli EPB238 was transformed with the single reporter implementation of the LTS, the obtained fluorescent values were low and hardly distinguishable from the baseline. Furthermore, the fluorescent values across several conditions resulting in low induction, when normalised by OD600 and corrected for baseline, were negative. Although a clear cause for the above was not identified, a major difference between the E. coli EPB238 and JT2 strains is that the former has replaced the native ompR gene with an ompR:yfp fusion. The authors have shown that this can still bind to its DNA recognition sites, however it remains unclear whether the affinity was affected (YFP could be interfering to some extent with the OmpR DNA binding site or OmpR dimerisation on the promoter).

7.2.3. Module 4: Light Tuneable Oscillator

Finally, the dual feedback oscillator (Stricker et al, 2008) was re-designed in order to be tuneable by light. The proposed re-design aimed at providing two control points through the two TCS, where one is used to tune the
period of the oscillation via Cph8/OmpR, while the other to tune the amplitude with CcaS/CcaR.

**Theoretical Work**

The proposed re-design process, in order to achieve the set objectives, introduced an additional negative feedback loop to the oscillator, facilitated by a non-IPTG responsive mutant LacI (LacI<sup>q</sup>) (Hasan & Szybalski, 1995) transcribed under control of the Plac/ompR. Preliminary simulations showed that this loop resulted in faster oscillations (shorter periods) with lower amplitude when red light is absent. The second modification is the replacement of the Plac/ara-1 promoter that drives the GFP output with the Plac/ccaR. When green light increased the transcription rate of the modelled GFP output, oscillations of larger amplitude were observed with no effect on the period.

However, the preliminary model was based on the mechanistic model published for the oscillator (Stricker et al, 2008) and did not take in account the congestion effect of GFP on the ClpXP. When GFP is modelled able to bind, through its SsrA degradation tag, it competes with the binding of all other regulators LacI, LacI<sup>q</sup> and AraC for the available free ClpXP binding sites. The result is an overall decrease in the degradation rates for all free or DNA bound regulators, in a mechanism that was extensively discussed more recently by the same group (Cookson et al, 2011). Apart from the obvious effect on the time evolution of the simulations, this also introduced an additional feedback mechanism from the reporter to the rest of the system something that makes the prospect of a purely independent amplitude tuning impossible. In fact, this effect was more recently used, to couple post-translationally two oscillatory circuits (Prindle et al, 2014).

In order to investigate this effect and possible ways to overcome it, the model was expanded in order to explicitly simulate the complex formation of all the SsrA tagged species with the ClpXP machinery and the resultant degradation. An additional modification was introduced into the LacI loop formation mechanism in order to be more representative of the physical mechanism, where a single LacI tetramer can form the loop instead of two that were initially modelled. The latter increased the range of chemical inducers that can be used and obtain robust and sustained oscillations (Figure 6.2.2 compared to Figure 6.1.5).

From the analysis of the model, in terms of sensitivity defined as the impact of red light on the amplitude and impact of green light on the period, a number of possible solutions can be implemented in order to decouple or minimise the correlation between amplitude and period. Since the primary cause for this is the coupling of the reporter to the rest of the system through the proteosome competition, all proposed solutions aimed at minimising that feedback via decongestion of ClpXP. It is worth noting that all sensitivity analysis was performed in a multidimensional space, at the extremities of the region bound by all four possible light combinations (dark, red only, green only and both wavelengths), across a range of IPTG (0-5 mM) and arabinose (0-10 %) concentration.

The first proposed solution, is the relatively fast and practical tuning of the RBS of the reporter. If this RBS is lowered 10-fold (compared to the initial model value), it was shown in the simulations that the amplitude can be modulated by green light with very small impact on the period (Figure 6.3.2). The sensitivity was slightly higher in the absence of red light.
The second proposed way, is the control of ClpXP from a synthetic genetic construct so that its maximum degradation capacity is controlled. The numerical simulations showed that while increasing amounts of ClpXP in the system can indeed lower the period sensitivity to green light, the sensitivity of amplitude to the red light will increase. Based on that an optimal amount of ClpXP molecules was identified to be 700, but this number is subject to the uncertainty of the parameters. Hence, in a physical implementation of the system, tuning would be required to identify possibly a different optimal point. The second observation from the results shown in Figure 6.3.3 is that for increasing ClpXP, the tuneable range of amplitude and period via light induction is decreased, eventually collapsing to fast oscillations highly sensitive to either wavelength (Figure 6.3.3 (a)). This solution is intriguing for implementation not just for the specific system, but also for understanding and implementing a more controlled regulation for the commonly used SsrA tagged TFs in general. Modifying the activity of ClpXP can in fact have a dramatic impact on the cell division mechanism as shown and discussed in Chapter 5. Since the model does not take in account this crosstalk with housekeeping mechanisms, this solution is less likely to work.

The most effective solution that was identified was the complete decoupling of the reporter (or output) from the ClpXP machinery. This requires the use of a different protease or fast degradation mechanism, otherwise the oscillatory behaviour cannot be sustained, since the output will accumulate during each cycle, which will most certainly be faster than the natural half-life of GFP (24 h). Such proteases were reviewed extensively by (Gottesman & Maurizi, 1992; Gottesman, 1996). For example the Lon protease can achieve similarly fast degradation resulting in half-lives of 10 min for proteins that are tagged on their N-terminus, C-terminus or internally (Gur & Sauer, 2008). Simulating such a mechanism (Figure 6.3.4) suggested not only that the sensitivity of the period to the green light was minimised, but also an improvement in the sensitivity of amplitude to red light. The amplitude sensitivity was minimal when green light was off (hence lower amplitudes) but was more significant when higher amplitudes were induced by green light.

The last aspect of the modelling work on the oscillator was to assess the variability in the periods or amplitude when the system was simulated stochastically. The analysis, which included all possible extreme light conditions (no light, green only, red only and both wavelengths), showed that the oscillator can exhibit clear oscillatory patterns over long periods under most chemical or light conditions. The trends in the light induction agree with the deterministic predictions, where red light can primarily shorten the period and decrease the amplitude, while green light results in increasing amplitude (and period if GFP is coupled to ClpXP). However, for some conditions, where red light was absent and chemical concentrations where such that high amplitude oscillations are generated, the period and the amplitude recorded, were highly variable. The cause for the skewed, abnormal or even bimodal distributions, were caused primarily by the low copy number vector, from which the red light TCS was modelled to be expressed from. When the copy number for the particular genes was increased, all the distributions for the light conditions that OmpR-p has elevated levels, appeared more tight and unimodal. From this, it is strongly suggested that for a physical implementation, low copy number vectors or genomic integration in a single copy should be avoided and instead a two plasmid system should be employed. The disadvantage of the latter, would be the large constructs to be cloned into a single vector,
something that can make the cloning process very challenging, but also the risk of decreasing the stability of these plasmids or the recombination frequency.

7.3. Implications and Future Prospects

7.3.1. Module 1: Light Receptor Characterisation

The mechanistic model produced in this work can become a useful modelling framework not just for the light inducible systems, but two-component systems in general since they all share similar network structure. This model apart from being able to capture the basic signal transduction from the input to the response promoter, is detailed enough to provide access to a range of biologically relevant parameters, which can be used as fine tuning dials. This is achieved through the rigorous reduction method followed, which although it is able to reduce the number of reactions and lump parameters, still retains their physical meaning, which can be mapped back to the biochemical reactions. These include RBS strength, promoter strength, degradation rates and gene copy numbers. In addition, the retroactivity effect from the OmpR-p binding site that naturally occurs in the derivation of the model can be a helpful tool for connecting downstream networks.

From the experimental side, the single plasmid based Cph8 system even if it was underperforming in terms of induction fold difference, is more compatible to be used with other circuits since it has eliminated the crosstalk of 'popular' TF like TetR, LacI and AraC with the TCS components, i.e. the original two plasmid system (Levskaya et al, 2005) utilised regulated promoters from the aforementioned TFs for expression of the chromophore producing enzymes for making the TCS functional. In addition, it provides a single linear construct that can be integrated into the genome, eliminating the need for multiple plasmid systems to control circuits with this TCS. Finally, the fact that the measured heme levels appeared up regulated when ho1 and pcyA were expressed, gives an indication why the expression of pl-PCB is toxic and unstable to the cells as was previously reported (Ferry et al, 2011).

Although the proposed structure of the model in this work was shown able to capture the characteristics of the red light induction module, a number of parameters were uncertain and fitted on limited data. In order to inform the parameter set better, more diverse measurements of the dynamics need to become available. These could include longer time courses, preferably in a suitable microfluidic setup, where the system is induced with a range of time pulses, of variable intensity or length. A significant contribution, would be the possibility of using additional reporters in order to quantify the expression rates for Cph8, OmpR or the PCB production enzymes. Phosphorylation assays similar to the method applied previously for YF1 (Möglichs et al, 2009), can inform better the Cph8 or CcaS mediated phosphorylation rates under different light conditions, while single molecule chase assays (Hammar et al, 2014), can measure the association and dissociation rate constants for OmpR or CcaR to their promoters or receptors. The predictive ability of the model, can then be evaluated further by introducing a series of re-designs and consequently improve the performance of the system.
For example, multiple constructs of the PompC promoter with different reporters at different copy numbers can give insight into the retroactivity mechanism that the model suggested and would allow the experimental estimation of the optimal copy number of OmpR-p binding sites that can be used. Alternative re-designs can address the relatively small fold difference of the light induction. The ways this can be achieved include using a series of RBS strengths or promoters to shift the ratios of Cph8/CcaS to OmpR/CcaR, protein engineering in order to reintroduce the EnvZ phosphatase activity back to Cph8 and finally PompC mutagenesis to reduce its basal expression rate and increase its maximum transcription rate. An interesting avenue could also be the use of PompF instead of PompC, which would invert the effect of light, i.e. activate transcription when red light is present. Finally, the TCS systems if optimised from the above, can then be integrated into the genome in order to obtain a strain that can exert light control without the need for cloning and expression of larger constructs in multi-plasmid systems. A new synthetic circuit that is designed to be controlled by light can do so in such a strain, simply by including a promoter such as Plac/ompR, Plac/ccaR or other possible variants in the network design. Finally, a very intriguing prospect yet to be realised is to expand the multichromatic control system (Tabor et al, 2011) in order to build a Red-Green-Blue (RGB) responsive version by combining Cph8, CcaS and YF1, although taking in account the implications that overexpression of transmembrane proteins can have on the host (Wagner et al 2007; Gubellini et al, 2011).

7.3.2. Module 2: Dual Input Promoters

This work, based on a previously engineered LacI and AraC dual regulated promoter (Lutz & Bujard, 1997), proved the principle that light and chemical responsive promoters can be rationally engineered to be regulated by both types of induction. This provides for a useful tool to connect the TCS with other synthetic circuits, without the need of expressing additional transcription factors. This makes the connection of modules less resource demanding for the cell and, although not demonstrated in this work, it is assumed to be faster than mediating the connection through 'third-party' TFs, due to the delay for their transcription and translation. The relevant model provides a base for a theoretical design and modelling of this connection via the multiple occupancy states of the promoter (Stricker et al, 2008), which can be used for future re-designs of light controllable circuits. However, as it was not possible to account for the mapping between GFP molecules and fluorescence units, in its current form it can only be used for qualitative analysis.

Finally, a useful application and future prospect for the dual input promoters, is the fast prototyping of other synthetic circuits. A very common approach for optimising synthetic circuits is to build multiple versions of the same system using different RBS strength or promoters in order to fine tune the balance between a network’s components. With the use of a dual input promoter instead, a single prototype can be built and light used to modulate the expression rate of a particular component. If the model is sufficiently parameterised, then it can be used to compute the required RBS or promoter strength based on the light intensity used to reach the optimal point.

The work on these promoters can also be extended in a number of ways. Following the same approach,
binding sites for other TF can be used. This will potentially generate a library of connectors in order to make the TCS compatible with most of the existing synthetic circuits. An interesting, yet more challenging aspect is to integrate multichromatic signals on to the same promoter, e.g. OmpR-p and CcaR-p binding sites together. Arranging the specific sites in different ways can generate logic functions that are controlled by light. For example, if a constitutive promoter is engineered to include these sites downstream of the transcription start site (Dr. James Arpino-unpublished work), the binding of the TFs can block the RNA polymerase, rendering a NOR gate. Alternatively, PompC and Pcpc-G2 placed in tandem could yield a promoter which gives an OR function. Finally, the light regulated promoters can be hybridised with TAL effector sites (Blount et al, 2012), such that they render diverse or application specific, dual input promoters. In fact, combinations of dual input promoters harbouring the same light regulated sites but different repressor sites could be used in order to render higher order functions.

7.3.3. Module 3: Light Tuneable Toggle Switch

The implementation of the light tuneable version of the toggle switch is a demonstration of how light can be connected and used to control other synthetic circuits though the dual input promoters. For the system demonstrated, the data indicated that it was bistable in the dark but monostable to the LacI side when light was used. However, in the absence of IPTG, light was only able to drive the system from the high to low state but not the opposite. If this behaviour is confirmed in constant growth conditions (i.e. bioreactor or microfluidics), it suggests that the system can be utilised as a memory controller, where IPTG and phosphorylated OmpR-p (produced in the dark) can activate high expression rates turning the system 'ON'. When IPTG is removed, expression rates fall, but are stabilised at intermediate expression levels (due to the native LacI). If light is switched on then the device resets to the low or 'OFF'. Apart from the potential use to build more complex synthetic networks, this can be scaled to a useful application for light based genetic expression control at an industrial level, e.g. for biofermentation or biofuel production, Lee et al. (2008) suggested that cells could be engineered to "switch from a cellulose digestion mode to fuel production" in response to the environment. Given that two synthetic pathways are each controlled by a side of a toggle switch, using chemical inducers in large quantities, to control a population in a bioreactor can be costly solution compared to the electricity needed to operate a few LED devices.

A future task for this system is its expression in a suitable ΔlacI ΔenvZ strain, in order to confirm the model’s prediction that once in high state, the system should be able to maintain the same levels upon IPTG removal. However, for building a version of this toggle switch where switching the system from the low to the high state only by switching of the light is possibly not feasible. The reason is that the LacI binding sites will be occupied if LacI is starting at high levels, regardless of the light effect, inhibiting transcription, or at least delay significantly the switching time point. Instead, an interesting prospect is the use of the CcaS/CcaR TCS to control the LacI side of the toggle switch. Such a bi-chromatic light tuneable toggle switch would require a dual input promoter that is regulated by TetR instead of LacI, by replacing the appropriate binding sites. This could potentially allow for one wavelength to switch the system to one state and another wavelength to move...
it towards the opposite direction.

7.3.4. Module 4: Light Tuneable Oscillator

This work is the first to my knowledge, to have addressed the prospect of decoupling the amplitude and the period of a genetic oscillator and use light for their control. It was shown that this can be feasible within a range of biological parameters and the introduction of specific modifications to the dual feedback oscillator (Stricker et al, 2008). The detailed mechanistic model that was build was shown useful for investigating the effect of a number of modifications of the initial network and of course it can be further utilised in the future for the theoretical testing of a large number of other possible modifications. More importantly, this study can be a solid building block for a physical implementation of such a system, which would be one of the largest synthetic networks built to date, comprising 10 genes. A significant amount of further work is needed to realise the system experimentally. Cloning what was already shown as a large network, but also implementing a microfluidic platform equipped with a light induction setup e.g. an array of LEDs or a laser is needed. Furthermore, parameter identification can potentially allow for better predictions. The parameter regime that allows this oscillator to operate as discussed was within biologically relevant range of values. On the other hand, a better identification can reveal a wider or more constrained range of parameter values that can be used.

A number of potential applications can then been implemented using this system. It has been previously proposed that next generation biosensors can use frequency encoding (oscillatory output) (Danino et al, 2010) instead of the classic fluorescent intensity as an output. This proposed oscillator with a decoupled amplitude and frequency could potentially combine both, giving a more robust or two dimensional output, one encoded in the frequency and another one in the amplitude. For example, the frequency could be increased by using the trinitrotoluene (TNT) receptor RPB (Looger et al, 2003), another TCS that phosphorylates OmpR, while green light can be used as a gain to the oscillatory output signal in order to increase the dynamic range of the biosensor. Alternatively, the sensory domains of both EnvZ and CcaS can be engineered (Looger et al, 2003; French et al, 2011; Goers et al, 2013) to detect different molecules or pollutants. When one of them e.g. arsenic is detected, it would trigger a shift in the period, while if another one e.g. mercury is also present it would result in a shift in the amplitude.

Another interesting prospective is the implementation of a robust oscillator through feedback control. Based on the control theory application in synthetic circuits as discussed in Chapter 1, light can mediate a feedback control loop which can regulate and correct deviation from a setpoint of amplitude or frequency, due to fluctuations in the environment. For this, a live closed loop electronic circuit is needed, where a fluorescence sensor can monitor the waveform at small time intervals and feed the information into a computer. For any deviation in the amplitude or the frequency, the computational model can compute the intensity and duration of a light pulse of the appropriate wavelength that is needed to correct the output. Finally, an LED array can shine the required amount of green light for minimising the error in the amplitude or red light for the period. Alternatively, constant illumination at different intensities can allow a computer controlled waveform.
Chapter 7. Conclusions and Discussion

7.3.5. Other Future Prospects

As the field of synthetic biology progresses, the realisation amongst synthetic biologists that large complex systems are not easily feasible within the same host has led to the idea of using consortia of cells for facilitating these networks becoming more favourable (Amos, 2014). With this approach, the cells are relieved from burden, but in addition crosstalk can be minimised, as a specific TF can be used in an infinite number of different circuits as long as they are not expressed within the same host. The challenging aspect for the above is how the cells can communicate and exchange information with each other. Small molecules (AHL, HSL) that mediate the natural quorum sensing in bacteria have previously been used in a synthetic biology context (Basu et al, 2005; Tabor et al 2009; Danino et al, 2011) as a reliable tool. Yet, the 'channels' of communication are restricted by the fact that the cells need to exchange matter in the same medium, while the number of orthogonal HSL variants is limited.

Light induction, can open up a new dimension in how microbial population communication can be implemented and the findings in this work can be a contributing factor. This can be done by interfacing electronics and populations so that fluorescent or light signals can be sent and received from each isolated population. Considering the aspects covered by this work, such a network could involve an electronic interface, which operates specific LEDs that can induce one or both TCSs, in a culture that expresses a light tuneable toggle switch. Once the signal is received in one population, OmpR-p or CcaR-p can alter the behaviour of the specific circuit, by activating one of the two double input promoters. The output can be received by an electronic sensor, which can feedback to the same culture through another light pulse, or on a different isolated culture that can express the same or a different circuit. In other words, each culture can perform a small function, communicates through light with another one and collectively perform a more complex computation. The above concept is work in progress and was proposed initially by Dr Guy-Bart Stan and discussed further with Dr Karen Polizzi, Dr James Arpino, Dr Jordan Ang and Prof Mauricio Barahona.

Last, on demand oscillations in a repressilator-like network (Elowitz et al, 2000), is also an interesting prospect. In fact, this work was inspired from a theoretical prediction, that even-number repressilator rings operating at quasi stable state, can give rise to long live transients and oscillations (Strelkowa & Barahona, 2011). The authors, showed that despite the current believes on stability analysis of these rings, that even gene number repressor rings can only have stable equilibrium points, is potentially not the case, when delayed transcriptional waves are propagating around such a large ring (>6 repressors). Their proposed design, of a closed live feedback loop controller, that can read a fluorescent output and signal back via light induction, was predicted to be able to initiate and sustain these oscillations. At the time, the number of available repressors, for an experimental realisation of the system was limited. However, new technologies e.g. TAL effectors (Blount et al, 2012) combined with the findings of this work, regarding the dual regulated promoters, modelling and characterisation of light induction can provide a solid foundation for delivering the specific system in vivo.
Chapter 8

Bibliography

23. Baumgartner, J. W. et al. Transmembrane signalling by a hybrid protein: Communication from the domain of


118 Govan, J. M., Lively, M. O. & Deiters, A. Photochemical control of DNA decoy function enables precise regulation


141 Heyne, K. *et al.* Ultrafast dynamics of phytochrome from the cyanobacterium synechocystis, reconstituted with


## Chapter 9

### Supplementary Information

#### 9.1. Appendix A: Light Receptors

#### 9.1.1. Cph8 TCS Reduced Computational Model

<table>
<thead>
<tr>
<th>Species Table</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
</tr>
<tr>
<td>Intracellular Heme</td>
</tr>
<tr>
<td>Cph8 copy number</td>
</tr>
<tr>
<td>OmpR copy number</td>
</tr>
<tr>
<td>730 nm Light</td>
</tr>
<tr>
<td>635 nm Light</td>
</tr>
<tr>
<td>PCB copy number</td>
</tr>
<tr>
<td>PCB at steady state</td>
</tr>
<tr>
<td>Total OmpR</td>
</tr>
<tr>
<td>Total Cph8</td>
</tr>
<tr>
<td>Fully occupied PompC at steady state</td>
</tr>
<tr>
<td>LacZ</td>
</tr>
<tr>
<td>OmpR-P</td>
</tr>
<tr>
<td>Total PompC:LacZ</td>
</tr>
<tr>
<td>OmpR at steady state</td>
</tr>
<tr>
<td>Free Cph8 at steady state</td>
</tr>
<tr>
<td>Total PcyA and Ho1</td>
</tr>
<tr>
<td>Biliverdin at steady state</td>
</tr>
<tr>
<td>Total holophytochrome</td>
</tr>
<tr>
<td>Cph8:OmpR complex at steady state</td>
</tr>
<tr>
<td>Free Cph8:OmpR complex at steady state</td>
</tr>
<tr>
<td>Total PompC for secondary reporter</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reactions Table</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
</tr>
<tr>
<td>LacZ production reaction</td>
</tr>
<tr>
<td>OmpR-p reaction</td>
</tr>
</tbody>
</table>

---

*Imperial College London, September 2014*

Marios Tomazou *PhD Thesis* 243
<table>
<thead>
<tr>
<th>#</th>
<th>Name</th>
<th>Rule</th>
<th>RuleType</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Function L</td>
<td>TFL = kl<em>Red/(al+kl</em>FarRed+Dp)</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>2</td>
<td>MM 2</td>
<td>km2 = [co*Dp+k2r]/([k2]</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>3</td>
<td>CPHPCB subset SS</td>
<td>CphPCBred = (Tr<em>BCph8</em>PCBred<em>kl)/([PCBred</em>kl]+Dp)</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>4</td>
<td>Total OmpR</td>
<td>OmpRtotss = To*gOmpR</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>5</td>
<td>Total PCB enzymes</td>
<td>Etot = Tr*gPCB</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>6</td>
<td>BV real root</td>
<td>BVss = (sqrt(((-KB+cb)<em>Heme-km1</em>KB)^2<em>Etot^2+(2</em>((-KB+cb)<em>Heme-km1</em>KB)<em>Etot</em>Heme+km1))/ (2<em>Dp</em>Heme+km1))</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>7</td>
<td>Total Cph8</td>
<td>Cph8totss = Tr*gCph8</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>8</td>
<td>PCB root</td>
<td>PCBred = (sqrt(Dp^4*(Heme+km1)^2+2<em>k1</em>(Heme+km1)^2*(Tr<em>gCph8-BVss)<em>Dp+Etot^2</em>Heme+BVss</em>Heme+PCBred))/ (2<em>Dp</em>Heme+km1))</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>9</td>
<td>Free Cph8 for OmpR rate substitution</td>
<td>fCph8ss = OmpRss<em>Tr</em>gCph8/(PCBred<em>kl+Dp)</em>[OmpRss<em>co</em>FarRed+Dp]</td>
<td>Repeated Assignment</td>
</tr>
</tbody>
</table>

**Rules Table**

<table>
<thead>
<tr>
<th>#</th>
<th>Name</th>
<th>Rule</th>
<th>RuleType</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Function L</td>
<td>TFL = kl<em>Red/(al+kl</em>FarRed+Dp)</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>2</td>
<td>MM 2</td>
<td>km2 = [co*Dp+k2r]/([k2]</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>3</td>
<td>CPHPCB subset SS</td>
<td>CphPCBred = (Tr<em>BCph8</em>PCBred<em>kl)/([PCBred</em>kl]+Dp)</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>4</td>
<td>Total OmpR</td>
<td>OmpRtotss = To*gOmpR</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>5</td>
<td>Total PCB enzymes</td>
<td>Etot = Tr*gPCB</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>6</td>
<td>BV real root</td>
<td>BVss = (sqrt(((-KB+cb)<em>Heme-km1</em>KB)^2<em>Etot^2+(2</em>((-KB+cb)<em>Heme-km1</em>KB)<em>Etot</em>Heme+km1))/ (2<em>Dp</em>Heme+km1))</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>7</td>
<td>Total Cph8</td>
<td>Cph8totss = Tr*gCph8</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>8</td>
<td>PCB root</td>
<td>PCBred = (sqrt(Dp^4*(Heme+km1)^2+2<em>k1</em>(Heme+km1)^2*(Tr<em>gCph8-BVss)<em>Dp+Etot^2</em>Heme+BVss</em>Heme+PCBred))/ (2<em>Dp</em>Heme+km1))</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>9</td>
<td>Free Cph8 for OmpR rate substitution</td>
<td>fCph8ss = OmpRss<em>Tr</em>gCph8/(PCBred<em>kl+Dp)</em>[OmpRss<em>co</em>FarRed+Dp]</td>
<td>Repeated Assignment</td>
</tr>
</tbody>
</table>
Chapter 9. Supplementary Information

Parameter Table

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>Value Units</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>kf (called u730 in the text)</td>
<td>3670</td>
<td>1/(minute*molecule)</td>
<td>f(I730) calculated with z</td>
</tr>
<tr>
<td>kl (called u640 in the text)</td>
<td>3220.18</td>
<td>1/(minute*molecule)</td>
<td>f(I730) calculated with z</td>
</tr>
<tr>
<td>k2</td>
<td>0.001383782</td>
<td>1/(minute*molecule)</td>
<td>Yoshida et al, 2002</td>
</tr>
<tr>
<td>k2r</td>
<td>1</td>
<td>1/minute</td>
<td>Yoshida et al, 2003</td>
</tr>
<tr>
<td>cb</td>
<td>0.013</td>
<td>1/minute</td>
<td>Migita et al, 2003</td>
</tr>
<tr>
<td>ar</td>
<td>0.011352948</td>
<td>1/minute</td>
<td>Siryaporn &amp; Goulian, 2008</td>
</tr>
<tr>
<td>a</td>
<td>1.04E-04</td>
<td>1/minute</td>
<td>Siryaporn &amp; Goulian, 2008</td>
</tr>
<tr>
<td>b1</td>
<td>0.215654405</td>
<td>1/(minute*molecule)</td>
<td>Based on Kd, Head et al, 1998</td>
</tr>
<tr>
<td>b1r</td>
<td>1</td>
<td>1/minute</td>
<td>Based on Kd, Head et al, 1999</td>
</tr>
<tr>
<td>b2</td>
<td>0.005830544</td>
<td>1/(minute*molecule)</td>
<td>Based on Kd, Head et al, 2000</td>
</tr>
<tr>
<td>b2r</td>
<td>0.02</td>
<td>1/minute</td>
<td>Based on Kd, Head et al, 2001</td>
</tr>
<tr>
<td>b3</td>
<td>0.001372199</td>
<td>1/(minute*molecule)</td>
<td>Based on Kd, Head et al, 2002</td>
</tr>
<tr>
<td>b3r</td>
<td>1</td>
<td>1/minute</td>
<td>Based on Kd, Head et al, 2003</td>
</tr>
<tr>
<td>al (called kl in the text)</td>
<td>1</td>
<td>1/minute</td>
<td>Fitted on Light transfer Function</td>
</tr>
<tr>
<td>km1</td>
<td>5491.97861</td>
<td>molecule</td>
<td>Cornejo &amp; Beale, 1997</td>
</tr>
<tr>
<td>Tl2</td>
<td>204.1081715</td>
<td>Dimensionless</td>
<td>Fitted</td>
</tr>
<tr>
<td>sa</td>
<td>0.023721985</td>
<td>1/(minute*molecule)</td>
<td>Based on Kd Head et al, 2003</td>
</tr>
<tr>
<td>sar</td>
<td>1</td>
<td>1/minute</td>
<td>Based on Kd Head et al, 2003</td>
</tr>
<tr>
<td>Tr</td>
<td>0.81289847</td>
<td>Dimensionless</td>
<td>Fitted</td>
</tr>
<tr>
<td>To</td>
<td>2869.09743</td>
<td>Dimensionless</td>
<td>Cai &amp; Inouye 2002</td>
</tr>
<tr>
<td>Dp</td>
<td>0.014</td>
<td>1/minute</td>
<td>Dilution rate</td>
</tr>
<tr>
<td>co</td>
<td>0.10364186</td>
<td>1/minute</td>
<td>Fitted</td>
</tr>
<tr>
<td>km3</td>
<td>15142.85714</td>
<td>molecule</td>
<td>Calculated based on k1</td>
</tr>
<tr>
<td>KB</td>
<td>1.014</td>
<td>1/minute</td>
<td>Calculated from Km1 and k1</td>
</tr>
<tr>
<td>k1</td>
<td>6000</td>
<td>1/(minute*molecule)</td>
<td>Assumed fast (Migita et al, 2003; Cornejo &amp; Beale, 1997)</td>
</tr>
<tr>
<td>v</td>
<td>0.02552582</td>
<td>1/minute</td>
<td>Fitted</td>
</tr>
<tr>
<td>Tp</td>
<td>28.27255819</td>
<td>Dimensionless</td>
<td>J23110 calculated based on Golding et al, 2005</td>
</tr>
<tr>
<td>TFL</td>
<td>N/A</td>
<td>Light Function L (Rule)</td>
<td></td>
</tr>
<tr>
<td>km2</td>
<td>N/A</td>
<td>Initial Assignmend Rule</td>
<td></td>
</tr>
<tr>
<td>j</td>
<td>N/A</td>
<td>Rule (Load)</td>
<td></td>
</tr>
<tr>
<td>BA</td>
<td>N/A</td>
<td>Rule (Load)</td>
<td></td>
</tr>
<tr>
<td>BAa</td>
<td>N/A</td>
<td>Rule (Load)</td>
<td></td>
</tr>
</tbody>
</table>
9.1.2. Parameter Fitting and Supplementary Information

Light to LacZ Transfer Function Model Fitting

For the above fitting, $z$ was set to $1E-4$

INLMFIT Fitting summary

<table>
<thead>
<tr>
<th>Group</th>
<th>$T_0$</th>
<th>$T_{tr}$</th>
<th>$u$</th>
<th>$v$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0410850</td>
<td>123.2079039</td>
<td>0.1013046</td>
<td>0.0074346</td>
</tr>
<tr>
<td>G</td>
<td>0.7899362</td>
<td>154.33795276</td>
<td>0.1072156</td>
<td>0.00371407</td>
</tr>
<tr>
<td>1</td>
<td>0.0253080</td>
<td>104.6574787</td>
<td>0.1013512</td>
<td>0.00375566</td>
</tr>
</tbody>
</table>

Mean Standard Deviation
- 0.01209047
- 104.10817346
- 0.1054116
- 0.005253519

Standard Errors

<table>
<thead>
<tr>
<th>Group</th>
<th>$T_0$</th>
<th>$T_{tr}$</th>
<th>$u$</th>
<th>$v$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0027903</td>
<td>81.0001300</td>
<td>0.0017504</td>
<td>0.00075020</td>
</tr>
<tr>
<td>G</td>
<td>0.00273045</td>
<td>122.29803137</td>
<td>0.00172511</td>
<td>0.00171312</td>
</tr>
<tr>
<td>1</td>
<td>0.00160405</td>
<td>45.25337435</td>
<td>0.0017704</td>
<td>0.00165541</td>
</tr>
</tbody>
</table>

Variability for high and low copy promoter for limited OmpR-p

GFP and CFP Stochastic Simulation for Dual Reporter OmpR-p Regulated System

For equal transcription rate but different copy numbers

- $gGFP = 50$ (ColE1)
- $gCFP = 15$ (p15A)

For equal steady state level by different transcription rates and copy numbers

- $gGFP = 50$ (ColE1)
- $gCFP = 15$ (p15A)
### Appendix B: Other Models and Information

#### 9.2.1. Dual Input Promoter Model

<table>
<thead>
<tr>
<th>#</th>
<th>Name</th>
<th>Initial Amount</th>
<th>Initial Amount Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>LacI Tetramer</td>
<td>r4</td>
<td>17</td>
<td>molecule</td>
</tr>
<tr>
<td>LacI mRNA</td>
<td>mr</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>LacI unfolded</td>
<td>ruf</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>LacI</td>
<td>r</td>
<td>3</td>
<td>molecule</td>
</tr>
<tr>
<td>LacI dimmer</td>
<td>r2</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Plac/OmpR free</td>
<td>p10000</td>
<td>20</td>
<td>molecule</td>
</tr>
<tr>
<td>Plac/OmpR OmpR</td>
<td>p1000</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Plac/OmpR LacI4</td>
<td>p10100</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Plac/OmpR OmpR + LacI4</td>
<td>p1010</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Plac/OmpR 2XLacI4</td>
<td>p10020</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Plac/OmpR Looped</td>
<td>pL00</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Plac/OmpR Looped with LacI</td>
<td>pL10</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Plac/ccar free</td>
<td>pg0000</td>
<td>20</td>
<td>molecule</td>
</tr>
<tr>
<td>Plac/ccar Ccar</td>
<td>pg1000</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Plac/ccar LacI4</td>
<td>pg0100</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Plac/ccar Ccar + LacI4</td>
<td>pg1010</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Plac/ccar 2XLacI4</td>
<td>pg0200</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Plac/ccar Ccar + 2XLacI4</td>
<td>pg1020</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Plac/ccar Looped</td>
<td>pgL00</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Plac/ccar Looped with LacI</td>
<td>pgL10</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>GFP mRNA</td>
<td>mg</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>GFP unfolded23</td>
<td>guf</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>GFP</td>
<td>gf</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>OmpR-p</td>
<td>Op</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>OmpR</td>
<td>O</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Cph8</td>
<td>C8</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Cph8:OmpR complex</td>
<td>C8c</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Red light</td>
<td>Red</td>
<td>0</td>
<td>N/A (Intensity)</td>
</tr>
<tr>
<td>Inactive Cph8</td>
<td>C8Pfr</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Phosphatase active CcaS</td>
<td>CG</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Ccar</td>
<td>R</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Kinase CcaS Ccar complex</td>
<td>CGR</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Ccar-R-p34</td>
<td>Rp</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Green light</td>
<td>GREEN</td>
<td>0</td>
<td>N/A (Intensity)</td>
</tr>
<tr>
<td>Kinase active CcaS</td>
<td>CGPr</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Phosphatase CcaS Ccar complex</td>
<td>CGr</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Native LacI gene</td>
<td>Gr</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Secondary GFP mRNA</td>
<td>mgf</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Secondary GFP unfolded</td>
<td>gulf</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>Secondary GFP</td>
<td>gfpl</td>
<td>0</td>
<td>molecule</td>
</tr>
<tr>
<td>#</td>
<td>Reaction</td>
<td>Reaction Rate</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>------------------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>mr -&gt; mr + ruf</td>
<td>tr*mr</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>ruf -&gt; r</td>
<td>kfr*ruf</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2 r &lt;=&gt; r2</td>
<td>kdr<em>r2 / 2 - krd</em>r2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2 r2 &lt;=&gt; r4</td>
<td>kr<em>r2 / 2 - ktr</em>r4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>mr -&gt; null</td>
<td>df*mr</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>ruf -&gt; null</td>
<td>df*ruf</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>r -&gt; null</td>
<td>df*r</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>r2 -&gt; null</td>
<td>df*r2</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>r4 -&gt; null</td>
<td>df*r4</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>pl0000 + 2 Op &lt;-&gt; pl1000</td>
<td>bal<em>pl0000</em>Op^2 - bal*pl1000</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>pl0010 + 2 Op &lt;-&gt; pl1010</td>
<td>bal<em>pl0010</em>Op^2 - baar*pl1010</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>pl0020 + 2 Op &lt;-&gt; pl1020</td>
<td>bal<em>pl0020</em>Op^2 - bal*pl1020</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>pl0000 + r4 &lt;-&gt; pl0010</td>
<td>k2<em>r</em>pl0000<em>r4 - kr</em>pl0010</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>pl1000 + r4 &lt;-&gt; pl1010</td>
<td>k2<em>r</em>pl1000<em>r4 - kr</em>pl1010</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>pl0010 + r4 &lt;-&gt; pl0020</td>
<td>kr<em>pl0010</em>r4 - k2<em>r</em>pl0020</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>pl1010 + r4 &lt;-&gt; pl1020</td>
<td>kr<em>pl1010</em>r4 - k2<em>r</em>pl1020</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>pl1010 -&gt; pl10 + 2 Op</td>
<td>kl*pl1010</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>pl0010 -&gt; pl10</td>
<td>kl*pl0010</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>pl1000 -&gt; pl0000</td>
<td>ku*pl1000</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>pl0000 -&gt; pl1000 + mg</td>
<td>ba*pl0000</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>pl1000 -&gt; pl1000 + mg</td>
<td>baar*pl1000</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>pl1000 -&gt; pl0000</td>
<td>dpo*pl1000</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>pl1010 -&gt; pl0000</td>
<td>dpo*pl1010</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>pl1020 -&gt; pl0000</td>
<td>dpo*pl1020</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>pl0010 -&gt; pl0000</td>
<td>df*pl0010</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>pl0020 -&gt; pl0010</td>
<td>df*pl0100</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>pl10 -&gt; pl0000</td>
<td>df*pl110</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>pg0000 + 2 Rp &lt;-&gt; pg1000</td>
<td>kc<em>pg0000</em>Rp^2 - kcr*pg1000</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>pg0010 + 2 Rp &lt;-&gt; pg1010</td>
<td>kc<em>pg0010</em>Rp^2 - kcr*pg1010</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>pg0020 + 2 Rp &lt;-&gt; pg1020</td>
<td>kc<em>pg0020</em>Rp^2 - kcr*pg1020</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>pg0000 + r4 &lt;-&gt; pg0010</td>
<td>k2<em>pg0000</em>r4 - kr*pg0010</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>pg1000 + r4 &lt;-&gt; pg1010</td>
<td>k2<em>pg1000</em>r4 - kr*pg1010</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>pg0010 + r4 &lt;-&gt; pg0020</td>
<td>kr<em>pg0010</em>r4 - k2<em>r</em>pg0020</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>pg1010 + r4 &lt;-&gt; pg1020</td>
<td>kr<em>pg1010</em>r4 - k2<em>r</em>pg1020</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>pg1010 -&gt; pg110 + 2 Rp</td>
<td>kl*pg1010</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>pg0100 -&gt; pg10</td>
<td>kl*pg0100</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>pg10 -&gt; pg0000</td>
<td>ku*pg1000</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>pg0000 -&gt; pg0000 + mgf</td>
<td>bg*pg0000</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>pg1000 -&gt; pg1000 + mgf</td>
<td>bag*pg1000</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>pg1000 -&gt; pg0000</td>
<td>dc*pg1000</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>pg1010 -&gt; pg0000</td>
<td>dc*pg1010</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>pg1020 -&gt; pg0000</td>
<td>dc*pg1020</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>pg0100 -&gt; pg0000</td>
<td>df*pg0010</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>pg0020 -&gt; pg0000</td>
<td>df*pg0020</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>pg110 -&gt; pg10</td>
<td>df*pg110</td>
<td></td>
</tr>
</tbody>
</table>
RULE Table

<table>
<thead>
<tr>
<th>#</th>
<th>Rule</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ds = df</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>2</td>
<td>kr = krr*(((cmaxr-cminr)<em>2</em>(iptg^{2})+cminr)</td>
<td>Repeated Assignment</td>
</tr>
<tr>
<td>3</td>
<td>k2r = 2*kr</td>
<td>Repeated Assignment</td>
</tr>
<tr>
<td>4</td>
<td>k2rr = 2*krr</td>
<td>Initial Assignment</td>
</tr>
<tr>
<td>5</td>
<td>GFP = d*gfpl</td>
<td>Repeated Assignment</td>
</tr>
<tr>
<td>6</td>
<td>GFPl = d*gfpl</td>
<td>Repeated Assignment</td>
</tr>
</tbody>
</table>

PARAMETER Table

<table>
<thead>
<tr>
<th>#</th>
<th>Name</th>
<th>Value</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>krr</td>
<td>1.8</td>
<td>Stricker et al, (2008)</td>
</tr>
<tr>
<td>2</td>
<td>tr</td>
<td>90</td>
<td>Stricker et al, (2008)</td>
</tr>
<tr>
<td>3</td>
<td>da</td>
<td>0.54</td>
<td>Stricker et al, (2008)</td>
</tr>
<tr>
<td>4</td>
<td>dr</td>
<td>0.54</td>
<td>Stricker et al, (2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>5</td>
<td>kfr</td>
<td>0.9</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>6</td>
<td>klg</td>
<td>0.9</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>7</td>
<td>kdr</td>
<td>0.018</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>8</td>
<td>kt</td>
<td>0.018</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>9</td>
<td>kdr</td>
<td>1.80E-04</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>10</td>
<td>ktr</td>
<td>1.80E-04</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>11</td>
<td>ki</td>
<td>0.62</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>12</td>
<td>kul</td>
<td>0.18</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>13</td>
<td>kr</td>
<td>1</td>
<td>Stricker et al. (2008) (IPTG Rule)</td>
</tr>
<tr>
<td>14</td>
<td>k2r</td>
<td>1</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>15</td>
<td>k2rr</td>
<td>1</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>16</td>
<td>df</td>
<td>0.014</td>
<td>Dilution</td>
</tr>
<tr>
<td>17</td>
<td>ds</td>
<td>0.014</td>
<td>Dilution</td>
</tr>
<tr>
<td>18</td>
<td>bal</td>
<td>4.21E-04</td>
<td>Fitted</td>
</tr>
<tr>
<td>19</td>
<td>balr</td>
<td>169.8215878</td>
<td>Fitted</td>
</tr>
<tr>
<td>20</td>
<td>kc</td>
<td>1.06E-04</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>21</td>
<td>kcr</td>
<td>117.4608485</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>22</td>
<td>dc</td>
<td>6.00E-04</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>23</td>
<td>tg</td>
<td>9</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>24</td>
<td>cmxrr</td>
<td>0.2</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>25</td>
<td>cmnnr</td>
<td>0.01</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>26</td>
<td>kr1</td>
<td>0.035</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>27</td>
<td>kro</td>
<td>0.36</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>28</td>
<td>k2ro</td>
<td>1</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>29</td>
<td>exo</td>
<td>40</td>
<td>Calculated based on Chapter 3</td>
</tr>
<tr>
<td>30</td>
<td>dpo</td>
<td>0.014</td>
<td>Dilution</td>
</tr>
<tr>
<td>31</td>
<td>exa</td>
<td>1.055</td>
<td>Assumed within range</td>
</tr>
<tr>
<td>32</td>
<td>ko1</td>
<td>0.001383782</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>33</td>
<td>ko1r</td>
<td>1</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>34</td>
<td>co</td>
<td>0.55</td>
<td>Fitted</td>
</tr>
<tr>
<td>35</td>
<td>ar</td>
<td>0.00987082</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>36</td>
<td>af</td>
<td>1.14E-05</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>37</td>
<td>kred</td>
<td>3000</td>
<td>Chapter 3 (rounded value)</td>
</tr>
<tr>
<td>38</td>
<td>kal</td>
<td>1</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>39</td>
<td>kgreen</td>
<td>25000</td>
<td>Calculated based on photon flux and z</td>
</tr>
<tr>
<td>40</td>
<td>exc</td>
<td>1.055</td>
<td>Assumed within range</td>
</tr>
<tr>
<td>41</td>
<td>exr</td>
<td>10</td>
<td>Assumed within range</td>
</tr>
<tr>
<td>42</td>
<td>br</td>
<td>0.5</td>
<td>Assumed within range</td>
</tr>
<tr>
<td>43</td>
<td>iptg</td>
<td>0</td>
<td>IPTG concentration mM</td>
</tr>
<tr>
<td>44</td>
<td>d</td>
<td>0.04356813</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>45</td>
<td>ba</td>
<td>0.99861511</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>46</td>
<td>baa</td>
<td>4.07630413</td>
<td>Stricker et al. (2008)</td>
</tr>
<tr>
<td>47</td>
<td>bg</td>
<td>0.10629439</td>
<td>Assumed within range</td>
</tr>
<tr>
<td>48</td>
<td>bag</td>
<td>2.7394751</td>
<td>Assumed within range</td>
</tr>
<tr>
<td>49</td>
<td>kal</td>
<td>1</td>
<td>Calculated based on photon flux and z</td>
</tr>
<tr>
<td>50</td>
<td>dur</td>
<td>5</td>
<td>Assumed within range</td>
</tr>
</tbody>
</table>
### 9.2.2. LTS Model

#### Reaction Table

<table>
<thead>
<tr>
<th>#</th>
<th>Reaction</th>
<th>Reaction Rate</th>
<th>#</th>
<th>Reaction</th>
<th>Reaction Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pt00 + l4 &lt;-&gt; Pt01</td>
<td>kr2<em>Pt00</em>l4 - ks*Pt01</td>
<td>30</td>
<td>Pt00 + Op &lt;-&gt; Pt10</td>
<td>B<em>Pt00</em>Op - Br*Pt10</td>
</tr>
<tr>
<td>2</td>
<td>Pt10 + t2 &lt;-&gt; Pt11</td>
<td>w5<em>Pt10</em>t2 - w6*Pt11</td>
<td>31</td>
<td>Pt01 + Op &lt;-&gt; Pt11</td>
<td>B<em>Pt01</em>Op - Br*Pt11</td>
</tr>
<tr>
<td>3</td>
<td>l + l &lt;-&gt; l2</td>
<td>ks<em>1</em>1 - ks<em>1</em>2</td>
<td>32</td>
<td>Pt10 -&gt; Pt00</td>
<td>bs*Pt10</td>
</tr>
<tr>
<td>4</td>
<td>l2 + l2 &lt;-&gt; l4</td>
<td>ks<em>2</em>12<em>12 - ksr</em>14</td>
<td>33</td>
<td>Pt11 -&gt; Pt00</td>
<td>bs*Pt11</td>
</tr>
<tr>
<td>5</td>
<td>t + t &lt;-&gt; t2</td>
<td>k3<em>t</em>t - k3r*t2</td>
<td>34</td>
<td>Pt00 -&gt; Pt00 + mt</td>
<td>bw2*Pt00</td>
</tr>
<tr>
<td>6</td>
<td>ml -&gt; ufl + uRFP + ml</td>
<td>w10*ml</td>
<td>35</td>
<td>null &lt;-&gt; O</td>
<td>To - br*O</td>
</tr>
<tr>
<td>7</td>
<td>mt -&gt; uft + uGFP + mt</td>
<td>w11*mt</td>
<td>36</td>
<td>null &lt;-&gt; C8</td>
<td>Te - bs*C8</td>
</tr>
<tr>
<td>8</td>
<td>1 &lt;-&gt; null</td>
<td>b1*1</td>
<td>37</td>
<td>C8 + O --&gt; C8 + Op</td>
<td>c<em>C8</em>O</td>
</tr>
<tr>
<td>9</td>
<td>t &lt;-&gt; null</td>
<td>b1*t</td>
<td>38</td>
<td>0 &lt;-&gt; Op</td>
<td>a<em>C8 + ar</em>Op</td>
</tr>
<tr>
<td>10</td>
<td>ml &lt;-&gt; null</td>
<td>gat<em>I</em>ml</td>
<td>39</td>
<td>C8 + Red + Pfr + Red</td>
<td>kl<em>C8</em>Red</td>
</tr>
<tr>
<td>11</td>
<td>mt + ml -&gt; ufl + uRFP + mt</td>
<td>w10*ml</td>
<td>40</td>
<td>Pfr --&gt; C8</td>
<td>al*Pfr</td>
</tr>
<tr>
<td>12</td>
<td>l2 + null</td>
<td>b1*12</td>
<td>41</td>
<td>Pfr + null</td>
<td>bt*Pfr</td>
</tr>
<tr>
<td>13</td>
<td>l4 + null</td>
<td>b1*l4</td>
<td>42</td>
<td>Op + null</td>
<td>bt*Op</td>
</tr>
<tr>
<td>14</td>
<td>t2 + null</td>
<td>b1*t2</td>
<td>43</td>
<td>Pt10 + l4 &lt;-&gt; Pt11</td>
<td>ksr<em>Pt10</em>14 - ksr*Pt11</td>
</tr>
<tr>
<td>15</td>
<td>ufl + l</td>
<td>kfl*ufl</td>
<td>44</td>
<td>Pt01 &lt;-&gt; Pt11</td>
<td>1F<em>Pt11 - 1r</em>Pt11</td>
</tr>
<tr>
<td>18</td>
<td>uft + null</td>
<td>kft*uft</td>
<td>47</td>
<td>Pt10 + l4 &lt;-&gt; Pt11</td>
<td>ksr<em>Pt10</em>14 - ksr*Pt10</td>
</tr>
<tr>
<td>19</td>
<td>1 + 1 + 12</td>
<td>ks<em>1</em>1 - ks<em>1</em>2</td>
<td>48</td>
<td>Pt00 + Op &lt;-&gt; Pt10</td>
<td>B<em>Pt00</em>Op - Br*Pt10</td>
</tr>
<tr>
<td>20</td>
<td>l + l2 + l</td>
<td>ks<em>1</em>12<em>12 - ksr</em>14</td>
<td>49</td>
<td>Pt01 + Op &lt;-&gt; Pt11</td>
<td>B<em>Pt01</em>Op - Br*Pt11</td>
</tr>
<tr>
<td>21</td>
<td>l + l2 + l + 1 + 12</td>
<td>ks<em>1</em>12<em>12</em>12 - ksr<em>14</em>12</td>
<td>50</td>
<td>Pt02 + Op &lt;-&gt; Pt11 + 4 + Op</td>
<td>1F*Pt12</td>
</tr>
<tr>
<td>22</td>
<td>l + l2 + l + 1 + 12 + 12</td>
<td>ks<em>1</em>12<em>12</em>12<em>12 - ksr</em>14<em>12</em>12</td>
<td>51</td>
<td>Pt12 + Op &lt;-&gt; Pt11 + 4 + Op</td>
<td>1F*Pt12</td>
</tr>
<tr>
<td>23</td>
<td>uGFP + GFP</td>
<td>fg*uGFP</td>
<td>52</td>
<td>Pt02 + Op &lt;-&gt; Pt12</td>
<td>B<em>Pt02</em>Op - Br*Pt12</td>
</tr>
<tr>
<td>24</td>
<td>uRFP + GFP</td>
<td>fr*uRFP</td>
<td>53</td>
<td>Pt12 + Op &lt;-&gt; Pt12</td>
<td>B*Pt12</td>
</tr>
<tr>
<td>25</td>
<td>uGFP + null</td>
<td>bg*uGFP</td>
<td>54</td>
<td>Pt02 + Op &lt;-&gt; Pt12</td>
<td>B*Pt12</td>
</tr>
<tr>
<td>26</td>
<td>uRFP + null</td>
<td>bg*uRFP</td>
<td>55</td>
<td>null + uGFP</td>
<td>w10b</td>
</tr>
<tr>
<td>28</td>
<td>GFP + null</td>
<td>bg*GFP</td>
<td>57</td>
<td>Pt01 + Op + Pfr + Pfr1 + Op + 12</td>
<td>1F*Pt11 + Op + Pfr1 + Op + 12</td>
</tr>
<tr>
<td>Parameter Table</td>
<td>Rule Table</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>---------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Name</strong></td>
<td><strong>Value</strong></td>
<td><strong>Name</strong></td>
<td><strong>Value</strong></td>
<td><strong>#</strong></td>
<td><strong>Rule</strong></td>
</tr>
<tr>
<td>w10</td>
<td>0.1</td>
<td>Br</td>
<td>10</td>
<td>1</td>
<td>kr = kr²*(cmmaxr-cminr)/((1+(IPTG/krl)/b11))&lt;br&gt;kr1 = 0.035</td>
</tr>
<tr>
<td>w11</td>
<td>0.085</td>
<td>bw2</td>
<td>0.05</td>
<td>2</td>
<td>k2rr = 2*kr</td>
</tr>
<tr>
<td>kr</td>
<td>697.4263225</td>
<td>a1</td>
<td>1</td>
<td>3</td>
<td>kr2 = 2*kr</td>
</tr>
<tr>
<td>krr</td>
<td>1.8</td>
<td>kl</td>
<td>3220.18</td>
<td>4</td>
<td>w5 = w6*(cmmaxr-cminr)*(1/((1+(aTc/kra1)/b1)))/&lt;br&gt;cmaxr</td>
</tr>
<tr>
<td>w5</td>
<td>0.278970529</td>
<td>ar</td>
<td>0.25708</td>
<td>5</td>
<td>bw2 = w2/20</td>
</tr>
<tr>
<td>w6</td>
<td>0.4</td>
<td>al</td>
<td>1</td>
<td>6</td>
<td>bastb = w2/20</td>
</tr>
<tr>
<td>kl</td>
<td>0.18</td>
<td>c</td>
<td>0.55</td>
<td>7</td>
<td>baslb = w1/500</td>
</tr>
<tr>
<td>k1r</td>
<td>1.80E-04</td>
<td>Tr</td>
<td>1.05</td>
<td>8</td>
<td>bast = w2/3</td>
</tr>
<tr>
<td>k2</td>
<td>0.18</td>
<td>To</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k2r</td>
<td>1.80E-04</td>
<td>lf</td>
<td>0.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k3</td>
<td>0.18</td>
<td>lr</td>
<td>0.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k3r</td>
<td>1.80E-04</td>
<td>ul</td>
<td>0.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kft</td>
<td>0.9</td>
<td>k2rr</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kfl</td>
<td>0.9</td>
<td>bl</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w2</td>
<td>2.2</td>
<td>cmaxr</td>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bl</td>
<td>0.01</td>
<td>cminr</td>
<td>1.00E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bt</td>
<td>0.01</td>
<td>krl</td>
<td>0.035</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gtet</td>
<td>0.1733</td>
<td>kr2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>glacI</td>
<td>0.1733</td>
<td>cmxt</td>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fg</td>
<td>0.092</td>
<td>cmint</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fr</td>
<td>0.046</td>
<td>kra1</td>
<td>1.54E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w1</td>
<td>2.2</td>
<td>w10b</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>basl</td>
<td>1.5</td>
<td>bastb</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.01</td>
<td>bastb</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Br</td>
<td>10</td>
<td>baslb</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 9.2.1. Oscillatory Phase Table

<table>
<thead>
<tr>
<th>Condition</th>
<th>Phase</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IPTG mM</strong></td>
<td><strong>Arabinose %</strong></td>
<td><strong>Red Light W/m²</strong></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>N/A</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>N/A</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>#</td>
<td>Name</td>
<td>Value</td>
</tr>
<tr>
<td>----</td>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>1</td>
<td>ba</td>
<td>0.36</td>
</tr>
<tr>
<td>2</td>
<td>br</td>
<td>0.36</td>
</tr>
<tr>
<td>3</td>
<td>baa</td>
<td>7.2</td>
</tr>
<tr>
<td>4</td>
<td>bar</td>
<td>7.2</td>
</tr>
<tr>
<td>5</td>
<td>krr</td>
<td>1.8</td>
</tr>
<tr>
<td>6</td>
<td>krr</td>
<td>1.8</td>
</tr>
<tr>
<td>7</td>
<td>ta</td>
<td>90</td>
</tr>
<tr>
<td>8</td>
<td>tr</td>
<td>90</td>
</tr>
<tr>
<td>9</td>
<td>da</td>
<td>0.54</td>
</tr>
<tr>
<td>10</td>
<td>dr</td>
<td>0.54</td>
</tr>
<tr>
<td>11</td>
<td>kfa</td>
<td>0.9</td>
</tr>
<tr>
<td>12</td>
<td>kfr</td>
<td>0.9</td>
</tr>
<tr>
<td>13</td>
<td>kfg</td>
<td>0.9</td>
</tr>
<tr>
<td>14</td>
<td>kda</td>
<td>0.018</td>
</tr>
<tr>
<td>15</td>
<td>kdr</td>
<td>0.018</td>
</tr>
<tr>
<td>16</td>
<td>kt</td>
<td>0.018</td>
</tr>
<tr>
<td>17</td>
<td>kdar</td>
<td>1.80E-04</td>
</tr>
<tr>
<td>18</td>
<td>kdr</td>
<td>1.80E-04</td>
</tr>
<tr>
<td>19</td>
<td>ktr</td>
<td>1.80E-04</td>
</tr>
<tr>
<td>20</td>
<td>kl</td>
<td>0.36</td>
</tr>
<tr>
<td>21</td>
<td>kul</td>
<td>0.18</td>
</tr>
<tr>
<td>22</td>
<td>g</td>
<td>1080</td>
</tr>
<tr>
<td>23</td>
<td>c</td>
<td>0.1</td>
</tr>
<tr>
<td>24</td>
<td>l</td>
<td>2.5</td>
</tr>
<tr>
<td>25</td>
<td>ka</td>
<td>1.694117647</td>
</tr>
<tr>
<td>26</td>
<td>k</td>
<td>0.36</td>
</tr>
<tr>
<td>27</td>
<td>k2r</td>
<td>0.72</td>
</tr>
<tr>
<td>28</td>
<td>k2rr</td>
<td>3.2</td>
</tr>
<tr>
<td>29</td>
<td>df</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>ds</td>
<td>6.00E-04</td>
</tr>
<tr>
<td>31</td>
<td>bal</td>
<td>0.0012</td>
</tr>
<tr>
<td>32</td>
<td>balr</td>
<td>100</td>
</tr>
<tr>
<td>33</td>
<td>ko</td>
<td>0.0012</td>
</tr>
<tr>
<td>34</td>
<td>kcr</td>
<td>100</td>
</tr>
<tr>
<td>35</td>
<td>dc</td>
<td>0.014</td>
</tr>
</tbody>
</table>
### Reaction Table

<table>
<thead>
<tr>
<th>#</th>
<th>Reaction</th>
<th>ReactionRate</th>
<th>#</th>
<th>Reaction</th>
<th>ReactionRate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
<tr>
<td>2</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
<tr>
<td>3</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
<tr>
<td>4</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
<tr>
<td>5</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
<tr>
<td>6</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
<tr>
<td>7</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
<tr>
<td>8</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
<tr>
<td>9</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
<tr>
<td>10</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
<tr>
<td>11</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
<tr>
<td>12</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
<tr>
<td>13</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
<tr>
<td>14</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
<tr>
<td>15</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
<tr>
<td>16</td>
<td>( p_{\text{id1}} \leftrightarrow p_{\text{id10}} )</td>
<td>( k_{p_{\text{id10}}} \cdot q_{p_{\text{id1}}} - k_{p_{\text{id1}}} \cdot p_{\text{id10}} )</td>
<td>23</td>
<td>( p_{\text{id10}} \leftrightarrow p_{\text{id1}} + 4 \cdot D )</td>
<td>Dcat*( p_{\text{id10}} )</td>
</tr>
</tbody>
</table>

**Reaction Table**

- **Reaction Rate Formulas**
  - \( r_{\text{id1}} \rightarrow r_{2} \): \( k_{r_{2}} \cdot r_{\text{id1}} \)
  - \( r_{2} \rightarrow r_{\text{id1}} \): \( k_{r_{\text{id1}}} \cdot r_{2} \)
- **Other Relevant Equations**
  - \( r_{\text{id10}} \rightarrow r_{\text{id1}} + 4 \cdot D \)
  - Dcat*\( r_{\text{id10}} \)
  - \( r_{\text{id1}} \rightarrow r_{2} + 4 \cdot D \)
  - Dcat*\( r_{2} \)
  - \( r_{2} \rightarrow r_{\text{id1}} + 4 \cdot D \)
  - Dcat*\( r_{\text{id1}} \)
  - \( r_{\text{id10}} \rightarrow r_{\text{id1}} + 4 \cdot D \)
  - Dcat*\( r_{\text{id10}} \)

**September, 2014**

*Marios Tomazou, PhD Thesis*
Chapter 9. Supplementary Information

\[
\begin{align*}
\text{pa1020} &\rightarrow \text{pa10} \times r_4 \times p_2 \\
\text{pl1011} &\rightarrow \text{pl10} \times r_4 \times p_2 \\
\text{pa0011} &\rightarrow \text{pa00} \times r_4 \times p_2 \\
\text{pg0020} &\rightarrow \text{pg00} \times r_4 \times p_2 \\
\text{pl1020} &\rightarrow \text{pl10} \times r_4 \times p_2 \\
\text{pa1001} &\rightarrow \text{pa10} \times r_4 \times p_2 \\
\text{pl0011} &\rightarrow \text{pl00} \times r_4 \times p_2 \\
\text{pa0001} &\rightarrow \text{pa00} \times r_4 \times p_2 \\
\text{pg1002} &\rightarrow \text{pg10} \times r_4 \times p_2 \\
\text{pg0011} &\rightarrow \text{pg00} \times r_4 \times p_2 \\
\text{pg1020} &\rightarrow \text{pg10} \times r_4 \times p_2 \\
\text{pg0001} &\rightarrow \text{pg00} \times r_4 \times p_2 \\
\text{pa1010} &\rightarrow \text{pa10} \times r_4 \times p_2 \\
\text{pg1011} &\rightarrow \text{pg10} \times r_4 \times p_2 \\
\text{pl1011} &\rightarrow \text{pl10} \times r_4 \times p_2 \\
\text{pa0010} &\rightarrow \text{pa00} \times r_4 \times p_2 \\
\text{pg0010} &\rightarrow \text{pg00} \times r_4 \times p_2 \\
\text{pa1000} &\rightarrow \text{pa10} \times r_4 \times p_2 \\
\text{pg1000} &\rightarrow \text{pg10} \times r_4 \times p_2 \\
\text{paL01} &\rightarrow \text{paL0} \times r_4 \times p_2 \\
\text{pgL01} &\rightarrow \text{pgL0} \times r_4 \times p_2 \\
\text{plL01} &\rightarrow \text{plL0} \times r_4 \times p_2 \\
\text{paL02} &\rightarrow \text{paL0} \times r_4 \times p_2 \\
\text{pgL02} &\rightarrow \text{pgL0} \times r_4 \times p_2 \\
\text{plL02} &\rightarrow \text{plL0} \times r_4 \times p_2 \\
\text{paL0D0} &\rightarrow \text{paL0} \times r_4 \times p_2 \\
\text{pgL0D0} &\rightarrow \text{pgL0} \times r_4 \times p_2 \\
\text{plL0D0} &\rightarrow \text{plL0} \times r_4 \times p_2 \\
\end{align*}
\]
LTS Model

Rule Table
#

Rule

RuleType

1

G = 2*pa1000 + 4*pa0010 + 6*pa1010 + 4*pa0001 + 6* pa1001 + 8*pa0011 +
10*pa1011 + 8*pa0020 + 10*pa1020 + 8*pa0002 + 10 * pa1002 + 4* r4 + 4*r4q +
a + 2*a2 + auf + 2*r2 + 2*r2q + ruf + rufq + 4*paL10 + 4*paL01 + 4*pr0010 +
6*pr1010 + 4*pr0001 + 6* pr1001 + 8*pr0011 + 10*pr1011 + 8*pr0020 + 10*pr1020
+ 8*pr0002 + 10 * pr1002 + 4*prL10 + 4*prL01 + 2*pr1000 + 4*pg0010 + 4*pg1010
+ 4*pg0001 + 4* pg1001 + 8*pg0011 + 8*pg1011 + 8*pg0020 + 8*pg1020 + 8*pg0002
+ 8*pg1002 + 4*pg1001 + 4*pgL10 + 4*pgL01 + 4*pl0010 + 4*pl1010 + 4*pl0001 +
4* pl1001 + 8*pl0011 + 8*pl1011 + 8*pl0020 + 8*pl1020 + 8*pl0002 + 8*pl1002 +
4*pl1001 + 4*plL10 + 4*plL01+rq+r

repeatedAssignment

2

df = g/(c + X) + dc

repeatedAssignment

3

ds = df

repeatedAssignment

4

ka = kar*(((cmaxa-cmina)*((ara^2)/((ka1^2)+(ara^2)))*(1/(1+((iptg/
kr2)^2))))+cmina)

initialAssignment

5

kr = krr*(((cmaxr-cminr)*1/(1+((iptg/kr1)^2)))+cminr)

initialAssignment

6

k2rq = krq*2

initialAssignment

7

dss = g/(c+gfp+guf)

repeatedAssignment

8

LacItot = 4*r4 + 2*r2 + r + ruf

repeatedAssignment

9

k2r = 2*kr

initialAssignment

10

11

LacItot = 4*pa0010 + 4*pa1010 + 4*pa0001 + 4* pa1001 + 8*pa0011 + 8*pa1011 +
8*pa0020 + 8*pa1020 + 8*pa0002 + 8 * pa1002 + 4* r4 + 4*r4q + 2*r2 + 2*r2q +
ruf + rufq + 4*paL10 + 4*paL01 + 4*pr0010 + 4*pr1010 + 4*pr0001 + 4* pr1001
+ 8*pr0011 + 8*pr1011 + 8*pr0020 + 8*pr1020 + 8*pr0002 + 8 * pr1002 + 4*prL10
+ 4*prL01 + 4*pg0010 + 4*pg1010 + 4*pg0001 + 4* pg1001 + 8*pg0011 + 8*pg1011
+ 8*pg0020 + 8*pg1020 + 8*pg0002 + 8*pg1002 + 4*pg1001 + 4*pgL10 + 4*pgL01 +
4*pl0010 + 4*pl1010 + 4*pl0001 + 4* pl1001 + 8*pl0011 + 8*pl1011 + 8*pl0020 +
8*pl1020 + 8*pl0002 + 8*pl1002 + 4*pl1001 + 4*plL10 + 4*plL01 + rq + r
AraCtot = 2*pa1000 + 2*pa1010 + 2* pa1001 + 2*pa1011 + 2*pa1020 + 2* pa1002
+ a + 2*a2 + auf + 2*pr1010 + 2* pr1001 + 2*pr1011 + 2*pr1020 + 2 * pr1002 +
2*pr1000

repeatedAssignment

FALSE

repeatedAssignment

FALSE

Additional computational material on older or intermediated versions of the models
presented is available upon request.

Imperial College London, September 2014

Marios Tomazou PhD Thesis

259

