

Design and implementation of a mammalian synthetic gene oscillator

Hugo Sant'Ana Pereira
Supervised by Prof. Rob Krams and Dr. Karen Polizzi

Department of Bioengineering
Faculty of Engineering
Imperial College London

Submitted in partial fulfilment of the requirements for the degree of Doctor of
Philosophy (Ph.D.)

Abstract

The core goal of synthetic biology as a discipline is to design, develop and characterize biological parts in order to precisely control cellular behaviour. Much of the research in this field has been focused on the development of gene regulatory networks, namely switches and oscillators. The study of synthetic gene oscillators has attracted significant attention in the past decade due to their intriguing dynamics and relevance in controlling inflammatory, metabolic and circadian signalling pathways. Additionally, the precise expression dynamics and molecular mechanisms that underlie the mammalian circadian clock structure are not fully understood.

The work presented herein regards the design and implementation of a tuneable mammalian synthetic gene oscillator with a novel biological structure. To this end, an approach based on a combination of *in silico* design and *in vivo* part validation, in conjunction with a comparative analysis of previously implemented synthetic gene oscillators, was taken when assembling the proposed system. The topology of the system relies on a delayed negative feedback loop, consisting of the coupled regulatory activities of the transcription regulators LacI, tTA, and Gal4.

The numerical solution and stability analysis of an ODE-based model describing the dynamics of the system are indicative that the proposed system is capable of generating sustained oscillations across a wide range of parameter values. The biological parts that comprise the system have been monitored and validated in HEK293T cells through time-lapse fluorescence microscopy and image analysis. The *in vivo* performance of the proposed mammalian synthetic gene oscillator was also assessed in the HEK293T cell line, and monitored using time-lapse fluorescence microscopy. Damped fluorescence oscillations were observed: these could be tuned by a differential IPTG concentration gradient and abolished by doxycycline.

The proposed mammalian synthetic gene oscillator provides valuable insight into the gene expression regulatory processes leading to oscillatory behaviour, and has the potential to foster progress in future synthetic biology-based therapies.

Acknowledgements

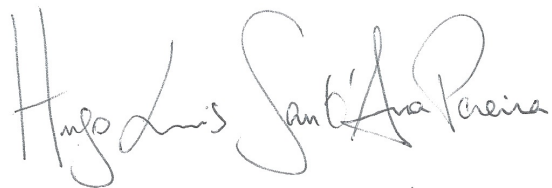
I would firstly like to thank Prof. Rob Krams and Dr. Karen Polizzi, my Ph.D. supervisors. Your support and advice were invaluable throughout the project. Extended gratitude to Dr. Karen Polizzi, for her availability in helping with thesis corrections and proofreading. Likewise, the BBSRC is gratefully acknowledged for funding this work. Dr. Zoltan Kis is hereby acknowledged for his amazing input and help in initially welcoming me to the group and offering his pZK11, pZK13, and pZK14 plasmids. Likewise, Dr. Takayuki Homma is hereby thanked for his brilliant and supportive guidance. To Dr. Ryan Pedrigi, Dr. Vikram Mehta, Daniel Baeriswyl, Miten Patel, Dr. Leila Towhidi, Dr. Nataly Maimari, Dr. Sandra Bovens, Dr. Cristoforo Silvestri, Nisrine Zoghaib, Mohammed Ahmed: you made everyday the most enjoyable it could ever be. The MRes and MSc students I supervised are kindly acknowledged: Joshua Sachs, Dafni Glinos, Jaisa Ahmed, and Xiaoyu Chen. Dr. Bruno Cotta, Dr. Simon Hepworth, and Dr. Chris Haley are thanked for their great support and advice during my PIPS placement at the Imperial Corporate Partnerships and Enterprise Strategy team.

I would also like to acknowledge Marta Garcia Bellmunt, Shanas Choudhury, and Ken Keating for their great help in managing the Tissue Culture and Genomics laboratories. Likewise, Dr. Debora Keller is hereby thanked for her great help and training in the FILM facility. This work has also benefitted from the activities of a great administrative team: Allan Nyunt, Anna Kubik, Melanie Albright, Edit Toth, Nazib Ahmed, and Kate Hobson. I would also like to thank Dr. Benoit Chachuat for his collaboration in the stability analysis work. Dr. Tom Ellis, Dr. Geoff Baldwin, and Dr. Sylvain Ladame are hereby thanked for their availability as ESA and LSR examiners. Thanks to Prof. Martin Fussenegger for his availability in spending one day in our group and sharing so much of his expertise in the field of mammalian synthetic biology. Special thanks to Dr. Mario Eduardo Villanueva: your talent and academic brilliance can only be surpassed by your friendship. Deniz, Gerson, Pi, Shane, Naveed, Jai, Tine and Channarong, thank you for everything.

To my mother, father, Tomás, Filipa and Maria: thank you for your never-ending support and unconditional love; I could not have done it without you.

Declaration of Originality

I, Hugo Sant'Ana Pereira, hereby declare that this thesis, as well as all corresponding work presented herein, is the result of my own original research, unless otherwise noted, acknowledged or referenced. I hereby state that I am the sole author of this document and that I have written it in my own words.

A handwritten signature in black ink that reads "Hugo Luis Sant'Ana Pereira". The signature is written in a cursive style with a large initial 'H'.

March, 2016

Copyright Declaration

'The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives license. 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 license terms of this work'.

Table of Contents

Abstract	2
Acknowledgements.....	3
Declaration of Originality	4
Copyright Declaration	5
Table of Contents	6
List of Publications	10
List of Abbreviations	11
List of Figures	13
List of Tables	21
List of Equations.....	22
Chapter 1 - Introduction	25
1. INTRODUCTION	26
Chapter 2 – Background and Literature Review	31
2. INTRODUCTION TO GENE OSCILLATORS.....	32
2.1. FUNDAMENTAL PRINCIPLES UNDERLYING THE DESIGN OF BIOCHEMICAL OSCILLATORS	34
2.2. THE GOODWIN OSCILLATOR: SINGULAR SELF-REPRESSION	38
2.3. REPRESSILATORS: DELAYED NEGATIVE FEEDBACK IN MULTI-REPRESSOR SYSTEMS.....	40
2.4. THE ATKINSON OSCILLATOR: AMPLIFIED NEGATIVE FEEDBACK	43
2.5. THE FUSSENEGGER OSCILLATORS.....	45

2.6. THE SMOLEN / HASTY OSCILLATOR: DESIGN AND IMPLEMENTATION OF ROBUST OSCILLATORS	
.....	48
2.7. THE NEED FOR NEW OSCILLATOR DESIGNS	51
Chapter 3 – Design of a novel mammalian synthetic gene oscillator	54
3.1. INTRODUCTION	55
3.2. SYSTEM STRUCTURE AND TOPOLOGY	57
3.3. INTRODUCTION TO AN ODE-BASED MODEL OF THE PROPOSED SYNTHETIC GENE OSCILLATOR	
.....	62
3.4. A BRIEF REVIEW OF BIFURCATION THEORY AND ITS APPLICABILITY TO MODEL ANALYSIS	65
3.5. RESULTS	69
<i>3.5.1. Numerical solution of the model describing the proposed synthetic gene oscillator</i>	69
<i>3.5.2. Bifurcation analysis of the proposed synthetic gene oscillator model</i>	72
<i>3.5.3. Considerations about promoter leakiness</i>	76
3.6. DISCUSSION	78
3.7. CONCLUDING REMARKS	81
Chapter 4 – Construction and <i>in vitro</i> validation of system parts	83
4.1. INTRODUCTION	84
4.2. ORGANIZATION OF BIOLOGICAL ELEMENTS AND SYSTEM MODULES	86
4.3. METHODS	90
<i>4.3.1. Bacterial cell culture conditions</i>	90
<i>4.3.2. In silico plasmid design</i>	90
<i>4.3.3. Bacterial transformations</i>	91

4.3.4. Plasmids for system part validation	91
4.3.5. Plasmid preparation and DNA sequencing	92
4.3.6. Endotoxin-free plasmid preparation	93
4.3.7. Agarose gel electrophoresis	94
4.3.8. Mammalian cell culture conditions.....	94
4.3.9. Turbofect™ transfection and in vivo implementation of system components	95
4.3.10. Microscopy imaging and image processing	96
4.4. RESULTS AND DISCUSSION	98
4.4.1. Construction of synthetic gene circuits for validation of system components	98
4.4.2. In vivo implementation of the LacI validation synthetic gene circuit.....	108
4.4.3. In vivo implementation of the tTA validation synthetic gene circuit	113
4.4.4. In vivo implementation of the Gal4 validation synthetic gene circuit.....	116
4.5. CONCLUDING REMARKS	120
Chapter 5 – In vivo implementation of a novel mammalian synthetic gene oscillator	122
5.1. INTRODUCTION	123
5.2. REORGANIZATION AND COUPLING OF SYSTEM MODULES	124
5.3. METHODS	126
5.3.1. Plasmids for system in vivo implementation	126
5.3.2. Turbofect™ transfection and system in vivo implementation.....	127
5.3.3. Microscopy imaging, image processing and single cell analysis.....	127
5.4. RESULTS AND DISCUSSION	129

5.4.1. <i>In vitro construction of system modules</i>	129
5.4.2. <i>In vivo implementation of the proposed mammalian synthetic gene oscillator</i>	137
5.4.3. <i>Exploring system tunability with chemical inhibitors</i>	144
5.5. CONCLUDING REMARKS	149
Chapter 6 – Future work and general conclusions	151
References	158
Appendix A.3.	171
Appendix A.4.	202
Appendix A.5.	223

List of Publications

Peer-reviewed journal articles

1. Zoltán Kis, Hugo Sant'Ana Pereira, Takayuki Homma, Rob Krams (2015) Mammalian synthetic biology: emerging medical applications, *Journal of the Royal Society Interface*, 12, 106;
2. Mario Eduardo Villanueva, Hugo Sant'Ana Pereira, Jai Rajyaguru, Rob Krams, Benoit Chachuat (2015) Bifurcation and stability analysis of nonlinear dynamic systems using complete search, *Proc. AIChE Annual Meeting 2015*, 508d;
3. Hugo Sant'Ana Pereira, Mario Eduardo Villanueva, Rob Krams, Benoit Chachuat (2016) Bifurcation analysis in systems and synthetic biology using complete-search methods, *In preparation*;
4. Mario Eduardo Villanueva, Hugo Sant'Ana Pereira, Benoit Chachuat (2016) Bifurcation and Stability analysis of process models using complete-search methods, *In preparation*;
5. Hugo Sant'Ana Pereira, Rob Krams (2016) A damped mammalian synthetic gene oscillator, *In preparation*;

Poster presentations at conferences

6. Hugo Sant'Ana Pereira, Jaisa Ahmed, Rob Krams (2014) Design and implementation of a mechanosensitive gene oscillator in mammalian cells, *MEC Bioeng 2014*, Poster 221, London, U.K.

List of Abbreviations

Bp: Base-pair

CMV: Cytomegalovirus

DDE: Delay Differential Equation

DMEM: Dulbecco's Modified Eagle Medium

DNA: deoxyribonucleic acid

EDTA: Ethylenediaminetetraacetic acid

FBS: Foetal Bovine Serum

GFP: Green Fluorescent Protein

GRN: Gene Regulatory Network

HEK: Human Embryonic Kidney

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HSV: Herpes Simplex Virus

IPTG: Isopropyl β -D-1-thiogalactopyranoside

Kbp: Kilo base-pair

lacO: lactose operator

LB medium: Luria-Bertani Medium

mRNA: messenger ribonucleic acid

NF- κ B: Nuclear Factor Kappa B

NLS: Nuclear Localization Signal

NRI: Nitrogen Regulator I

NR1p: Phosphorylated Nitrogen Regulator I

ODE: Ordinary Differential Equation

ORF: Open Reading Frame

PBS: Phosphate Buffered Saline

PCR: Polymerase Chain Reaction

PIT: Pristinamycin Transactivator

PolyA: Poly-Adenylation

RFU: Relative Fluorescence Unit

List of Abbreviations

RNA: ribonucleic acid

RSV: Rous Sarcoma Virus

ROI: Region of Interest

SDE: Stochastic Differential Equation

SGO: Synthetic Gene Oscillator

siRNA: small interfering ribonucleic acid

SOC medium: Super Optimal Medium with Catabolite Repression

SV40: Simian Virus 40

TBE: Tris-Borate EDTA

TetR: Tetracycline Repressor

tetO: tetracycline operator

TRE: Tetracycline Responsive Element

tTA: Tetracycline Transactivator

UAS: Upstream Activation Sequence

yemGFP: Modified Yeast Enhanced Green Fluorescent Protein

List of Figures

Figure 2.1. Diagram of a feedback loop.	35
Figure 2.2. Adapted from Purcell <i>et al.</i> , 2010. (Left) Schematic representation of the Goodwin oscillator topology. (Right) The circuit design by Lutz and Bujard (1997) for <i>in vivo</i> implementation.....	38
Figure 2.3. Adapted from Elowitz and Leibler, 2000; and Purcell <i>et al.</i> , 2010. (Top) Diagrammatic representations of the network topology, promoters and genes used in the Elowitz and Leibler repressilator. (Bottom) Representation of the plasmids containing the system core regulatory modules and the reporter construct.	41
Figure 2.4. Adapted from Purcell <i>et al.</i> , 2010. (Left) Diagrammatic representation of the topology Atkinson oscillator topology. (Right) Illustration of the activator (red) and repressor modules (blue).	43
Figure 2.5. Adapted from Tigges <i>et al.</i> (2009) and Purcell <i>et al.</i> (2010). (Left) Interaction map of the components that comprise the original Fussenegger oscillator design. (Right) Diagrammatic representation of the system topology, where A and C represent tTA and PIT, respectively; B represents antisense tTA mRNA.....	44
Figure 2.6. Adapted from Stricker <i>et al.</i> (2008) and Purcell <i>et al.</i> (2010). (Left) Diagrammatic representation of the gene network constructed by Stricker <i>et al.</i> (2008) for the <i>in vivo</i> implementation of the theoretically designed Smolen oscillator (Smolen, Baxter, & Byrne, 1998).	50
Figure 3.1. Diagrammatic representation of the proposed oscillator design.....	61

Figure 3.2. Concentration profiles of LacI, tTA, Gal4 and GFP proteins in the absence of chemical inhibitors IPTG and tetracycline, assuming equimolar gene dosage ratios (1:1:1:1). Results are shown for the 1500 – 7500 minute time window. 70

Figure 3.3. Concentration profiles of LacI, tTA, Gal4 and GFP mRNA transcripts in the absence of chemical inhibitors IPTG and tetracycline, assuming equimolar gene dosage ratios (1:1:1:1). Results are shown for the 1500 – 7500 minute time window. 70

Figure 3.4. Preliminary analysis of frequency and amplitude modulation by tetracycline (top) and IPTG (bottom). 71

Figure 3.5. Tunability potential of parameters variable through experimental setup. 75

Figure 4.1. Diagrammatic representation of the *lacI*, *tTA*, *gal4* and *gfp* modules. 86

Figure 4.2. Diagrammatic representation of *lacI*-, *tTA*-, and *gal4*-based synthetic gene circuits for individual module *in vivo* validation. 89

Figure 4.3. pCMVLacI plasmid map. 99

Figure 4.4. pOPRS_GFP plasmid map. 100

Figure 4.5. pCMVGal4 plasmid map. 101

Figure 4.6. pUAS-GG plasmid map..... 102

Figure 4.7. Agarose gel electrophoresis of pUAS-GG, pCMVGal4, pOPRS_GFP and pCMVLacI..... 103

Figure 4.8. pCMVtTA plasmid map.	104
Figure 4.9. pZK14 plasmid map.....	105
Figure 4.10. InFusion® PCR-mediated extraction of the <i>tTA</i> ORF	106
Figure 4.11. Agarose gel electrophoresis result of a pZK11 and pZK14 restriction enzyme digestions with HindIII and BamHI, respectively.....	107
Figure 4.12. Fluorescence channel images of HEK cells transfected with LacI validation synthetic gene circuit components.....	110
Figure 4.13. Fluorescence intensity time evolution of HEK293T cells co-transfected with pCMVLacI and pOPRS_GFP, in DMEM medium supplemented with 1 mM IPTG.	111
Figure 4.14. Fluorescence channel images of HEK cells transfected with tTA validation synthetic gene circuit components.....	114
Figure 4.15. Fluorescence intensity time evolution of HEK293T cells co-transfected with pCMVtTA and pZK14, in DMEM medium supplemented with and without 1 ug/mL doxycycline.....	115
Figure 4.16. Fluorescence channel images of HEK cells transfected with Gal4 validation synthetic gene circuit components.....	117
Figure 4.17. Fluorescence intensity time evolution of HEK293T cells co-transfected with pCMVGal4 and pUAS-GG.....	118
Figure 5.1. Diagrammatic representation of the proposed mammalian synthetic gene oscillator topology and biological organization.....	125

Figure 5.2. pUAS_LacI plasmid map.....	130
Figure 5.3. pOPRS_tTA plasmid map.	131
Figure 5.4. pTRE3G_Gal4 plasmid map.....	132
Figure 5.5. pTRE3G_UbGFP plasmid map.	133
Figure 5.6. Agarose gel electrophoresis of pUAS_LacI (lanes 2 to 6) after BamHI restriction digestion.....	134
Figure 5.7. Agarose gel electrophoresis of pTRE3G_Gal4 (lanes 2 to 6) after EcoRI restriction digestion.....	135
Figure 5.8. Agarose gel electrophoresis of pOPRS_tTA after NheI restriction digestion (lanes 11 and 12), shown in conjunction with pUAS-GG (lanes 3 and 4), pCMVGal4 (lanes 5 and 6), pOPRS_GFP (lanes 7 and 8), and pCMVLacI (lanes 9 and 10).....	135
Figure 5.9. Agarose gel electrophoresis of pTRE3G_UbGFP after BamHI restriction digestion (lanes 2 to 6), shown in conjunction with pZK14 (lanes 7 to 11).	135
Figure 5.10. Single cell analysis of three cells exhibiting damped fluorescence oscillations after co-transfection with equimolar ratios of pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pTRE3G_UbGFP.....	138
Figure 5.11. Fluorescence channel images corresponding to the single-cell signals illustrated in figure 5.10.	139

Figure 5.12. Damped GFP oscillatory signal obtained by model simulation assuming 0.05:1:1, 0.1:1:1, 0.15:1:1, and 0.2:1:1 molarity ratios between pUAS_LacI, pOPRS_tTA, and pTRE3G_Gal4, respectively. 142

Figure 5.13. Sustained oscillatory signal obtained by simulating the dynamic model of the proposed synthetic gene oscillator with updated maximal transcription rates: 22.8 min⁻¹, 30 min⁻¹, and 153 min⁻¹ for tTA, Gal4 and LacI, respectively.143

Figure 5.14. Single cell analysis of three cells exhibiting damped fluorescence oscillations after co-transfection with equimolar ratios of pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pTRE3G_UbGFP, in complete DMEM médium supplied with 1 mM IPTG. 145

Figure 5.15. Single cell analysis of three cells exhibiting damped fluorescence oscillations after co-transfection with equimolar ratios of pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pTRE3G_UbGFP, in complete DMEM médium supplied with 2 mM IPTG. 146

Figure 5.16. Single cell analysis of three cells exhibiting damped fluorescence oscillations after co-transfection with equimolar ratios of pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pTRE3G_UbGFP, in complete DMEM medium supplied with 3 mM IPTG. 146

Figure 5.17. Detected single-cell fluorescence cycle periods for HEK293T cells co-transfected with pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pTRE3G_UbGFP at equimolar ratios, and supplied with 1 mM, 2mM, and 3 mM IPTG. 147

Figure A.3.1. Bifurcation diagram of parameter k1 with respect to GFP concentration. 181

Figure A.3.2. Bifurcation diagram of parameter k_2 with respect to GFP concentration.
..... 181

Figure A.3.3. Bifurcation diagram of parameter k_3 with respect to GFP concentration.
..... 182

Figure A.3.4. Bifurcation diagram of parameter k_4 with respect to GFP concentration.
..... 182

Figure A.3.5. Bifurcation diagram of parameter k_{dm1} with respect to GFP concentration. 183

Figure A.3.6. Bifurcation diagram of parameter k_{dm2} with respect to GFP concentration. 183

Figure A.3.7. Bifurcation diagram of parameter k_{dm3} with respect to GFP concentration. 184

Figure A.3.8. Bifurcation diagram of parameter k_{dp1} with respect to GFP concentration. 184

Figure A.3.9. Bifurcation diagram of parameter k_{dp2} with respect to GFP concentration. 185

Figure A.3.10. Bifurcation diagram of parameter k_{dp3} with respect to GFP concentration. 185

Figure A.3.11. Bifurcation diagram of tetracycline concentration with respect to GFP concentration. 186

Figure A.3.12. Bifurcation diagram of IPTG concentration with respect to GFP concentration. 186

Figure A.3.13. Bifurcation diagram of parameter g_1 with respect to GFP concentration. 187

Figure A.3.14. Bifurcation diagram of parameter g_2 with respect to GFP concentration. 187

Figure A.3.15. Bifurcation diagram of parameter g_3 with respect to GFP concentration, in the range of $0 < g_3 < 0.01$ 188

Figure A.3.16. Bifurcation diagram of parameter g_3 with respect to GFP concentration, in the range of $0 < g_3 < 50$ 188

Figure A.3.17. Bifurcation diagram of parameter a_1 with respect to GFP concentration. 189

Figure A.3.18. Bifurcation diagram of parameter a_2 with respect to GFP concentration. 189

Figure A.3.19. Bifurcation diagram of parameter a_3 with respect to GFP concentration. Red dot(s) indicate Hopf points. 190

Figure A.4.2.1. Fluorescence channel images superimposed to brightfield channel images of HEK cells transfected with LacI validation synthetic gene circuit components..... 220

Figure A.4.2.2. Fluorescence channel images superimposed to brightfield channel images of HEK cells transfected with tTA validation synthetic gene circuit components..... 221

Figure A.4.2.3. Fluorescence channel images superimposed to brightfield channel images of HEK cells transfected with Gal4 validation synthetic gene circuit components..... 222

Figure A.5.1. Fluorescence channel images corresponding to the single-cell signals illustrated in figure 5.13. 233

Figure A.5.2. Fluorescence channel images corresponding to the single-cell signals illustrated in figure 5.14. 234

Figure A.5.3. Fluorescence channel images corresponding to the single-cell signals illustrated in figure 5.15. 235

List of Tables

Table 2.1. Comparison between different <i>in vivo</i> implementations of synthetic gene oscillators.	53
Table 3.1. Core parameter values and respective Hopf intervals corresponding to oscillatory behaviour in the model described by equations 3.1 – 3.8.	73
Table 3.2. Hopf interval ranges of tetracycline, IPTG and gene dosage parameters..	74
Table 3.3. Hopf point locations of leakiness parameters for core system components.	77
Table 4.1. Concentration and purity of plasmids encoding system parts to be validated through synthetic gene circuit <i>in vivo</i> implementation.	108
Table 5.1. Concentration and purity of plasmids encoding the finalized system parts, i.e. pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4, and pTRE3G_UbGFP.....	137
Table A.3.1. List of parameters for the ODE-based model described by equations 3.1 to 3.10.....	172

List of Equations

Equation 2.1. ODE describing an example of a protein repressing its own production.	36
Equation 2.2. DDE equation derived from equation 2.1.	37
Equation 3.1. ODE describing the dynamics of LacI mRNA production in the context of the proposed mammalian synthetic gene oscillator.	63
Equation 3.2. ODE describing the dynamics of tTA mRNA production in the context of the proposed mammalian synthetic gene oscillator.	63
Equation 3.3. ODE describing the dynamics of Gal4 mRNA production in the context of the proposed mammalian synthetic gene oscillator.	63
Equation 3.4. ODE describing the dynamics of GFP mRNA production in the context of the proposed mammalian synthetic gene oscillator.	64
Equation 3.5. ODE describing the dynamics of LacI protein production in the context of the proposed mammalian synthetic gene oscillator.	64
Equation 3.6. ODE describing the dynamics of tTA protein production in the context of the proposed mammalian synthetic gene oscillator.	64
Equation 3.7. ODE describing the dynamics of Gal4 protein production in the context of the proposed mammalian synthetic gene oscillator.	64
Equation 3.8. ODE describing the dynamics of GFP protein production in the context of the proposed mammalian synthetic gene oscillator.	64

Equation 3.9. Transfer function describing the IPTG-LacI interaction kinetics..... 64

Equation 3.10. Transfer function describing the tetracycline-tTA interaction kinetics.
..... 65

Equation 3.11. Example of a two-variable ODE system..... 65

Equation 3.12. Steady state expression of the system represented by equation 3.11.
..... 66

Equation 3.13. General Taylor expansion form of the system described by equation
3.11..... 66

Equation 3.14. System described by equation 3.1 re-written in a linearized form, and
identification of its Jacobian matrix..... 66

Equation 3.15. Derivation of the *eigenvalues* of equation 3.14. 67

“Thus the system of the world only oscillates around a mean state from which it never departs except by a very small quantity. By virtue of its constitution and the law of gravity, it enjoys a stability that can be destroyed only by foreign causes, and we are certain that their action is undetectable from the time of the most ancient observations until our own day. This stability in the system of the world, which assures its duration, is one of the most notable among all phenomena, in that it exhibits in the heavens the same intention to maintain order in the universe that nature has so admirably observed on Earth for the sake of preserving individuals and perpetuating species.”

Pierre-Simon Laplace (1786), *Sur l'Equation Séculaire de la Lune*

Chapter 1

Introduction

1. Introduction

The core goal of synthetic biology as a discipline is to design, develop and characterize biological parts in order to precisely control cellular behaviour (Mukherji & van Oudenaarden, 2009; Purcell, Savery, Grierson, & di Bernardo, 2010). Specifically, synthetic biologists work in the development and synthesis of biological networks, circuits and devices that perform desired functions in a predictable way (Endy, 2005; Purcell et al., 2010; Serrano, 2007). A dynamic and continuous combination of *in vivo* and *in silico* analysis is required to achieve these goals (Purcell et al., 2010). Synthetic biology can thus be regarded as a multi-disciplinary field, combining concepts of engineering, mathematics and biology (Mukherji & van Oudenaarden, 2009; Purcell et al., 2010). Recent accomplishments for synthetic biology have been based on the engineering of cells resulting in practical applications, namely in the areas of biofuel production (Steen et al., 2008; Waks & Silver, 2009), bioremediation (Gilbert, Walker, & Keasling, 2003), and biosensing (Rajendran & Ellington, 2008), as well as clinical applications (Abil, Xiong, & Zhao, 2014; Anderson, Clarke, Arkin, & Voigt, 2006; Khalil & Collins, 2010; Khosla & Keasling, 2003; Kis, Sant'Ana Pereira, Homma, Pedrigi, & Krams, 2015; Lienert, Lohmueller, Garg, & Silver, 2014; Lu, Khalil, & Collins, 2009; Mukherji & van Oudenaarden, 2009; Ro et al., 2006; Ruder, Lu, & Collins, 2011; Serrano, 2007; Tavassoli, 2010; Ye & Fussenegger, 2014).

The development of such applications is usually accomplished through the design and synthesis of artificial gene networks that elicit specific biological functions (O'Brien, Van Itallie, & Bennett, 2012). Genetic networks can be defined as "intricate webs of interactions between regulatory elements controlling protein productions" (O'Brien, Van Itallie, et al., 2012). Given the topological similarities observed between these networks and electronic devices, the term "genetic circuit" is often used to refer to a genetic network (Hasty, McMillen, & Collins, 2002; McAdams & Shapiro, 1995; O'Brien, Van Itallie, et al., 2012). Generally, naturally occurring gene networks are complex systems, usually comprised of hundreds to thousands of interacting genes (Pedraza & van Oudenaarden, 2005). Conversely, synthetic gene

circuits are simple circuits, made of a few interacting genes (O'Brien, Van Itallie, et al., 2012). However, even being relatively simple systems, the introduction and subsequent behaviour of synthetic gene circuits is highly affected by inherent cellular noise (Pedraza & van Oudenaarden, 2005). The fluctuation of reaction rates (e.g. for transcription and translation) due to stochastic variation in the cellular pool of housekeeping genes is a good example of this, amongst many others (Pedraza & van Oudenaarden, 2005; Rao & Arkin, 2003). This intrinsic complexity in implementing synthetic gene circuits is usually dealt with by synthetic biologists through the creation of mathematical models as an aid to system design. Indeed, numerous theoretical gene circuit designs have been modelled and proposed for eukaryotic and prokaryotic *in vivo* implementation over the past few years (Purcell et al., 2010). However, only a fraction of these theoretical systems is successfully implemented in host cells.

Since the work developed by Becksei & Serrano (2000) and Elowitz & Leibler (2000), two main types of synthetic gene networks have drawn much attention in the field of synthetic biology: switches and oscillators (Purcell et al., 2010; Tyson, Albert, Goldbeter, Ruoff, & Sible, 2008):

- Switches and switch-based circuits are relatively simple networks that can output gene expression routines in response to a stimulus-based input. These systems have yielded most of the pioneering synthetic biology applications. Kis *et al.* (2015), Auslander & Fussenegger (2013) and Karlsson & Webber (2012) have extensively reviewed some of the most iconic switches and switch-based circuits developed to date (Ausländer & Fussenegger, 2013; Karlsson & Weber, 2012; Kis et al., 2015);
- Oscillators are comprised of biological elements that present periodic expression profiles. Unlike switches, their design takes into account dynamical systems properties such as *delay* and *feedback*. There is growing interest in the development of synthetic gene oscillators mainly due to their applicability to cyclic cellular processes, from circadian rhythms to anti-inflammatory responses

(Dibner, Schibler, & Albrecht, 2010; Goldbeter, 1995; Gonze, 2010; Hoffmann, Levchenko, Scott, & Baltimore, 2002; Kemler & Fontana, 1999; Kobayashi & Kageyama, 2010; Kwon, Choe, Son, & Kim, 2011; Monk, 2003; Nelson et al., 2004; Susaki, Stelling, & Ueda, 2010; Tigges, Marquez-Lago, Stelling, & Fussenegger, 2009; H. R. Ueda, 2007; Hiroki R Ueda et al., 2005; Ukai & Ueda, 2010; Ukai-Tadenuma, Kasukawa, & Ueda, 2008).

Pioneering synthetic biology studies focused on the implementation and standardization of network components in prokaryotic cells (Michael B Elowitz & Leibler, 2000; Gardner, Cantor, & Collins, 2000; Kis et al., 2015). However, more recent applications have emerged in eukaryotes, namely in yeast and mammalian cells (Church, Elowitz, Smolke, Voigt, & Weiss, 2014; Khalil et al., 2012; Kis et al., 2015; Purcell, Peccoud, & Lu, 2014; Tigges & Fussenegger, 2009). In contrast to prokaryotes, eukaryotic cells have developed innate complex pathways to resist the introduction of foreign genetic material, making them harder to work with in a synthetic biology context (Kis et al., 2015). Their highly compartmentalized morphology also poses an obstacle to the implementation of synthetic gene networks (Kis et al., 2015). Over the past decade, synthetic biology applied to mammalian cells has greatly evolved from the development of simple gene switches and networks to oscillators and therapy-oriented circuits (Church et al., 2014; Khalil et al., 2012; Kis et al., 2015; Purcell et al., 2014; Tigges & Fussenegger, 2009). Currently, mammalian synthetic biology enables the development of strategies and technologies that can be applied to gene- and cell-based therapies, namely in the areas of personalized medicine, artificial insemination, cardiovascular disease, and treatment of metabolic and immune disorders (Church, Elowitz, Smolke, Voigt, & Weiss, 2014; Kis et al., 2015; Tigges & Fussenegger, 2009).

Despite the continued progress in synthetic biology there are still challenges to be addressed. One of these is the development of a robust mammalian synthetic gene oscillator. Indeed, a viable implementation of such a system would (i) greatly increase our understanding of mammalian gene expression mathematical models, (ii) provide a framework to study cyclic cellular processes with unprecedented

accuracy, and (iii) potentially lead to the development of therapies in the fields of cell cycle regulation and inflammatory responses (Tigges, Déneraud, Greber, Stelling, & Fussenegger, 2010; Tigges et al., 2009). Several designs for such a system have been proposed (O'Brien, Van Itallie, et al., 2012), but only two systems have been implemented *in vivo* (Tigges et al., 2010, 2009). These implementations, however, lack the desired level of robustness to achieve the goals mentioned above: oscillatory behaviour was reported to be irregular and occurred only for a small percentage of the transfected cellular population (Purcell et al., 2010; Tigges et al., 2010, 2009). Several implementation-related aspects have been proposed to explain this, namely the dynamical complexity of bi-directional mRNA expression in mammalian cells, and encoding of the network in different plasmids thus reducing the probability of having all the system components in the same cell (Purcell et al., 2010; Tigges et al., 2009). Moreover, the period of previously implemented mammalian synthetic gene oscillators can only be altered through the variation of transfected system component relative dosages (Purcell et al., 2010; Tigges et al., 2009). This limits the potential physiological applicability of these systems: this frequency tuning strategy depends entirely on the chosen experimental setup, rather than in a tangible, real-time biochemical interaction between tuning agent and system components.

The aim of this work is to design and implement a novel mammalian synthetic gene oscillator that shows an increase in robustness and biochemical tunability range in comparison to previous designs. The remainder of this thesis is structured as follows:

- *Chapter 2* provides a comprehensive review of the most iconic synthetic gene oscillators developed to date. Additionally, a list of fundamental principles underlying the design of biochemical oscillators is established;
- *Chapter 3* addresses the *in silico* design of a novel synthetic gene oscillator from a topological and dynamical perspective. The numerical solution for a lumped

parameter kinetic ODE model is presented and its parametric space is analyzed through bifurcation analysis;

- *Chapter 4* addresses the biological design and construction of a novel mammalian synthetic gene oscillator. The construction of system pre-components, as well as the individual validation of their activity in mammalian cells, is presented;
- *Chapter 5* presents the *in vitro* assembly and *in vivo* implementation results of a synthetic gene oscillator system in mammalian cells, as well as an extensive discussion of experimental results;
- *Chapter 6* presents an overall conclusion of this work and proposes future research directions.

Chapter 2

Background and Literature Review

2. Introduction to gene oscillators

Gene oscillators are one of the most studied types of synthetic gene circuits (O'Brien, Van Itallie, et al., 2012). These circuits are comprised of collections of genes that are expressed periodically. Consequently, the concentration of the proteins encoded by these genes also varies in a periodic manner. Numerous implementations and theoretical designs of synthetic oscillatory gene circuits have been reported over the past years. These range from *in vitro* systems to *in vivo* implementations in both prokaryotic and eukaryotic host cells (O'Brien, Van Itallie, et al., 2012). The first successful attempt to design and implement a synthetic gene oscillator was reported by Elowitz and Leibler (2000), with the development of the "repressilator" (Michael B Elowitz & Leibler, 2000; O'Brien, Van Itallie, et al., 2012). Synthetic gene oscillators have greatly evolved since this report, with more recent reports showing circuits involving more complex regulatory logic and implementation techniques (Atkinson, Savageau, Myers, & Ninfa, 2003; Balagaddé, You, Hansen, Arnold, & Quake, 2005; Danino, Mondragón-Palomino, Tsimring, & Hasty, 2010; Michael B Elowitz & Leibler, 2000; Fung et al., 2005; Montagne, Plasson, Sakai, Fujii, & Rondelez, 2011; O'Brien, Van Itallie, et al., 2012; Stricker et al., 2008; Tigges et al., 2010, 2009).

There are two main reasons why gene oscillators are heavily studied. First, oscillatory patterns of gene expression have been found to play a major role at several levels of organismal development, ranging from neuronal rhythms to metabolic oscillations and circadian clocks (Goldbeter, 1995; Gonze, 2010). Moreover, it has been reported that some stress response signalling pathways respond to external stimuli with transient oscillations (Hoffmann et al., 2002; Kemler & Fontana, 1999; Kobayashi & Kageyama, 2010; Monk, 2003; Nelson et al., 2004). Although the oscillations do not persist indefinitely, the dynamical properties of the oscillations are able to specifically determine the downstream response (Hoffmann et al., 2002; Kobayashi & Kageyama, 2010; Nelson et al., 2004). This has been specifically reported for the NF- κ B, Hes1 and p53 pathways (Monk, 2003). Thus, understanding the regulatory mechanisms behind these naturally occurring gene

oscillators is an important step to understand cellular physiology (O'Brien, Van Itallie, et al., 2012). Furthermore, this knowledge is pivotal for the development of therapies for diseases that affect the dynamical properties of these systems at the regulatory level. The construction of synthetic gene oscillators is useful in this context as it allows for the breakdown of highly complex systems into isolated modules that can be easily analysed.

The second reason why genetic oscillators are heavily studied regards the general understanding of their theoretical behaviour. For instance, synthetic gene oscillators are often simple circuits presenting rich dynamical properties (O'Brien, Van Itallie, et al., 2012). This provides a solid framework for testing mathematical models of gene regulation.

The design principles and dynamical requirements for biochemical oscillators are discussed in this section. Additionally, a comprehensive review of the most iconic synthetic gene oscillators developed to date is provided.

2.1. Fundamental principles underlying the design of biochemical oscillators

Understanding the basis of molecular oscillations in a cellular context is more than a biochemistry or genetic experimentation exercise (Novák & Tyson, 2008). Indeed, oscillators have systems-level properties (such as *entrainment* and *periodicity*) that transcend the properties of their individual molecular components, and involve the full network topology (Novák & Tyson, 2008). These systems-level properties can only be fully understood upon analysing experimental data in combination with a theoretical perspective of the system based on quantitative mathematical modelling and model analysis (Novák & Tyson, 2008). Generally, these models address fundamental concepts of linear systems theory such as *feedback*, *delay* and *bistability*. This theoretical perspective of the dynamics of biochemical oscillators allows us to pinpoint some essential design requirements for these systems.

According to Novak and Tyson (2008), there are four main requirements for the design of biochemical oscillators. First, *negative feedback* is required for a network to be carried back to its 'oscillation start point'. Second, the *time delay* in which the negative feedback is effected must be sufficiently large not to allow the system to settle on a stable steady state. Time delay can be added to a biological system through several mechanisms, such as introduction of physical constraints to biological processes (e.g. bio-molecular transport of chemical species within a cell), existence of numerous reaction intermediates (e.g. such as in an enzymatic pathway), and/or positive feedback-induced dynamical hysteresis (Novák & Tyson, 2008). Third, *non-linearity* in the kinetic laws that describe the reactions in the system is essential to destabilize the steady state. Generally, in mathematical models of gene regulatory networks (GRNs), nonlinearity is fundamentally based on the cooperativity action of transcription factors (i.e. described by a Hill coefficient). Fourth, the reactions that describe the synthesis and degradation of the involved bio-chemical species must occur in *time-scales* that allow the system to generate and sustain oscillations.

Feedback can be described as a propagation of effects among the components of a (biological) system in which one component inevitably predicts an increase or decrease in its own rate of production (E. D. Conrad, Tyson, Rogers, & Day, 1999). The diagram in figure 1 represents a generalized example of a feedback loop. Each of the arrows directed from x_i to x_{i+1} represents either a positive or negative effect on the production rate of x_{i+1} (E. D. Conrad et al., 1999). Indeed, a positive effect can arise from activation of processes leading to the synthesis or from inhibition of processes leading to the degradation of x_{i+1} .

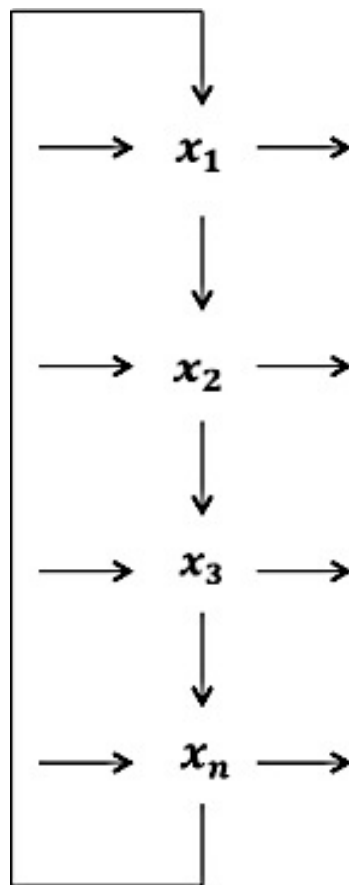


Figure 2.1. Diagram of a feedback loop. Arrows directed in and out of x_i represent processes that contribute to the synthesis and degradation of x_i . Arrows directed from x_i to x_{i+1} represent either a positive or negative effect of x_i on the dynamics of the next variable in the chain.

It would not make sense to begin illustrating the abovementioned concepts with fully detailed models of cellular rhythms for two main reasons. First, naturally occurring oscillatory systems are highly complex and generally coupled to one or more biochemical networks. Indeed, the network coupling properties, as well as the parametric space that shape them, remain obscure for most biological systems. Secondly, systems exhibiting complex cellular rhythms are likely to be overlaid by several control signals thereby disguising their core dynamical properties (Novák & Tyson, 2008). Thus, we will start illustrating the abovementioned concepts in a simple example of a negative feedback-based model of a single protein repressing its own production, by transcription down-regulation (Novák & Tyson, 2008). Let us consider that the rate of change in concentration of protein X , with respect to time, is given by:

$$\frac{dX}{dt} = k_1 \cdot S \frac{K_d^p}{K_d^p + X^p} - k_2 \cdot X \quad (2.1)$$

The first and second terms on the right hand side of this ordinary differential equation define the synthesis and degradation of component X , respectively. The synthesis term is proportional to signal S , multiplied by the Hill function $\frac{K_d^p}{K_d^p + X^p}$, which describes how protein X down-regulates its own transcription. The term K_d is the dissociation constant for the binding of X to the regulatory site upstream of its target gene (i.e. in this case, this gene encodes protein X); the term p is the Hill coefficient, it defines the multimerization index at which protein X binds to its regulatory site (monomer, dimer, tetramer, etc.); the term k_1 is the maximal rate of production of X , proportional to the strength of signal, S . In the degradation term, the constant k_2 is the turnover rate of protein X (Novák & Tyson, 2008). The above equation, like most models describing rate laws in biological systems, follows the *Law of Mass Action*, which states that the rate of change in the concentration of a component in a biochemical system is proportional to some product of

concentrations of the chemical components present in the system (E. D. Conrad et al., 1999).

Simulating this model with a biological plausible choice of parameter values (first proposed by Mackey and Glass in 1977 and adapted by Novák and Tyson, 2008) will show us that the system does not present oscillatory behaviour, or any overshoots and undershoots. Instead, the concentration of protein X will be drawn to its steady-state value X_0 , thus illustrating that the system enters a homeostatic behaviour.

However, let us now consider an explicit time delay in the system, by assuming that the rate of protein synthesis at time t depends on some value of concentration of protein X in the past ($t - \tau$), whereby τ is the time delay required for transcription and translation (Novák & Tyson, 2008). Thus, the above equation can be re-written as follows:

$$\frac{dX(t)}{dt} = k_1 \cdot S \frac{K_d^p}{K_d^p + X(t - \tau)^p} - k_2 \cdot X(t) \quad (2.2)$$

For specific choices of model parameters and time delay, this equation exhibits oscillations. Essentially, this is because the time delay forces the negative feedback to overshoot and undershoot the steady state repeatedly. Effectively, what is meant by “specific choices” of model parameters and time delay is that the kinetic parameters (i.e. S , p , Km , and τ) must satisfy three main constraints (Novák & Tyson, 2008). First, the time delay, τ , must be sufficiently long. That is, there is a minimum time delay value τ_{\min} below which oscillations are impossible, for fixed values of parameters S , p and Km (Novák & Tyson, 2008). Second, there must be sufficient nonlinearity in the system. Essentially, oscillations become “easier” if the nonlinearity (i.e. the p and Km values) in the system increases, which means that τ_{\min} becomes smaller. Third, the model parameters must be in the same timescale, in order to allow for an appropriate balance of the synthesis and degradation terms.

The applicability of these concepts is illustrated in subsections 2.2-2.6 through the review of previous designs and implementations of synthetic gene oscillators that can be found in the literature.

2.2. The Goodwin oscillator: singular self-repression

The Goodwin oscillator was theoretically conceived over 40 years ago – it was the first genetic oscillator to be studied and analyzed (Goodwin Waddington, Conrad Hal, 1963; Purcell et al., 2010). Unsurprisingly, it is also the simplest oscillator developed so far. It comprises a single gene encoding a repressor that down-regulates its own production (figure 2.2).



Figure 2.2. Adapted from Purcell *et al.*, 2010. (Left) Schematic representation of the Goodwin oscillator topology. (Right) The circuit design by Lutz and Bujard (1997) for *in vivo* implementation.

The first biological design of the Goodwin oscillator for *in vivo* implementation was based on a *LacI* gene under the regulation of a $P_{LlacO-1}$ promoter (Lutz & Bujard, 1997), capable of driving transcription at a sufficiently high level when unrepressed (Purcell et al., 2010). It was implemented in *E. coli* cells (Lutz & Bujard, 1997). The *lac* repressor, encoded by the *lacI* gene, down-regulates its own production by binding to the $P_{LlacO-1}$, thus forming a negative feedback loop (Purcell et al., 2010; Stricker et al., 2008). The *in vivo* performance of this system was assessed by placing the same promoter upstream of a gene encoding a fluorescent

protein (i.e. *yemGFP* – monomeric yeast-enhanced green fluorescent protein) (Lutz & Bujard, 1997). Additionally, this design included the presence of *ssra* tags in both the *lac* repressor and *yemGFP* protein (Lutz & Bujard, 1997). These tags are natively recognized by *E.coli* as signals for protein degradation; they were used by Lutz and Bujard (1997) to achieve faster dynamics in the system (Purcell et al., 2010). Indeed, the half-life of wild-type GFP has been reported to be between 15 and 26 hours, depending on the organism it is expressed in (Corish & Tyler-Smith, 1999). Such a long half-life leads to over-accumulation and “masking” of the oscillatory signal, and degradation tags are an efficient way of preventing this (Lutz & Bujard, 1997; Purcell et al., 2010). The system was transformed into *E.coli* cells that were deficient in the *lac* operon in order to minimize interference with native pathways encoded by the host genome (Lutz & Bujard, 1997; Purcell et al., 2010).

Some of the experimental results obtained by Lutz and Bujard (1997) were consistent with simulation results (Stricker et al., 2008): oscillations were irregular, and the period was largely unaffected by the addition of IPTG (i.e. a chemical inducer of *Lacl*, which acts by inhibiting it thus relieving repression) (Purcell et al., 2010). However, the period of oscillations reported by Lutz and Bujard (1997) was 30 minutes – slightly longer than the value obtained from simulations (Stricker et al., 2008). The percentage of cells that showed oscillations was not reported, which means that an assessment of robustness is not possible (Purcell et al., 2010).

In addition, a different *in vivo* implementation of the Goodwin oscillator was reported by Becksei and Serrano (2000) and, in further agreement with simulations, has also validated the mathematical results regarding the nonlinearity requirements governing repression dynamics that lead to oscillations (Purcell et al., 2010). Unlike the Lutz and Bujard (1997) implementation, which had *Lacl* as the negative feedback effector, the Becksei and Serrano implementation (2000) was based on the TetR repressor (Purcell et al., 2010). One of the essential differences between these two transcriptional repressors is that *Lacl* functions as a tetramer, whereas TetR functions as a dimer. An important implication of this is that TetR-mediated repression will thus be described by a lower Hill coefficient than *Lacl*-mediated repression. Indeed, the implementation using TetR as the negative feedback effector

did not oscillate, showing highly stable dynamics instead (Becskei & Serrano, 2000; Purcell et al., 2010).

2.3. Repressilators: delayed negative feedback in multi-repressor systems

A repressilator can be thought of as an extension of the Goodwin oscillator (Purcell et al., 2010). It can be defined as “a regulatory network of one or more genes, with each gene repressing its successor in the cycle” (Müller et al., 2006; Purcell et al., 2010). The term “repressilator” was first used by Elowitz and Leibler (2000) to describe a system containing a 3-gene closed feedback loop (Michael B Elowitz & Leibler, 2000; Purcell et al., 2010). Thus, a repressilator containing a single gene would be the abovementioned Goodwin oscillator.

The study of repressilators began over 30 years ago as a logical and conceptual extension to the Goodwin oscillator, using Boolean-like simulations (Fraser & Tiwari, 1974; Purcell et al., 2010). Subsequently, continuous models were studied using DDEs (Smith, 1987; Wang, Jing, & Chen, 2005), ODEs (Michael B Elowitz & Leibler, 2000; Müller et al., 2006; Samad, Vecchio, & Khammash, 2005), and discrete stochastic simulations based on the Gillespie algorithm (Purcell et al., 2010; Rajala, Häkkinen, Healy, Yli-Harja, & Ribeiro, 2010).

The first *in vivo* implementation of a multi-gene repressilator was published by Elowitz and Leibler (2000). The network was implemented in *E.coli*. This system was based on a three-gene topology, namely LacI (from *E.coli*), TetR (from the Tn10 transposon), and *ci* from the λ phage (Michael B Elowitz & Leibler, 2000; Purcell et al., 2010). The regulatory logic of this network and its effectors are illustrated in figure 2.3. The LacI repressor downregulates *tetR* expression, TetR downregulates *ci* expression, and *ci* downregulates *lacI* expression, thus completing the negative feedback.

The experimental results showed that 40% of cells exhibited oscillations, with a period of 60 minutes (± 40) – in agreement with deterministic simulations – and

significant amplitude variations – in agreement with stochastic simulations (Michael B Elowitz & Leibler, 2000). The oscillations, reported by GFP fluorescence, failed to return to a zero-level, as opposed to what was predicted in simulations (Michael B Elowitz & Leibler, 2000). This has been theorized to be due to the higher stability of GFP relative to the rapidly degrading repressor proteins in the system (Purcell et al., 2010). Indeed, it has been previously shown that the addition of *ssra* tags to the λ *cl* gene reduces the half-life of *cl* from more than 60 minutes to *c.* 4 minutes (Keiler & Sauer, 1996). The addition of these same tags to the *gfp* gene reduces the GFP half-life to 30 to 40 minutes (Andersen et al., 1998).

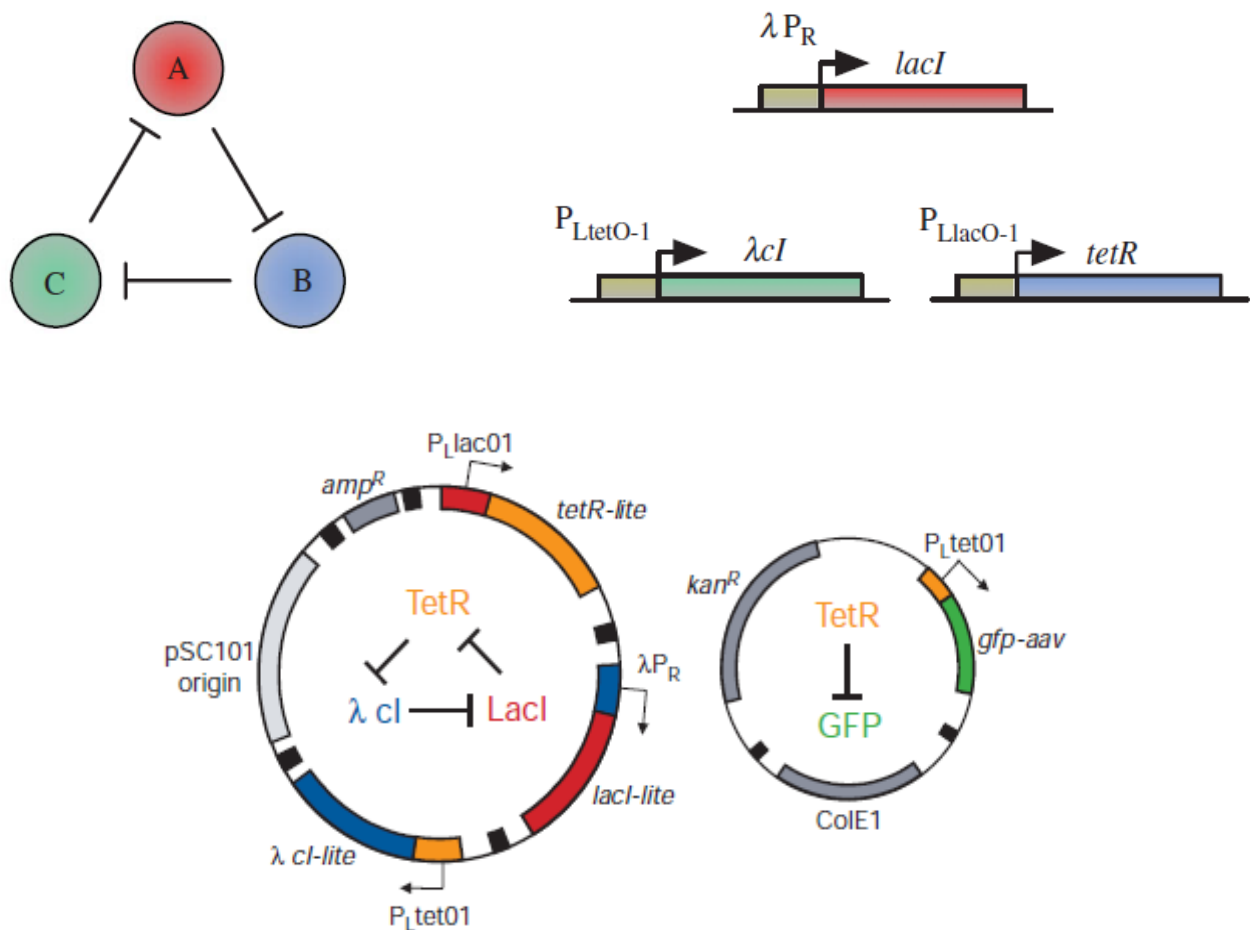


Figure 2.3. Adapted from Elowitz and Leibler, 2000; and Purcell *et al.*, 2010. (Top) Diagrammatic representations of the network topology, promoters and genes used in the Elowitz and Leibler repressilator. (Bottom) Representation of the plasmids containing the system core regulatory modules and the reporter construct.

Despite the variability in amplitude, oscillations in progeny were correlated for a significant amount of time after cell division, thus showing that the network state is passed on to daughter cells (Michael B Elowitz & Leibler, 2000; Purcell et al., 2010). The timing of oscillations, as well as frequency and amplitude, were not correlated with cell division dynamics (Michael B Elowitz & Leibler, 2000). This shows that the repressilator dynamics were de-coupled from the cell division-regulating pathways. However, entry of *E.coli* cells in stationary phase arrested oscillations, indicating that the dynamics of the system are coupled to the global metabolic and growth cellular pathways. A possible explanation for this interference with the system is the concentration increase in σ^{38} transcription initiation factor during stationary phase (Purcell et al., 2010). This would compete with σ^{70} (i.e. responsible for transcription initiation from the repressilator promoters) for RNA polymerase and may affect transcription dynamics significantly (Purcell et al., 2010).

Even though it was conceived after the Goodwin oscillator, the Elowitz and Leibler (2000) repressilator was considered the first successful *in vivo* implementation of a synthetic gene oscillator (Purcell et al., 2010). However, the low percentage of cells exhibiting oscillations suggests that this implementation lacked robustness (Michael B Elowitz & Leibler, 2000; Purcell et al., 2010). Although not fully understood, noise may play a role on this (Michael B Elowitz & Leibler, 2000; O'Brien, Van Itallie, et al., 2012; Purcell et al., 2010). As mentioned above, possible interference with native, global cellular pathways could also explain the lack of robustness in this implementation.

In addition, theoretical studies have shown that the addition of positive feedback loops to any of the repressilator components increases the parametric region over which the system exhibits oscillations (O'Brien, Van Itallie, et al., 2012; Purcell et al., 2010; Tsai et al., 2008; Wang et al., 2005). This addition of positive feedback loops to the system has also been shown to allow the repressilator to oscillate with a wide range of different frequencies for a given amplitude, thus increasing its tunability (Tsai et al., 2008).

2.4. The Atkinson oscillator: amplified negative feedback

The abovementioned genetic oscillators in 2.2 and 2.3 are exclusively formed by repressive links. The next logical step in the area of gene oscillator design was to incorporate positive interactions between genes and thus obtain new system topologies. The simplest of these amplified negative feedback topologies contains two genes, say *A* and *B*. Gene *A* upregulates itself and also gene *B*, while gene *B* downregulates gene *A*, effecting the negative feedback (Purcell et al., 2010). Three distinct versions of this topology have been investigated in the literature: (a) repression by transcriptional control, (b) repression by proteolysis, and (c) repression through dimerization-based sequestration (Atkinson et al., 2003; Barkai & Leibler, 2000; E. D. Conrad et al., 1999; E. Conrad, Mayo, Ninfa, & Forger, 2008; Guantes & Poyatos, 2006; Purcell et al., 2010). The only *in vivo* implementation of this class of oscillators was the Hopf-driven Atkinson oscillator (2003), a system based on repression by transcriptional control.

This oscillator consists of two modules. One module encodes an activator, NRI, the other encodes a repressor, *lacI* (Atkinson et al., 2003). Figure 2.4 illustrates the topology and genetic components of this system.

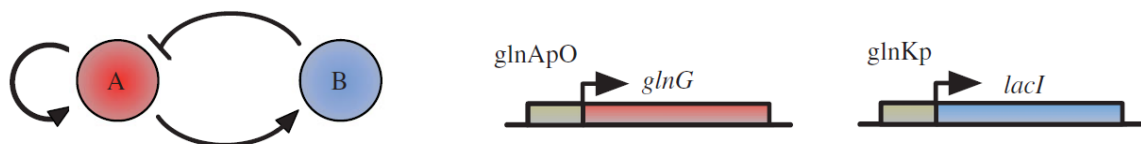


Figure 2.4. Adapted from Purcell *et al.*, 2010. (Left) Diagrammatic representation of the topology Atkinson oscillator topology. (Right) Illustration of the activator (red) and repressor modules (blue).

The activator module was constructed by fusing the *glnG* gene (encoding NRI), to an upstream control sequence based on a *glnA* promoter (Atkinson et al., 2003; Purcell et al., 2010). This promoter was regulated by two upstream, high-affinity, adjacent binding sites for NRlp (i.e. phosphorylated NRI), forming an enhancer and two *lac* operators – one upstream and one downstream of the enhancer (Atkinson et al., 2003; Purcell et al., 2010). Indeed, the design of this system intended to utilize DNA looping as a contributor to regulatory dynamics: during repression, the LacI protein induces loop formation by binding to the two operator sites, ensuring stable repression; during activation, NRlp interacts with promoter-bound RNA polymerase also through a DNA loop (Atkinson et al., 2003). These two loops act antagonistically, as their formation is mutually exclusive (Atkinson et al., 2003; Purcell et al., 2010).

The repressor module was built by fusing the *lacI* gene to a regulatory control sequence based on the *glnK* promoter and upstream enhancer formed of adjacent low-affinity and high-affinity NRlp binding sites (Atkinson et al., 2003; Purcell et al., 2010). Thus, *lacI* expression from the *glnK* promoter is only fully induced upon high concentrations of NRlp (Atkinson et al., 2003). The system was transformed in *E.coli* cells that were *lacI* deficient, thus ensuring that the only source of LacI repressor was the repressor module (Atkinson et al., 2003). On a population level, LacI dynamics were assayed through β -galactosidase activity, which was modulated by the interaction of LacI with the chromosomal-encoded *lacYZA* operon (Atkinson et al., 2003). At a single-cell level, dynamics were monitored by CFP (Cyan Fluorescent Protein) under regulation of a LacI-repressible promoter (Atkinson et al., 2003).

In vivo β -galactosidase dynamics closely matched modelling predictions (Purcell et al., 2010). Three damped oscillations were initially observed, as theoretically predicted (Atkinson et al., 2003). This was further increased to four damped oscillations by moving the activator module to an origin of replication where copy numbers were four times higher (Atkinson et al., 2003; Purcell et al., 2010). Additionally, amplitude and frequency were comparable between experimental results and modelling predictions; this was highlighted by investigating the effect of cellular doubling time (through controlled growth medium composition) on oscillator

dynamics (Atkinson et al., 2003). Single-cell observations matched those at the populational level, showing that damping is not a consequence of individual cells oscillating out of phase (Purcell et al., 2010). As in the Elowitz and Leibler repressilator implementation, the percentage of oscillating cells was not reported by Atkinson *et al.* (2003), meaning that robustness cannot be assessed.

Although this implementation showed damped oscillations, the significant agreement between experimental and theoretical results highlighted that simple mechanistic, dynamical models can capture the quantitative and qualitative behaviour of synthetic gene networks (Purcell et al., 2010). However, the quantitative agreement was only considerable at the population level. Indeed, further analysis is required to ascertain the applicability of such mathematical models to individual cell behaviour (Guantes & Poyatos, 2006; Purcell et al., 2010). Such an analysis would most likely help reveal the role of noise in the behaviour of synthetic gene networks, so far assumed to be of little importance (Guantes & Poyatos, 2006; Purcell et al., 2010).

2.5. The Fussenegger oscillators

The Fussenegger oscillators (published by Tigges *et al.*, 2009, 2010) are the only successful implementations of synthetic gene oscillators in eukaryotic hosts (Purcell et al., 2010). Indeed, Tigges *et al.* have chosen to implement their system in Chinese hamster ovary (CHO) cells (Tigges et al., 2009). The design of synthetic gene networks for eukaryotic cell implementation represents an important step forward in synthetic biology, as these cells exhibit a much more complex molecular biology landscape than prokaryotic cells.

The original Fussenegger oscillator design is based on two genes encoding transactivators PIT (Pristinamycin-Induced transactivator) and tTA (Tetracycline-dependent TransActivator), with both antisense and sense transcription occurring at the *tTA* gene (Purcell et al., 2010; Tigges et al., 2009). The components and topology of this system are illustrated in figure 2.5.

The sense tTA transcript is translated and the resulting protein upregulates its own expression (completing a positive feedback loop), while also upregulating expression of the *pit* gene. PIT upregulates antisense expression of the *tTA* gene (Tigges et al., 2009). The resulting transcript is not translated. Instead, it hybridizes with the sense tTA transcript, thus downregulating tTA expression at the translation level (Purcell et al., 2010; Tigges et al., 2009). This downregulating effect of antisense tTA mRNA completes the negative feedback. Hence, the Fussenegger oscillator is an amplified negative feedback oscillator (Purcell et al., 2010). However, unlike other previously designed amplified negative feedback oscillators (e.g. the Hopf-driven Atkinson oscillator), it contains an additional step before repression can occur, increasing the delay in negative feedback (Purcell et al., 2010).

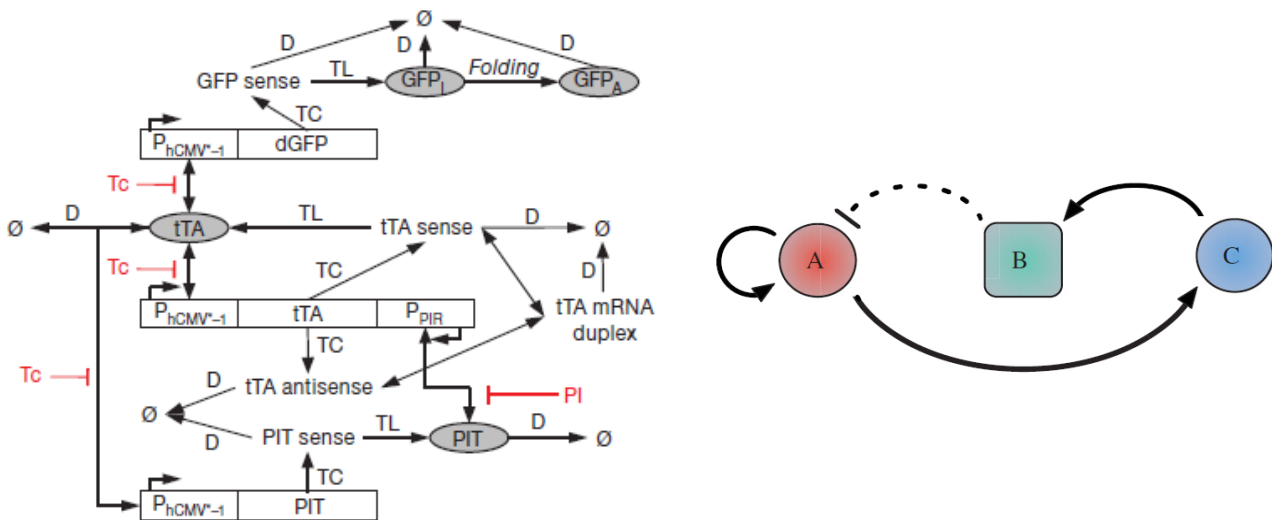


Figure 2.5. Adapted from Tigges *et al.* (2009) and Purcell *et al.* (2010). (Left) Interaction map of the components that comprise the original Fussenegger oscillator design. (Right) Diagrammatic representation of the system topology, where A and C represent tTA and PIT, respectively; B represents antisense tTA mRNA.

The system was constructed by placing the sense expression of tTA and PIT under the regulation of tTA-responsive promoters (i.e. $P_{hCMV^{*-1}}$) (Purcell et al., 2010; Tigges et al., 2009). Production of the antisense tTA transcript was put under regulation of a P_{PIR} promoter (PIT-responsive), a chimeric promoter constructed by fusing the P_{PTR} promoter from *Streptomyces pristinaespiralis* with the *Drosophila melanogaster* minimal hsp70 promoter (Fussenegger et al., 2000; Weber, Kramer, Fux, Keller, & Fussenegger, 2002). Expression of gene encoding fast-degrading GFP was also put under the regulation of $P_{hCMV^{*-1}}$, allowing monitoring of system dynamics (Purcell et al., 2010; Tigges et al., 2009). Additionally, the activity of transactivators tTA and PIT could be inhibited by tetracycline and pristinamycin, respectively (Tigges et al., 2009). The *tTA*, *pit* and *gfp* genes were each placed on different plasmids and further transfected in CHO-K1 cells (Tigges et al., 2009).

Sustained, undamped oscillations were observed, with periods of 170 ± 71 minutes (Purcell et al., 2010; Tigges et al., 2009). This confirmed the predicted effectiveness of considering extra delays in negative feedback-based networks (Purcell et al., 2010). However, the fact that eukaryotic implementations are predicted to have longer delays (e.g. in transcription initiation, diffusion, translation) may also play a role in the overall negative feedback delay (Purcell et al., 2010). Like in past oscillator implementations, significant cell-to-cell variability was observed, as predicted by stochastic simulations based on the Gillespie algorithm (Purcell et al., 2010). Additionally, the nonlinear relationship between amplitude and period, and relative gene dosage, also matched predictions from mathematical modelling, showing that the system could be predictably tuned (Purcell et al., 2010). In further agreement with predictions from theoretical work, tetracycline-mediated inhibition of tTA abolished oscillations (Tigges et al., 2009). However, in contrast with simulations, pristinamycin-mediated inhibition of PIT had a significant dynamical effect, leading to increasing fluorescence intensity, instead of oscillations (Purcell et al., 2010). Effectively, this occurs because pristinamycin-mediated inhibition of PIT prevents upregulation of antisense tTA mRNA, i.e. the negative feedback effector. This means that the *tTA* gene becomes unrepressed (while still in a positive feedback loop), which leads to an accumulation of its product and, consequently, GFP.

Additionally, Tigges *et al.* (2010) have constructed a 'low-frequency' variant of the Fussenegger oscillator, using direct interference by siRNAs instead of the indirect antisense mRNA-based repression mechanism (Purcell *et al.*, 2010; Tigges *et al.*, 2010). Although a step in repression has been removed in this alternative design, the delay in negative feedback was still enough to support sustained oscillations (Tigges *et al.*, 2010). Indeed, robust oscillations, with a period of *c.* 26 hours, were observed (Tigges *et al.*, 2010). To date, this is the longest oscillatory period of any synthetic gene oscillator (Purcell *et al.*, 2010). Unlike the original Fussenegger oscillator, the 'low-frequency' variant could not be tuned by altering relative plasmid dosages (Tigges *et al.*, 2010).

The implementation of both versions of the Fussenegger oscillator represent an important step forward in synthetic biology: these were the first oscillators to be implemented in eukaryotic systems and to make use of sense-antisense RNA interactions (Purcell *et al.*, 2010). However, these oscillators are not robust, having at most 18 % of cells exhibiting oscillations (Purcell *et al.*, 2010).

2.6. The Smolen / Hasty oscillator: design and implementation of robust oscillators

None of the abovementioned gene oscillators are robust. Indeed, the percentage of cells exhibiting oscillations is either not reported, or low (Purcell *et al.*, 2010). Effectively, robustness is a fundamental part of implementation if synthetic gene oscillators are to be coupled to large synthetic networks, or interact/regulate with natural biological systems (Lu *et al.*, 2009; Purcell *et al.*, 2010).

The Smolen oscillator comprises two genes (Smolen *et al.*, 1998). One of the genes (gene A) upregulates its own expression, while also upregulating expression of the other gene (gene B) (Smolen *et al.*, 1998). Gene B represses its own expression and that of gene A (Smolen *et al.*, 1998). In fact, the self-repression loop effected by the expression product of gene B differentiates the Smolen oscillator from amplified negative feedback oscillators in topological terms (Purcell *et al.*, 2010). The Smolen

oscillator was extensively studied from a theoretical point of view: ODE-based models can be found in Smolen *et al.* (1998), Hasty *et al.* (2002) and Stricker *et al.* (2008). Additionally, DDE- and SDE-based models were also presented in Smolen *et al.* (1998, 1999) and Wang *et al.* (2005), respectively. Figure 2.6 illustrates the Smolen oscillator topology, as well as the components used for its *in vivo* implementation (accomplished by Stricker *et al.*, 2008).

The *in vivo* implementation was highly based on the hybrid promoter $P_{lac/ara-1}$, first developed by Lutz and Bujard (1997). $P_{lac/ara-1}$ contains two adjacent AraC binding sites upstream of the promoter region (allowing AraC-mediated transactivation) and three LacI operator sites, two upstream and one downstream of the promoter region (Stricker *et al.*, 2008). The *araC* and *lacI* genes were independently placed under the control of $P_{lac/ara-1}$ promoters (Purcell *et al.*, 2010; Stricker *et al.*, 2008). Additionally, the *yemGFP* gene was also controlled by a $P_{lac/ara-1}$ promoter, allowing system performance to be monitored (Stricker *et al.*, 2008). All protein components contained *ssra* tags, allowing faster degradation times (Stricker *et al.*, 2008). The activator and repressor modules were placed in separate plasmids and transformed into *araC*- and *lacI*-deficient *E.coli* cells (Stricker *et al.*, 2008), thus minimizing interference with the host genome.

Over 99 % of cells exhibited oscillations, with a period of *c.* 40 minutes (Purcell *et al.*, 2010; Stricker *et al.*, 2008). This was in agreement with simulations and clearly demonstrated robustness (Purcell *et al.*, 2010). Oscillatory state was transmitted to progeny and population-level oscillatory synchrony was lost after a few periods (Purcell *et al.*, 2010; Stricker *et al.*, 2008). In further agreement with simulations, the oscillations were tunable, showing that the period could be altered between 13 and 58 minutes over a wide range of IPTG and arabinose concentrations, growth temperatures, and media sources (Stricker *et al.*, 2008).

Individual cells showed an unexpected gradual increase in period, possibly due to the dynamics of the reporter used (Purcell *et al.*, 2010; Stricker *et al.*, 2008). This was not accounted for in the mathematical models (Stricker *et al.*, 2008). Although cell doubling time in Luria Broth and minimal media differed by *c.* 1 hour,

the period of oscillations was comparable, demonstrating that system dynamics were de-coupled from cell cycle dynamics (Purcell et al., 2010; Stricker et al., 2008).

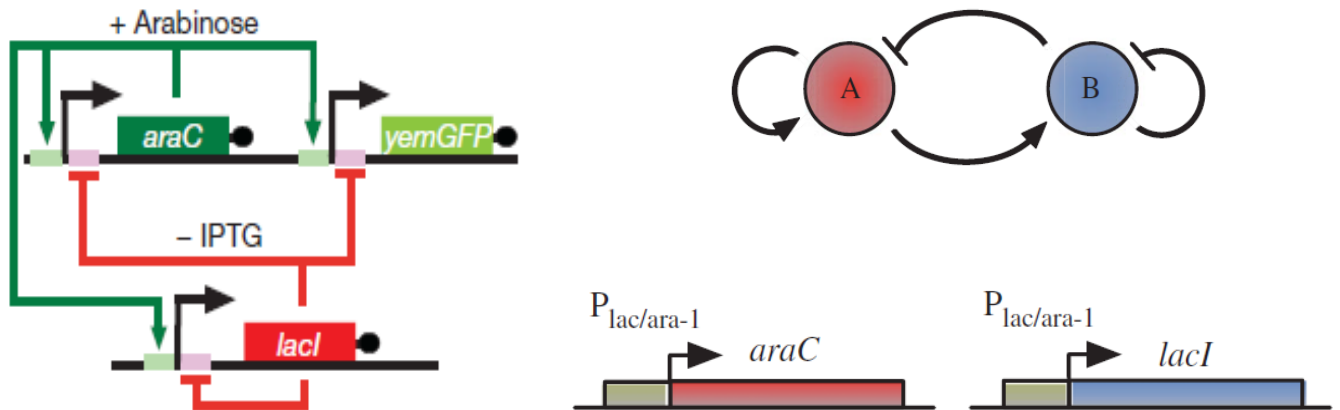


Figure 2.6. Adapted from Stricker *et al.* (2008) and Purcell *et al.* (2010). (Left) Diagrammatic representation of the gene network constructed by Stricker *et al.* (2008) for the *in vivo* implementation of the theoretically designed Smolen oscillator (Smolen et al., 1998). (Right) The Smolen oscillator topology (top); the activator and repressor modules used for *in vivo* implementation (Stricker *et al.*, 2008; Purcell *et al.*, 2010).

As mentioned above, the *in vivo* implementation of the Smolen oscillator is highly tunable and robust. These are typical characteristics of natural oscillators: tunability provides utility, while robustness provides reliability (Purcell et al., 2010). This is encouraging for the future construction of synthetic gene oscillators (Purcell et al., 2010). Indeed, the construction of the Smolen oscillator was followed up on by the assembly of a synchronized quorum of genetic clocks: the only synchronized synthetic gene network implemented to date (Danino et al., 2010). Studies on time delay are inconclusive at this stage, but the successful implementation of the Fussenegger oscillators suggests that an eukaryotic implementation of the Smolen oscillator may be successful (Purcell et al., 2010; Stricker et al., 2008).

2.7. The need for new oscillator designs

A path of increasing design complexity has been taken since the early theoretical design of the Goodwin oscillator to the recent construction of the Smolen oscillator and the sophisticated Fussenegger oscillators (Purcell et al., 2010). Mathematical models describing the dynamics of oscillatory gene expression have evolved thanks to the experimental construction of the systems reviewed above (O'Brien, Van Itallie, et al., 2012). However, there are still new topologies and implementations to be theoretically tested and experimentally constructed. This will provide considerable insight in the task of understanding and modelling genetic regulatory processes leading to oscillations (O'Brien, Van Itallie, et al., 2012). Additionally, comparisons between previously designed oscillators remain hindered by the lack of standardization and characterization of biological parts (Purcell et al., 2010). The design complexity that arises from this is further amplified by the variability in implementation features observed in the oscillators constructed so far, i.e. different biological components, topologies and host organisms. A summary of the implementation features reviewed throughout this chapter is presented in table 2.1.

In vivo implementations have almost exclusively been in prokaryotic hosts (i.e. *E.coli*), with the exception of Fussenegger oscillators (Purcell et al., 2010). Indeed, the molecular biology landscape of prokaryotes is simpler and easier to manipulate than that of eukaryotes. Additionally, the biological components required to assemble synthetic gene networks in prokaryotes are better characterized and standardized in comparison to eukaryotes.

Oscillatory periods vary between 13 minutes and 26 hours (for the Smolen and low-frequency Fussenegger oscillators, respectively). The reason for this significant variability is not yet well understood, and it has been hypothesized that the interplay of topological structure with noise plays a determining role in determining oscillatory period and amplitude (Purcell et al., 2010). In addition, period and amplitude of oscillations have been shown to be prone to tuning via several mechanisms, e.g. temperature, chemical inhibitors, medium composition, gene dosage. In terms of robustness (herein assessed as the percentage of host cells

exhibiting oscillatory behaviour), only the Smolen oscillator (implemented by Stricker *et al.*, 2008) can be considered to be robust (Purcell *et al.*, 2010). To date, there are still no available mammalian oscillator designs that are both robust and highly tunable. The main objective of the next chapters is to propose and highlight the theoretical and biological design features of such a system.

Table 2.1. Comparison between different *in vivo* implementations of synthetic gene oscillators.

Oscillator	Robustness (in % cells exhibiting oscillations)	Period (min)	Tunability	Host	Number of genes
Goodwin	Not reported	30	No	<i>E.coli</i>	1
Repressilator (Elowitz and Leibler)	40	160±40	Coupled to cell growth regulation	<i>E.coli</i>	3
Atkinson	Not reported	600 – 1200 (damped oscillations)	Yes (Period and amplitude, via cell doubling)	<i>E.coli</i>	2
Fussenegger	Not reported	170±71	Yes (Period and amplitude, via gene dosage)	CHO-K1	2
Fussenegger (low-frequency)	18	1560±510	No	CHO-K1	2
Smolen	99	13 - 58	Yes (Period, via IPTG and arabinose)	<i>E.coli</i>	2

Chapter 3

Design of a novel mammalian synthetic gene oscillator

3.1. Introduction

Synthetic gene oscillators have systems-level properties such as *frequency*, *amplitude*, *feedback* and *delay*. As such, the design of such a system requires more than a genetic engineering or molecular biology experimentation exercise (Novák & Tyson, 2008). Indeed, in order to fully understand and control these system-level properties one must have a theoretical perspective of the system based on quantitative mathematical modelling and analysis, in addition to experimental testing (Novák & Tyson, 2008). As established in 2.1, there are four fundamental requirements for the design of an oscillator:

- *Negative feedback* is required for a system to be carried back to its “starting point”;
- *Time delay* in the negative feedback prevents the system from settling on a stable steady state;
- *Non-linearity* in kinetic laws describing the system provide additional destabilization of the steady state;
- *Time-scales* of the synthesis and degradation reactions must be balanced across all system components;

Thus, the theoretical design of a synthetic gene oscillator must be based on a mathematical model that reflects these requirements while providing a comprehensive description of gene regulation processes within the system. The development of such a model for a novel synthetic gene oscillator is presented in this chapter.

In synthetic biology, descriptive mathematical models are unable to provide a fully detailed representation of all the processes happening within and around systems. This is largely due to the complex molecular biology landscape of cells, and the parametric uncertainty that they involve. Thus, the development of a mathematical model alone would not provide any insight regarding this uncertainty. Indeed, in order to explore the uncertain parametric space of such a mathematical

model, and have insight as to how it affects system dynamics, one must resort to a suitable type of model analysis. Bifurcation analysis pinpoints qualitative changes in the dynamic behaviour of a model when parameter values are changed (Strogatz, 1994): it allows one to correlate parametric space regions to dynamic behaviours. Because synthetic gene oscillators present complex dynamical properties that can translate into different dynamical behaviours, bifurcation analysis represents an invaluable tool for their design.

This chapter addresses the *in silico* design of a novel synthetic gene oscillator from a topological, dynamical and biological perspective. The first section focuses on presenting the structure and topology of the system, as well as its components. This is followed by the development of a lumped parameter kinetic ODE model that describes the proposed system and its dynamics. A brief review of bifurcation theory is then provided, followed by a comprehensive presentation and discussion of deterministic simulation and bifurcation analysis results. Concluding remarks regarding the design of the proposed system are presented at the end of this chapter.

3.2. System structure and topology

There is considerable variation in topology, structure and system components between past implementations of synthetic gene oscillators, from the simple *lacI*-based Goodwin oscillator to the complex Fussenegger mammalian oscillators (Novák & Tyson, 2008; O'Brien, Van Itallie, et al., 2012; Purcell et al., 2010; Tigges et al., 2010, 2009). Furthermore, alternative topologies have been theoretically proposed but not implemented *in vivo*, e.g. 4-component oscillators, chaotic oscillators (O'Brien, Van Itallie, et al., 2012; Purcell et al., 2010).

With this extensive design availability it becomes essential to define topological and structural selection criteria that pre-determine implementation feasibility before choosing a specific system topology. To this end, several synthetic gene oscillator topologies were characterized below, according to their suitability for mammalian cell implementation:

- 1-component oscillators are based on the Goodwin design (Goodwin Waddington, Conrad Hal, 1963). This system is known to generate irregular, non-robust oscillations in *E.coli* if its implementation is based on the Lac repressor system (Lutz & Bujard, 1997). Implementations based on other repressor systems are known to disrupt oscillatory behaviour, showing highly stable dynamics instead (Becskei & Serrano, 2000). Being the simplest type of synthetic gene oscillators, 1-component oscillators offer minimal implementation risk in comparison to other topologies. However, it has been pointed out that this topology would not be a good candidate for mammalian cell implementation (O'Brien, Van Itallie, et al., 2012). This is mainly due to the increased delay in mammalian gene expression: the negative feedback effected by LacI would not occur fast enough to avoid a stable steady state (Novák & Tyson, 2008; O'Brien, Van Itallie, et al., 2012; Purcell et al., 2010);
- 2-component oscillators have been shown to work in *E.coli* cells, mainly through the implementations carried out by Atkinson *et al.* (2003) and Stricker *et al.*

(2008) (Atkinson et al., 2003; Smolen et al., 1998; Stricker et al., 2008). Although these implementations consisted of 2 core components, their topologies are slightly different: the Hasty oscillator contains an additional negative feedback loop (Stricker et al., 2008). Indeed, the different dynamic behaviours observed in these two systems (transient *versus* sustained oscillations, respectively) are a reflection of this topological difference (Purcell et al., 2010). Both implementations relied on the function of well-characterized, prokaryotic, hybrid promoters that are responsive to more than one transcription factor (Atkinson et al., 2003; Purcell et al., 2010; Stricker et al., 2008). 2-component oscillators have been previously proposed for eukaryotic implementation (O'Brien, Van Itallie, et al., 2012; Purcell et al., 2010). However, it has been shown that, with an appropriate parameter selection, two component mammalian synthetic gene oscillators would not be able to generate enough negative feedback delay in order to sustain oscillations (Novák & Tyson, 2008; O'Brien, Van Itallie, et al., 2012; Purcell et al., 2010). Furthermore, hybridized promoters are not as well characterised and standardised in eukaryotes as they are in prokaryotes. This poses a considerable implementation risk because promoters and their respective operators / transcription factor binding sites represent key features in system design: they couple different system components through gene expression regulation;

- 3-component oscillators have been previously designed and implemented in both prokaryotic and eukaryotic hosts (Purcell et al., 2010). This is the case of the repressilator and the Fussenegger oscillators, respectively (Michael B Elowitz & Leibler, 2000; Tigges et al., 2010, 2009). The repressilator has been previously proposed for eukaryotic implementation (Purcell et al., 2010). However, for a plausible choice of mammalian cell parameters the system presents stable dynamics instead of sustained oscillatory behaviour, possibly due to increased delays in every step of repression throughout the system (O'Brien, Van Itallie, et al., 2012). Additionally, the dynamic mechanisms leading to the appearance of oscillations in the repressilator are not yet well understood (Kuznetsov &

Afraimovich, 2012). Indeed, it has been shown that the parameter regions in which limit cycles appear vary greatly if saturation, instead of linear, dynamics are assumed for protein degradation (Kuznetsov & Afraimovich, 2012). On the other hand, the Fussenegger oscillators are the only synthetic gene oscillators to be successfully implemented in mammalian cells. The topology of these systems is different from that of the repressilator: instead of comprising three consecutive repressive links, it consists of two up-regulating modules and one repressive link, mediated by either anti-sense mRNA synthesis or siRNA expression (Tigges et al., 2010, 2009). Although presenting a robustness of c. 18%, building upon the Fussenegger designs represents a lower implementation risk than any of the other previous topologies and systems due to their mammalian cell-optimized features, e.g. well-characterized transcription factors for mammalian cells, ubiquitin-tagged GFP reporter, orthogonality;

- Synthetic gene oscillators with 4 or more components, such as chaotic and multicellular oscillators, have also been proposed for both prokaryotic and eukaryotic implementation, but remain entirely theoretical to date with the exception of a prokaryotic synchronized genetic clock quorum developed by Danino *et al.* (2010) (Danino et al., 2010; O'Brien, Van Itallie, et al., 2012; Purcell et al., 2010). This is likely due to the implementation complexity they involve. Indeed, most synthetic gene oscillators are implemented in host cells through plasmid transformation or transfection (Atkinson et al., 2003; Michael B Elowitz & Leibler, 2000; Stricker et al., 2008; Tigges et al., 2010, 2009). For mammalian systems this is usually done following a one-component per plasmid rule, in order to achieve increased system modularity and experimental control (Tigges et al., 2010, 2009). Thus, there is a considerable implementation complexity and risk in designing a mammalian synthetic gene oscillator with any topology that is based on four or more system components.

Design of a novel mammalian synthetic gene oscillator

Taking the implementation risk characterization presented above into account, the mammalian synthetic gene oscillator design to be proposed herein is a 3-component delayed negative feedback system that follows the same topological features as the Fussenegger oscillators. However, the structure of the proposed system is based on three different genes instead of two genes with an additional anti-sense mRNA or siRNA module. This core structure has been suggested as a promising candidate for both prokaryotic and eukaryotic implementation (E. D. Conrad et al., 1999; Novák & Tyson, 2008; O'Brien, Van Itallie, et al., 2012; Purcell et al., 2010).

The core components of the system are *tTA*, *gal4* and *lacI*; the reporter module is an ubiquitin-tagged *gfp* construct. The three genes that make up the core of the system were selected because they are the most well-characterized transcription-factor encoding genes that are unpatented and functional in mammalian cells (Aubrecht, Manivasakam, & Schiestl, 1996; Cronin, Gluba, & Scrable, 2001; Gossen et al., 1995; Kakidani & Ptashne, 1988; Sadowski, Bell, Broad, & Hollis, 1992; Tigges et al., 2009; Webster, Jin, Green, Hollis, & Chambon, 1988). There are several alternative transcription factor encoding genes that are well-characterized for mammalian synthetic biology research (Kis et al., 2015). However, these are patented and, as such, were not considered for the design of this system (Aubel & Fussenegger, 2010; Fussenegger et al., 2000; Heng, Aubel, & Fussenegger, 2014; Nielsen et al., 2014; Weber & Fussenegger, 2009, 2010).

The structure and topology of the proposed mammalian synthetic gene oscillator are illustrated in figure 3.1. The design proposed herein is based on a delayed negative feedback loop: expression of the *tTA* gene leads up-regulation of *gal4* (at a *tetO* operator site), which in turn up-regulates expression of *lacI* (at an UAS– Upstream Activating Sequence – site). Negative feedback, essential for preventing the system from settling into a stable steady state, is carried out by LacI-mediated down-regulation of the *tTA* gene at a *lacO* operator site. In addition to its core regulatory role, tTA also up-regulates expression of an ubiquitin-tagged *gfp* gene (at a *tetO* operator site), allowing system dynamics to be monitored. The delay in negative feedback is essentially based on the intermediate tTA-*gal4* regulatory interaction: the LacI repressor is only produced after tTA-mediated up-regulation of

gal4, and the subsequent Gal4-mediated upregulation of *lacI*. The activities of LacI and tTA can be inhibited by IPTG and tetracycline (or doxycycline), respectively. This provides a potential source of frequency tuning similar to that in the Fussenegger oscillators (Tigges et al., 2010, 2009).

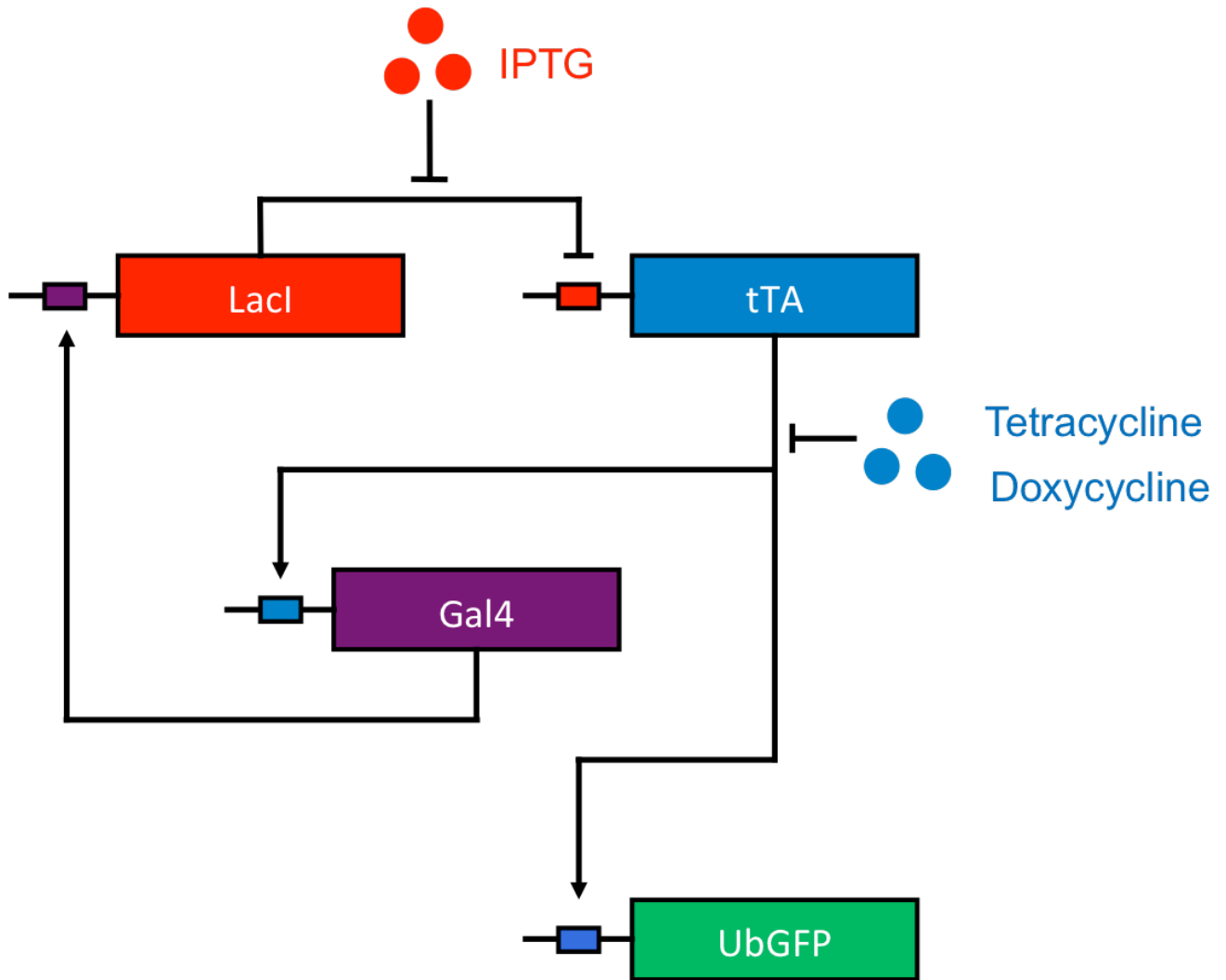


Figure 3.1. Diagrammatic representation of the proposed oscillator design. The core regulatory part of the system comprises three genes: *lacI*, *tTA* and *gal4*. The *tTA* and *gal4* genes encode the tetracycline-dependent (tTA) and Gal4 transactivators, respectively. The *lacI* gene encodes the LacI repressor. The *gfp* gene allows monitoring of dynamics, through the expression of an ubiquitin-tagged variant of Green Fluorescent Protein (GFP). Blunt-ended arrows represent down-regulation, whereas normal arrows represent up-regulation.

This network topology has been theoretically predicted to sustain

oscillations, and is thus regarded in the literature as a good candidate for *in vivo* implementation (E. D. Conrad et al., 1999; Novák & Tyson, 2008). The addition of a positive feedback loop to this topology has been predicted to further increase the size of parametric regions corresponding to oscillatory behaviour (Novák & Tyson, 2008; Purcell et al., 2010). However, this represents a source of complexity that might affect system robustness due to the inherent complexity of implementing synthetic hybrid promoters in mammalian cell systems (Purcell et al., 2010).

3.3. Introduction to an ODE-based model of the proposed synthetic gene oscillator

A lumped parameter kinetic ODE model was developed in order to predictively test if the proposed system is able to exhibit and sustain oscillatory behaviour with a realistic choice of parameter values. Moreover, it provides a modelling framework capable of predicting dynamic behaviour changes in response to specific biochemical stimuli, namely the concentrations of tetracycline and IPTG.

The fundamental structure of the model is based on the *Law of Mass Action*, which states that the rate of change in the concentration of a component in a biochemical system is proportional to some product of concentrations of the chemical components present in the system (E. D. Conrad et al., 1999; Novák & Tyson, 2008). Furthermore, several assumptions were made when developing this model; these are listed below. It is worth stating that the following assumptions are commonly made when developing ODE-based models for synthetic gene circuits. Indeed, these have also been made for all the ODE-based models that describe the synthetic gene oscillators reviewed in *Chapter 2* (Atkinson et al., 2003; Michael B Elowitz & Leibler, 2000; Goodwin Waddington, Conrad Hal, 1963; O'Brien, Itallie, & Bennett, 2012; Purcell et al., 2010; Stricker et al., 2008; Tigges et al., 2010, 2009).

- *Assumption 1.* The concentration of every DNA molecule in the system remains constant over time.

- *Assumption 2.* The modulation of transcription by tTA, Gal4 and LacI is described by Hill functions. Correspondingly, transcription factor co-cooperativity is described by Hill coefficients.
- *Assumption 3.* The rate of translation is generalized for all mRNA transcripts.
- *Assumption 4.* The degradation of mRNA transcripts and proteins follow linear kinetics.
- *Assumption 5.* The cell is assumed to be a perfectly mixed biochemical environment and no molecular transportation processes are required for transcription and translation to take place.
- *Assumption 6.* The interactions of LacI with IPTG and tTA with tetracycline follow saturation kinetics and are described by specific transfer functions (Alon, 2007).

The model has 8 variables and 27 parameters, and describes the dynamical interactions of the system illustrated in figure 3.1. The equations that describe it (3.1 to 3.10) are as follows:

$$\frac{d\lambda_m}{dt} = g_1 \cdot \frac{k_1 \cdot \gamma^{n_1}}{k_{d1}^{n_1} + \gamma^{n_1}} - k_{dm1} \cdot \lambda_m \quad (3.1)$$

$$\frac{d\theta_m}{dt} = g_2 \cdot \frac{k_2}{\left(1 + \frac{\lambda \cdot \omega(I)}{k_{d3}}\right)^{n_3}} - k_{dm2} \cdot \theta_m \quad (3.2)$$

$$\frac{d\gamma_m}{dt} = g_3 \cdot \frac{k_3 \cdot \theta^{n_2} \cdot \psi(T)}{k_{d2}^{n_2} + \theta^{n_2} \cdot \psi(T)} - k_{dm3} \cdot \gamma_m \quad (3.3)$$

$$\frac{d\Gamma_m}{dt} = g_4 \cdot \frac{k_3 \cdot \theta^{n_2} \cdot \psi(T)}{k_{d2}^{n_2} + \theta^{n_2} \cdot \psi(T)} - k_{dm4} \cdot \Gamma_m \quad (3.4)$$

$$\frac{d\lambda}{dt} = k_4 \cdot \lambda_m - k_{dp1} \cdot \lambda \quad (3.5)$$

$$\frac{d\theta}{dt} = k_4 \cdot \theta_m - k_{dp2} \cdot \theta \quad (3.6)$$

$$\frac{d\gamma}{dt} = k_4 \cdot \gamma_m - k_{dp3} \cdot \gamma \quad (3.7)$$

$$\frac{d\Gamma}{dt} = k_4 \cdot \Gamma_m - k_{dp4} \cdot \Gamma \quad (3.8)$$

$$\omega(I) = 1 + \left(\frac{I}{K_L}\right)^{n_4} \quad (3.9)$$

$$\psi(T) = 1 - \frac{T}{K_T + T} \quad (3.10)$$

Where $\theta_m, \gamma_m, \lambda_m$ and Γ_m , represent the mRNA concentrations of tTA, Gal4, LacI and GFP transcripts, respectively. The terms θ, γ, λ and Γ represent the concentrations of proteins tTA, Gal4, LacI and GFP, respectively. The terms I and T

represent the concentrations of IPTG and tetracycline, respectively; the kinetics of their respective interactions with LacI and tTA are described by transfer functions ω and ψ . A list describing all the kinetic parameters and their reference values found in the literature is presented in *Appendix A.3*.

3.4. A brief review of bifurcation theory and its applicability to model analysis

In the discipline of dynamic systems, the term *bifurcation* is used to describe the change in qualitative behaviour of a system when a parameter value is changed (Strogatz, 1994). This change in qualitative behaviour can be, for example, a dynamical system transition from a stable to unstable steady state. Various kinds of bifurcations may appear in dynamic systems, leading to different types of dynamic behaviour changes, e.g. pitchfork, saddle-node, Hopf.

Hopf bifurcations are of particular importance for the analysis of the model described by equations 2.1 – 2.10. In fact, a Hopf bifurcation can be defined as the “disappearance or appearance of a periodic orbit in a phase plane, as a result of a local change in the stability properties of a steady point” (Strogatz, 1994). Thus, there is a direct correlation between the existence of Hopf bifurcation regions and transitions in or out of sustained oscillatory behaviour. The appearance of periodic orbits through a Hopf bifurcation is illustrated in the following example of a two-variable ODE system:

$$\begin{aligned}\dot{x} &= f(x, y) \\ \dot{y} &= g(x, y)\end{aligned}\tag{3.11}$$

Where x and y are non-negative quantities representing the concentrations of biochemical components X and Y. The steady state solution (x_0, y_0) can be found by solving:

$$\begin{cases} f(x, y) = 0 \\ g(x, y) = 0 \end{cases} \quad (3.12)$$

A biochemical system will generally possess a stable steady state due to its underlying saturation kinetics assumptions (E. D. Conrad et al., 1999). However, the existence of a periodic steady state solution is guaranteed if stability is lost at a Hopf bifurcation (E. D. Conrad et al., 1999). The stability of (x_0, y_0) can be determined by linearizing 3.11; the Taylor expansion at (x_0, y_0) has the following general form:

$$\begin{aligned} \dot{x} &= f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\ &\quad + O((x - x_0)^2, (y - y_0)^2) \\ \dot{y} &= g(x_0, y_0) + g_x(x_0, y_0)(x - x_0) + g_y(x_0, y_0)(y - y_0) \\ &\quad + O((x - x_0)^2, (y - y_0)^2) \end{aligned} \quad (3.13)$$

The terms $f(x_0, y_0)$ and $g(x_0, y_0)$ vanish at the steady state (x_0, y_0) . Additionally, the nonlinear terms can be dropped by substituting $\hat{x} = x - x_0$ and $\hat{y} = y - y_0$ and setting $\vec{u} = \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix}$, provided that $|\hat{x}| \ll 1$ and $|\hat{y}| \ll 1$. Re-writing 3.13 as a linearized planar system yields:

$$\dot{\vec{u}} = \begin{pmatrix} f_x(x_0, y_0) & f_y(x_0, y_0) \\ g_x(x_0, y_0) & g_y(x_0, y_0) \end{pmatrix} \vec{u} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \vec{u} \quad (3.14)$$

The matrix $J_0 = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ is the *Jacobian matrix* of 3.11 at (x_0, y_0) . Setting $\tau = \text{trace}(J_0) = \alpha + \delta$ and $\Delta = \det(J_0) = \alpha\delta - \beta\gamma$ allows us to find the eigenvalues,

λ , of the Jacobian matrix, which satisfy $\det(J_0 - \lambda I_2) = 0$ and can be used to determine the stability of 3.14:

$$\begin{aligned}\det(J_0 - \lambda I_2) &= \det \begin{pmatrix} \alpha - \lambda & \beta \\ \gamma & \delta - \lambda \end{pmatrix} = (\alpha - \lambda)(\delta - \lambda) - \beta\gamma \\ &= \lambda^2 - (\alpha + \delta)\lambda + \alpha\delta - \beta\gamma = \lambda^2 - \tau\lambda + \Delta \\ \lambda &= \frac{\tau \pm \sqrt{\tau^2 - 4\Delta}}{2}\end{aligned}\tag{3.15}$$

According to linear systems theory, (x_0, y_0) is stable or unstable depending on whether the real part of λ is negative or positive, respectively (E. D. Conrad et al., 1999). When $\Delta < 0$, the discriminant of 3.15 is positive and λ is real. Moreover, $\sqrt{\tau^2 - 4\Delta} > |\tau|$, means that one solution is negative while the other is positive. Thus, the steady state is always a saddle point, i.e. unstable (E. D. Conrad et al., 1999; Strogatz, 1994). Conversely, when $\Delta > 0$, the sign of the discriminant in 3.15 must be considered in order to find $\Re(\lambda)$. If $\tau^2 \geq 4\Delta$, then λ is real, and $\sqrt{\tau^2 - 4\Delta} < |\tau|$ implies that both solutions are of the same sign (E. D. Conrad et al., 1999). In this case, (x_0, y_0) is always an unstable or stable node (E. D. Conrad et al., 1999). However, if $\tau^2 < 4\Delta$, then 3.15 is satisfied by a pair of complex conjugates, with τ as the real part. It is only possible for the real part of λ to change signs in this case, i.e. (x_0, y_0) changes stability if and only if the solutions for the eigenvalues of J_0 are a complex conjugate pair and $\text{trace}(J_0)$ changes sign (E. D. Conrad et al., 1999).

If $\lambda = \mu \pm i\omega$ then the solutions to the linear system 3.14 are of the form $c e^{\lambda t} = c e^{(\mu \pm i\omega)t} = c e^{\mu t} e^{\pm i\omega t} = c e^{\mu t} (\cos(\omega t) \pm i \sin(\omega t))$. Thus, when $\Delta > 0$ and $\tau^2 < 4\Delta$, ω will be nonzero and oscillatory solutions will exist (E. D. Conrad et al., 1999). Indeed, stability analysis guarantees that the stability of the original system (3.11) is determined by the linearized system, as long as the real part of the eigenvalues of the linearized system are nonzero (E. D. Conrad et al., 1999; Strogatz, 1994). In graphical terms, if the solutions for λ cross the imaginary axis in the complex plane anywhere other than the origin, the stability of 3.14, and consequently 3.11, will change and one might expect to see oscillatory behaviour (E. D. Conrad et al., 1999; Strogatz, 1994). This transition of a conjugate pair of

eigenvalues from the imaginary to the real axis is precisely what defines a Hopf bifurcation (E. D. Conrad et al., 1999; Strogatz, 1994).

There are three fundamental reasons why the theoretical concepts described above are directly applicable to the design of the proposed synthetic gene oscillator: (i) finding the parametric regions corresponding to Hopf bifurcations allows one to map specific ranges of parameters to sustained oscillatory behaviour; (ii) the analysis of periodic orbits in the phase plane of equations 3.1 – 3.8 provides critical insight regarding the frequency and amplitude modulation properties of the system, thus allowing an assessment of system tunability to be made; (iii) knowing the size of Hopf regions provides a solid framework for verifying if parameter values corresponding to oscillatory behaviour are near a bifurcation point (i.e. which would abolish or create periodic orbits when crossed) – this is highly relevant for the design of dynamic mammalian systems given their current parametric uncertainty.

There are several software packages that can be used to perform bifurcation analysis of dynamic systems (a. Dhooge, Govaerts, & Kuznetsov, 2003). However, most of these are written in a relatively low-level programming language, which makes the task of specifying, inputting and updating systems highly laborious (a. Dhooge et al., 2003). On the other hand, *MatCont*, developed by Dhooge *et al.* (2003), is a powerful and widely used environment for the study of differential equations and general scientific computing in MATLAB (a. Dhooge et al., 2003). *Matcont*, like most software packages used to perform stability analysis of dynamic systems, relies on numerical continuation (a. Dhooge et al., 2003). Numerical continuation is a computing method to approximate the solutions of systems of parametric nonlinear equations (a. Dhooge et al., 2003). This methodology is based on finding the solutions of a nonlinear system at a fixed point and continuing these throughout a parametric variation. There is great implementation variety within different numerical continuation methods (Allgower & Georg, 2003); however, the bifurcation analysis carried out in this work was performed with *Matcont* (version 6p2), given its proven reliability and widespread use in studying nonlinear biological systems (Bindel, Demmel, Friedman, Govaerts, & Kuznetsov, 2005; a. Dhooge et al., 2003; A. Dhooge, Govaerts, & Kuznetsov, 2004).

3.5. Results

3.5.1. Numerical solution of the model describing the proposed synthetic gene oscillator

Solving the dynamic model described by equations 3.1 – 3.10 provides critical insight into the dynamic behaviour of the proposed synthetic gene oscillator. Validating system functionality, that is, knowing that the proposed system is capable of sustaining oscillatory behaviour under realistic parameter values while being responsive to concentration gradients of tetracycline and IPTG, is the first criterion the model must meet in order to be considered as a plausible description of the proposed synthetic gene oscillator. To this end, the model was specified in MATLAB (version R2014_a) and solved using the *ode45* differential equation solver. The corresponding MATLAB code is presented in *Appendix A.3*.

The steady state dynamic concentration profiles of system protein and mRNA components are shown in figures 3.2 and 3.3, respectively. Figure 3.2 shows the evolution of LacI, tTA, Gal4 and GFP concentrations with time, in the absence of chemical inhibitors IPTG and tetracycline, and assuming equimolar gene dosage ratios between all components. Figure 3.3 illustrates the same evolution of system components with time, but for their respective mRNA transcripts.

It is clear from both figures that the system exhibits sustained oscillations under a realistic choice of parameters, with a period of approximately 1000 minutes. It can also be noted that the proportion of mRNA transcript concentrations to their respective proteins is maintained for all components, thus reflecting *Assumption 3*, i.e. generalized translation rate. Moreover, the oscillatory ranges for all components are in line with those of previous designs (Tigges et al., 2009). Figures 3.2 and 3.3 also illustrate that the tTA mRNA transcript and protein concentrations oscillate at considerably lower amplitudes than the remaining components, a feature that has also been reported in components of previous designs (Tigges et al., 2010, 2009).

Design of a novel mammalian synthetic gene oscillator

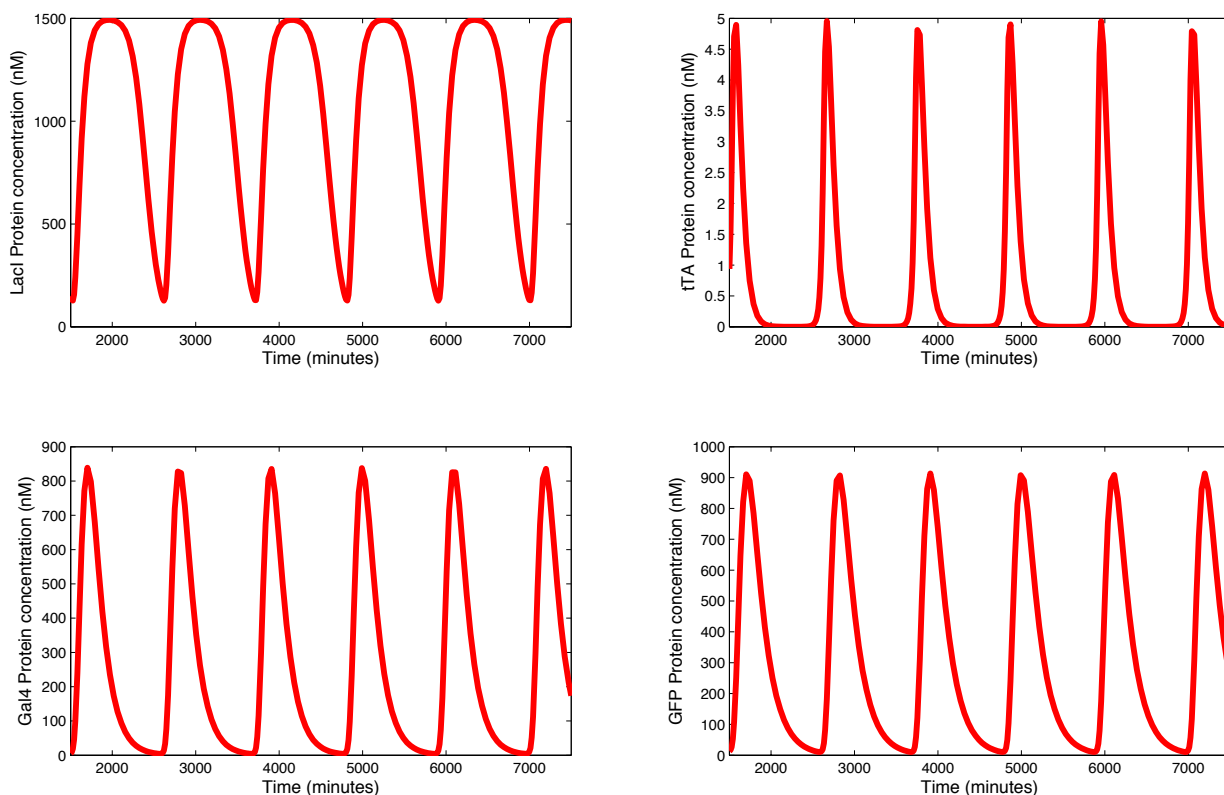


Figure 3.2. Concentration profiles of LacI, tTA, Gal4 and GFP proteins in the absence of chemical inhibitors IPTG and tetracycline, assuming equimolar gene dosage ratios (1:1:1:1). Results are shown for the 1500 – 7500 minute time window.

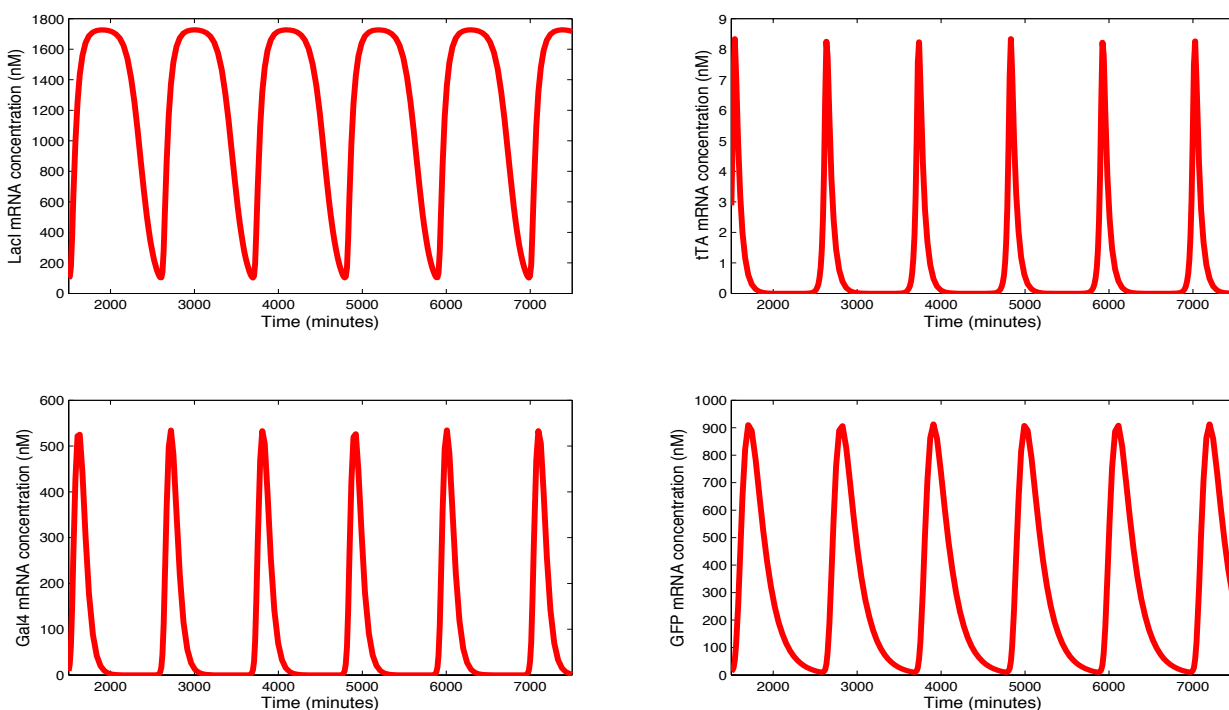


Figure 3.3. Concentration profiles of LacI, tTA, Gal4 and GFP mRNA transcripts in the absence of chemical inhibitors IPTG and tetracycline, assuming equimolar gene dosage ratios (1:1:1:1). Results are shown for the 1500 – 7500 minute time window.

This feature of tTA oscillations, as well as their resemblance to ‘fire and degrade’ peaks (i.e. ‘spiky’ appearance), can be attributed to the fact that the *tTA* gene is the only system component that is under direct repression, i.e. from LacI.

In addition to the simulation results presented in figures 3.2 and 3.3, a preliminary frequency and amplitude modulation test was carried out by simulating system dynamics under different concentrations of tetracycline and IPTG. These are illustrated in figure 3.4. Tetracycline and IPTG have a clear effect in system dynamics, at the frequency and amplitude level: increasing the concentration of tetracycline clearly reduces the amplitude of oscillations while simultaneously reducing the period; increasing the concentration of IPTG reduces the period of oscillations while simultaneously having a small amplitude reduction effect. These effects are further studied in the next section with a bifurcation analysis-based study of the periodic orbits that arise in the system described by equations 3.1 – 3.8.

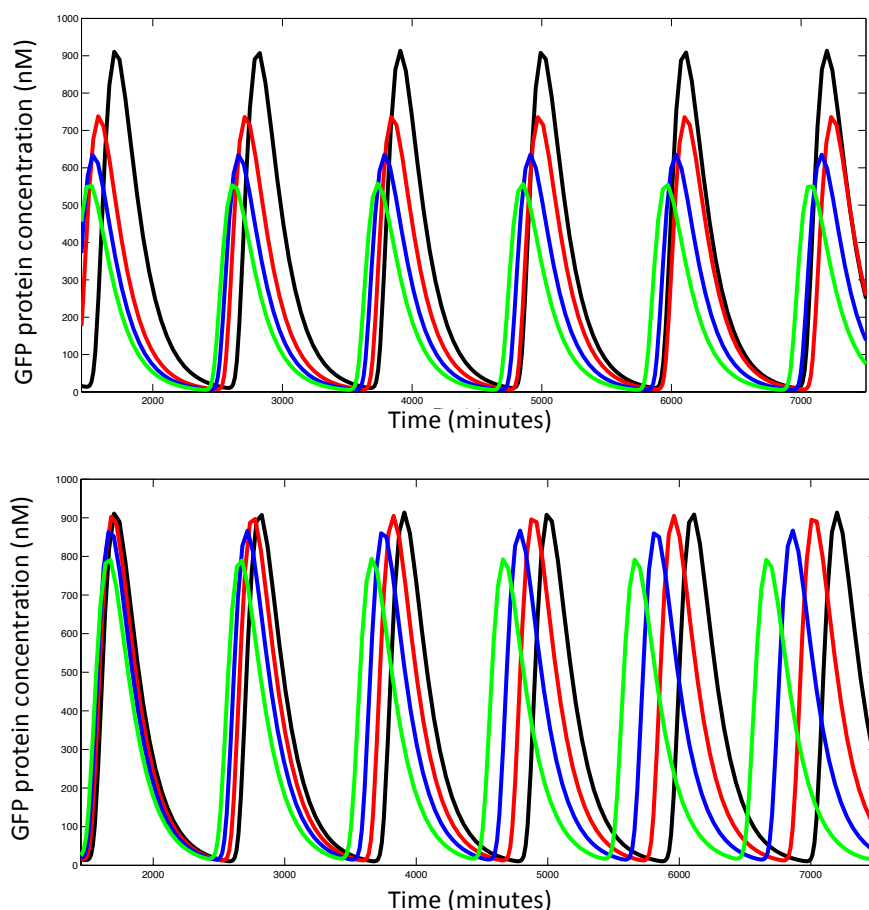


Figure 3.4. Preliminary analysis of frequency and amplitude modulation by tetracycline (top) and IPTG (bottom). (Top) System dynamics were simulated assuming the tetracycline concentrations 0 nM, 20 nM, 40 nM, and 60 nM (black, red, blue and green, respectively). (Bottom) System dynamics were simulated assuming IPTG concentrations 0 nM, 1.5 μM, 2 μM, 2.5 μM, (black, red, blue and green, respectively).

3.5.2. Bifurcation analysis of the proposed synthetic gene oscillator model

Analysing the parametric space of the model described by equations 3.1 – 3.10 is of critical importance to the design of the proposed synthetic gene oscillator. Numerical bifurcation analysis was performed to address this goal by scanning the model core parameters and correlating them to parametric space regions corresponding to oscillatory behaviour. Core parameters were considered to be those influencing the kinetics of the core components of the system: mRNA and protein degradation rates, generalized translation rate, and mRNA transcription rates.

Periodic solutions to the model were found to arise exclusively through Hopf bifurcations. Table 3.1 comprises a list of numerical ranges at which the core model parameters induce oscillatory behaviour. The bifurcation diagrams for the parameters listed in table 3.1 are presented in *Appendix A.3*, together with the Matcont continuation specifications used when each individual parameter. As expected, the reference values of all parameters listed in table 3.1 are within a Hopf bifurcation region. Indeed, with the exception of k_1 , k_{dp1} and k_{dp3} (table 3.1), all reference parameter values are located within their respective Hopf bifurcation regions with a distance of at least one order of magnitude to a Hopf point. These confidence interval values for k_1 , k_{dp1} and k_{dp3} are 62%, a factor of *c.* 5.7, and 67.2% of their corresponding reference values, respectively.

Moreover, only one Hopf bifurcation point was detected when scanning parameters k_1 , k_{dm1} , k_{dm3} and k_4 (table 3.1). This ensures that k_1 , k_{dm1} and k_4 can individually take any value higher than $11.387196 \text{ min}^{-1}$, 0.00108 min^{-1} and $0.003006 \text{ min}^{-1}$, respectively, and k_{dm3} any value lower than $6.875593 \text{ min}^{-1}$, without abolishing oscillations in the model. The remaining parameters listed in table 3.1 must take a value between their corresponding Hopf points in order to not abolish oscillations. Bifurcation analysis was also carried out for model parameters that are considered variable during experimental setup, namely the gene dosage ratios g_1 , g_2 , g_3 and the concentrations of tetracycline and IPTG.

Table 3.1. Core parameter values and respective Hopf intervals corresponding to oscillatory behaviour in the model described by equations 3.1 – 3.8.

Parameter	Symbol	First Hopf point	Last Hopf point	Reference value
Lacl mRNA				
transcription rate (min ⁻¹)	k_1	11.387196	N/A	30
tTA mRNA				
transcription rate (min ⁻¹)	k_2	0.000714	1445.609784	30
Gal4 mRNA				
transcription rate (min ⁻¹)	k_3	0.059681	1463.426501	30
Lacl mRNA				
degradation rate (min ⁻¹)	k_{dm1}	0.00108	N/A	0.0173
tTA mRNA				
degradation rate (min ⁻¹)	k_{dm2}	0.002466	585.70643	0.0173
Gal4 mRNA				
degradation rate (min ⁻¹)	k_{dm3}	N/A	6.875593	0.0173
General translation				
rate (min ⁻¹)	k_4	0.003006	N/A	0.02
Lacl degradation				
rate (min ⁻¹)	k_{dp1}	0.001004	0.056846	0.01003
tTA degradation				
rate (min ⁻¹)	k_{dp2}	0.002491	806.804175	0.0231
Gal4 degradation				
rate (min ⁻¹)	k_{dp3}	0.002262	3.328788	0.0069

The main objective of this was to determine if these parameters could take values that (i) lead to the appearance or disappearance of oscillatory behaviour, and (ii) are able to change the frequency of oscillations, thus contributing for system tunability.

Tetracycline and IPTG concentrations, as well as gene dosage parameters g_1 , g_2 , g_3 , were found to induce oscillatory behaviour in the model, with periodic orbits also arising from Hopf bifurcations. Their respective Hopf bifurcation intervals are listed in table 3.2. Furthermore, apart from being able to induce and abolish oscillatory behaviour, these parameters were found to dynamically change the frequency of oscillations when scanned throughout their respective Hopf intervals, further corroborating the results presented in figure 3.4. The tunability potential of these parameters is illustrated in figure 3.5. It is clear from table 3.2 that the model is able to produce oscillatory behaviour in the absence of tetracycline and IPTG. Indeed, these parameters are only able to abolish oscillatory behaviour if they take any value higher than 203.943297 nM and 3144.910751 nM, respectively. Gene dosage parameters g_1 , g_2 , and g_3 are also able to induce and abolish oscillatory behaviour in the model. Indeed, the model presents periodic solutions for any value of g_1 higher than 0.000024, and for any values of g_2 and g_3 that are within their specified Hopf intervals.

Table 3.2. Hopf interval ranges of tetracycline, IPTG and gene dosage parameters.

Parameter	Symbol	First Hopf point	Last Hopf point
Tetracycline concentration (nM)	T	N/A	203.943297
IPTG concentration (nM)	I	N/A	3144.910751
Lacl gene dosage (adimensional)	g_1	0.379573	N/A
tTA gene dosage (adimensional)	g_2	0.000024	48.186993
Gal4 gene dosage (adimensional)	g_3	0.001989	48.780886

Desian of a novel mammalian svnthetic aene oscillator

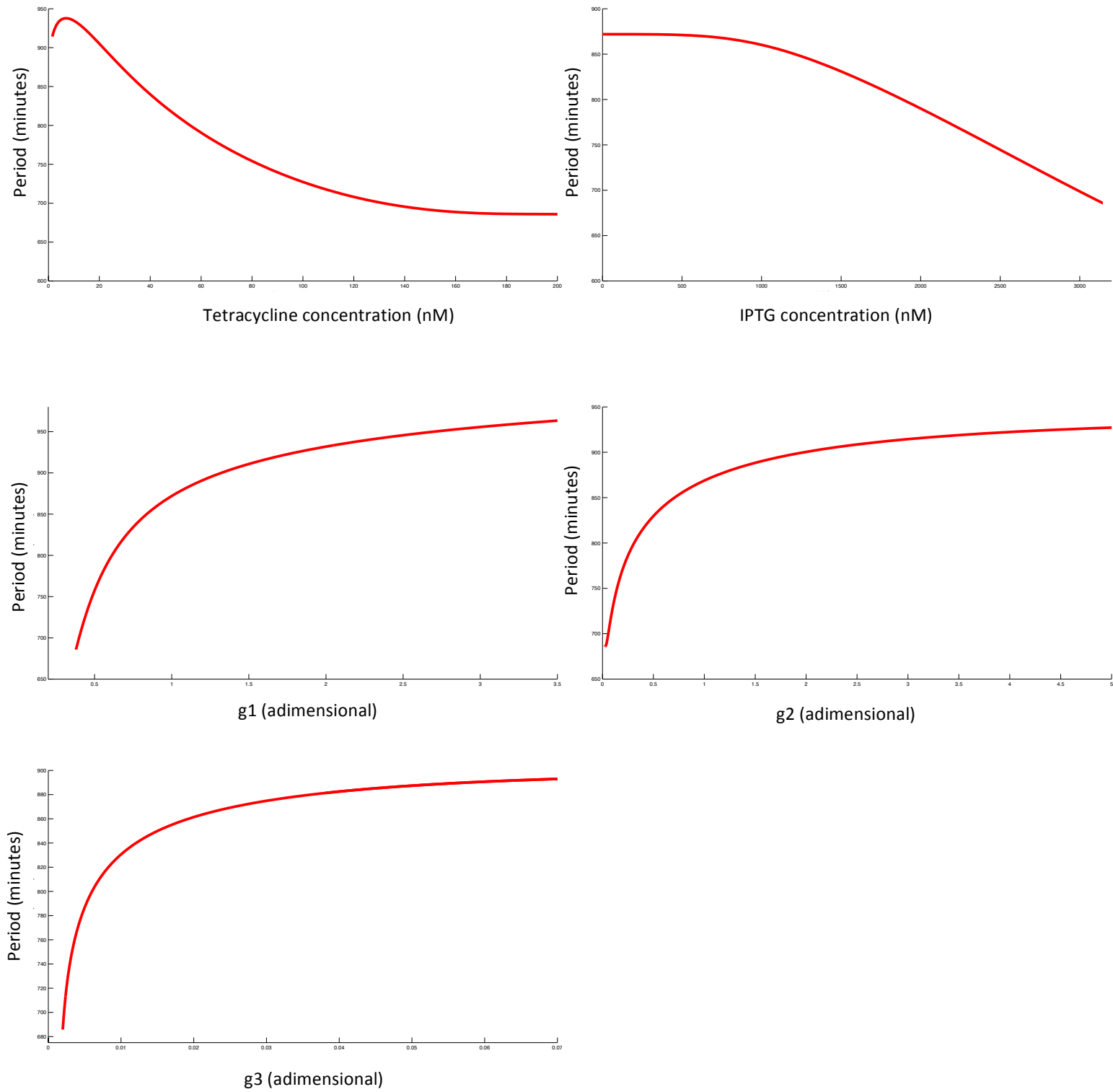


Figure 3.5. Tunability potential of parameters variable through experimental setup. (Top left) Period variation with increasing tetracycline concentration. (Top right) Period variation with increasing IPTG concentration. (Middle left, Middle right, bottom left, respectively) Period variation with increasing gene dosages g_1 , g_2 , and g_3 .

Figure 3.5 illustrates that all the parameters that are changeable through the experimental setup of system implementation have the potential of dynamically modifying the frequency of oscillations. This effect is more pronounced in tetracycline-, g_1 - and g_2 - mediated frequency tunability with respective period tuning ranges of c. 260, 265 and 230 minutes. Although less pronounced, this effect is also clearly present in IPTG- and g_3 -mediated tunability, with respective period tuning ranges of c. 195 and 175 minutes.

3.5.3. Considerations about promoter leakiness

Leakiness is defined as the basal expression level of a gene that is under repression (Ausländer, Ausländer, Müller, Wieland, & Fussenegger, 2012; Michael B Elowitz & Leibler, 2000; Greber & Fussenegger, 2007; Weber & Fussenegger, 2009). Additionally, it is a commonly used term to refer to the residual expression of genes that have upstream transactivator-responsive regulatory elements. In essence, leakiness can be thought of as a measure of (i) how ineffective transcriptional repressors are, and (ii) the expression level from transactivator-responsive regulatory elements in the absence of their corresponding transactivator (Ausländer et al., 2012; Greber & Fussenegger, 2007; Weber & Fussenegger, 2009). Leakiness from promoters and other gene regulatory elements has been proposed as a problematic feature of synthetic gene oscillators, and synthetic gene networks in general. This is mostly because leakiness provides an unwanted residual level of gene expression, which hinders the goal of achieving full controllability of system dynamics (Purcell et al., 2010; Serrano, 2007; Weber & Fussenegger, 2009).

Leakiness is relatively well characterized in prokaryotic synthetic gene networks, namely in past synthetic gene oscillator implementations (Michael B Elowitz & Leibler, 2000; Purcell et al., 2010; Stricker et al., 2008). However, this is not the case for synthetic gene networks implemented in eukaryotic hosts, especially mammalian cells, where leakiness is known to greatly vary in a cell line-specific manner (Ausländer et al., 2012; Greber & Fussenegger, 2007; Weber & Fussenegger,

2009). This parametric uncertainty involving the contribution of leakiness to the overall gene expression from system components is, in fact, the reason why leakiness parameters were not considered when developing the model described by equations 3.1 to 3.10.

However, these parameters were considered from a bifurcation analysis standpoint with the objective of determining the maximum leakiness that individual system components can withstand while preserving oscillatory behaviour. The results of a bifurcation analysis carried out on equations 3.1 to 3.10 considering parameters a_1 , a_2 and a_3 as the respective leakiness from *lacI*, *tTA* and *gal4* are presented in table 3.3.

It is apparent from tables 3.1, 3.2 and 3.3 that the model described by equations 3.1 to 3.10 is able to produce oscillatory behaviour in the absence of leakiness. Indeed, table 3.3 illustrates that only one Hopf point was discovered for each of the leakiness parameters considered (i.e. for positive values of a_1 , a_2 and a_3). Moreover, the maximum permitted leakiness from upstream regulatory elements greatly varies between system components. The system is able to produce sustained oscillatory behaviour even when the basal expression of *lacI* is 6 times greater than its standard expression level under Gal4-mediated transactivation. The maximum permitted basal expression levels from the regulatory elements upstream of *tTA* and *gal4* is considerably lower: c. 0.1% and 4.8%, respectively.

Table 3.3. Hopf point locations of leakiness parameters for core system components.

Module	Maximum leakiness from upstream regulatory element (adimensional)
<i>lacI</i>	6.006596
<i>tTA</i>	0.001013
<i>gal4</i>	0.048426

3.6. Discussion

The results presented in section 3.5 illustrate that the proposed synthetic gene oscillator design is able to produce sustained oscillatory behaviour with a realistic choice of parameter values. This is reflected in the numerical solution of the model describing the system, represented by equations 3.1 to 3.10 (figures 3.2 and 3.3). Different system components were found to oscillate with different amplitudes, but preserving the same oscillatory period, i.e. *c.* 1000 minutes. This effect is more pronounced in tTA, which was found to oscillate at considerably lower amplitude than LacI, Gal4 and GFP. This feature has been observed in previous synthetic gene oscillator implementations: system components under direct repressive activity oscillate at lower amplitudes (O'Brien, Itallie, et al., 2012; Purcell et al., 2010; Stricker et al., 2008; Tigges et al., 2010, 2009). Moreover, the development of the dynamic model described by equations 3.1 to 3.10 provided a theoretical framework in which system dynamics under the influence of chemical inhibitors (i.e. IPTG and tetracycline) could be tested. Indeed, a simple simulation exercise allows one to verify that IPTG and tetracycline are able to influence the amplitude and frequency of oscillations (figure 3.4). This feature was also present in previous oscillator designs: the original Fussenegger oscillator was tunable through varying tetracycline concentrations; the Smolen / Hasty oscillator was tunable across a wide range of arabinose and lactose concentrations (Purcell et al., 2010; Stricker et al., 2008; Tigges et al., 2009).

The appearance of periodic orbits in the phase space of equations 3.1 to 3.10 was further analyzed using bifurcation analysis. Moreover, this technique was used to investigate the tunability potential of the proposed system. Periodic solutions were found to arise through Hopf bifurcations, as outlined in section 3.4, thus corroborating that the model produces sustained oscillatory behaviour resulting from the transition of pairs of complex conjugate eigenvalues from the imaginary to the real plane. An analysis of periodic orbits with varying concentrations of IPTG and tetracycline revealed that these parameters, controllable in an experimental context, not only could abolish oscillatory behaviour if brought to high enough values, but

could also vary the oscillatory period by as much as c. 260 and 190 minutes. This tunability range is also in line with previous synthetic gene oscillator implementations: the original Fussenegger oscillator model predicted oscillations with a period of 170 minutes and a tunability range of c. 71 minutes; the low-frequency Fussenegger oscillator model predicted a period of 1560 minutes and a tunability range of 510 minutes; the Smolen / Hasty oscillator model predicted a period of c. 40 minutes with a tunability range of c. 20 minutes (Purcell et al., 2010; Stricker et al., 2008; Tigges et al., 2010, 2009). The dynamic frequency tuning effect of IPTG and tetracycline was also found to have an effect on amplitude of oscillations (figure 3.4). This has also been reported in previous designs, and can be explained by the fact that de-coupled frequency and amplitude modulation is highly dependent on the entrainment of oscillator populations rather than on individual parameters (Mondragón-Palomino, Danino, Selimkhanov, Tsimring, & Hasty, 2011; Tsimring et al., 2010).

In addition to the dynamic effect of tetracycline and IPTG, the gene dosages of core system components were tested as potential tunability factors. Like tetracycline and IPTG, the relative gene dosages g_1 , g_2 and g_3 also had a frequency alteration effect on system dynamics. Specifically, g_1 , g_2 and g_3 were found to have a tuning potential of c. 265, 230 and 175 minutes. The ability of tuning the frequency of oscillations by varying relative gene dosages was also explored in the original Fussenegger design, and validated in an experimental context (Tigges et al., 2009).

The tunability potential of the proposed system can thus be explored using five different experimentally controllable parameters. This feature implies a modular controllability in tuning potential that is not shared with previous designs: bacterial implementations of synthetic gene oscillators had, at most, two tuning parameters; the original Fussenegger design was tunable across a range of tetracycline concentrations, and relative dosages of two components (Purcell et al., 2010; Stricker et al., 2008; Tigges et al., 2009).

It is relevant to state that apart from being able to dynamically tune the frequency of oscillations, the ranges of tetracycline concentration, IPTG concentration and relative gene dosages across which tunability is possible, remain

valid from a biological perspective. Indeed, tetracycline- and IPTG-mediated tunability occurs in the ranges of 10 nM to 150 nM and 750 nM to 3000 nM, respectively. These concentration ranges are non-toxic for mammalian cells, and usually supplied in *in vivo* experiments of tetracycline- and IPTG-responsive synthetic gene networks (Aubrecht et al., 1996; Cronin et al., 2001; Gossen & Bujard, 1992; Gossen et al., 1995). Likewise, the relative gene dosage value ranges where g_1 , g_2 and g_3 -mediated tunability occurs are, respectively, 0.4 to 3, 0.0001 to 0.003, and 0.009 to 0.05. This also reflects that tuning system frequency using these parameters is realistic from an experimental perspective, as shown in previous implementations (Tigges et al., 2009): modular transfection-based *in vivo* system implementation allows for this adaptability of system component dosage ratios.

In addition to determining the tunability potential of the proposed system, the modelling framework developed herein allowed for an exploration of the parametric space of equations 3.1 to 3.8 (table 3.1). This exploration, carried out using bifurcation analysis, revealed that with the exception of k_1 , k_{dp1} and k_{dp3} all reference values of core model parameters are located at least one order of magnitude away from their respective closest Hopf point. These confidence interval values for k_1 , k_{dp1} and k_{dp3} are 62%, a factor of *c.* 5.7, and 67.2% of their corresponding reference values, respectively. This parametric exploration provides critical insight into the *topological robustness* of the proposed system. That is, the existence of large parameter confidence intervals with respect to the size of Hopf regions (table 3.1) provides a shielding strategy against parametric uncertainty: the model described by equations 3.1 to 3.10 will produce oscillatory behaviour even if the parameter values considerably vary from their reference values. This feature is relevant for the design of a mammalian synthetic gene oscillator, where parameter values greatly vary according to implementation strategies, experimental setups, and intracellular noise (Ausländer et al., 2012; Greber & Fussenegger, 2007; O'Brien, Van Itallie, et al., 2012; Pedraza & van Oudenaarden, 2005).

Promoter leakiness was also explored using a bifurcation analysis approach to equations 3.1 to 3.8. Parameters governing the kinetics of promoter leakiness were not considered when simulating model behaviour given that they show great

variation in mammalian cells across different cell lines, and its dynamics remain poorly characterized in these hosts in comparison to prokaryotic hosts (Ausländer et al., 2012; M B Elowitz & Leibler, 2000; Greber & Fussenegger, 2007; Purcell et al., 2010; Weber & Fussenegger, 2009). Instead, leakiness parameters were considered from a parametric space exploration standpoint, in order to provide insight into how much basal activity each of the system modules could withstand while still sustaining oscillatory behaviour (table 3.3). It is apparent that *lacI* can withstand a great level of basal activity from its upstream regulatory element in comparison to *tTA* and *gal4*. This is because expression from *lacI*, resulting in the synthesis of the LacI repressor, carries out the negative feedback in the system and the subsequent *tTA:gal4*-induced delay is large enough to permit a temporary accumulation of LacI, while sustaining oscillatory behaviour. This topological feature has also been identified in dynamic models of *in vitro* oscillators, circadian rhythms, and inflammatory pathways (Bodnar & Bartomiejczyk, 2012; Momiji & Monk, 2008; Monk, 2003; Rajala et al., 2010; Rateitschak & Wolkenhauer, 2007).

3.7. Concluding remarks

The main objective of this chapter was to provide a theoretical description and address the *in silico* design component of a novel mammalian synthetic gene oscillator. After comprehensively reviewing the advantages and disadvantages of potential system structures and topologies while taking into account their implementation feasibility, a 3-component delayed negative feedback oscillator based on *lacI*, *tTA* and *gal4* was the chosen topology for the proposed system.

A lumped parameter kinetic ODE model that describes the dynamics of the system was developed, taking underlying synthetic gene oscillator design principles into account. The numerical solution of the model showed that the system produces oscillatory behaviour under a realistic choice of parameter values. Moreover, simulation results showed that varying the concentrations of the two chemical inhibitors of the system, IPTG and tetracycline, could alter amplitude and frequency

of oscillations. This effect was further analyzed through bifurcation analysis, which revealed that IPTG, tetracycline, and relative gene dosages could all be used as tunability factors across a wide range of values.

Bifurcation analysis was also used to explore the parametric space of the ODE model. The results from this analysis show that all core parameter reference values are located in Hopf regions and have sizeable confidence intervals with respect to Hopf bifurcation points. The potential effect of promoter leakiness on system dynamics was also assessed in this way, revealing comprehensive maximum leakiness values that the individual system modules can withstand, while sustaining oscillatory behaviour.

The next chapter focuses on the biological construction and *in vivo* validation of the individual system parts: *lacI*, *tTA*, *gal4* and *gfp*. The first sections address the methodological steps and specifications taken to assemble individual modules. Results of their *in vivo* implementation are then presented, followed by a comprehensive discussion.

Chapter 4

Construction and *in vivo* validation of system parts

4.1. Introduction

This chapter addresses the biological construction and validation of the elements that make up the proposed mammalian synthetic gene oscillator design. A 3-component delayed negative feedback topology based on the periodic interactions of *lacI*, *tTA* and *gal4* has been shown to provide considerable tunability potential and topological robustness to the system.

The interactions between system components are based on the structural organization of upstream regulatory elements, namely promoters and operator sites, i.e. transcription-factor binding sites. Indeed, promoters and operator sites essentially drive the logical structure and topology of the system by determining the target binding sites for transcription factors LacI, tTA and Gal4. Thus, the biological construction of each component was carried out in modules, whereby each module is defined by a promoter, an operator site, the gene Open Reading Frame (ORF), and downstream elements (i.e. poly-adenylation and nuclear localization signals).

Furthermore, the biological assembly of each module must take the feasibility of *in vivo* mammalian cell implementation into account. Specifically, each system module is required to have a genetic element organization that allows for mammalian cell expression. To this end, several biological design core principles based on previous implementations of mammalian synthetic gene networks, were taken into account (Kis et al., 2015; Purcell et al., 2010):

- Plasmids containing individual modules must be endotoxin-free mammalian expression vectors;
- Each module must have downstream mRNA maturation signals, namely mammalian cell-recognizable poly-adenylation signals;
- The ORFs of each gene must be codon-optimized for mammalian cell expression in order to avoid additional delays within the system;
- Promoters and operator sites must be functional in mammalian cells;

Thus, the biological construction of each system module was based on a design framework that ensures implementation feasibility. Additionally, each

Construction and in vivo validation of system parts

assembled module was individually validated *in vivo*, in Human Embryonic Kidney (HEK293T) cells. This validation is fundamental to corroborate that all parts are functional before implementing the assembled system *in vivo*.

The next section provides a comprehensive review of the individual elements used to construct all system modules. This is followed by a detailed description of the underlying experimental methodology. *In vitro* and *in vivo* validation results are then presented and discussed, followed by a conclusion.

4.2. Organization of biological elements and system modules

The core regulatory part of the proposed synthetic gene oscillator comprises three genes (*lacI*, *tTA* and *gal4*) that interact according to a delayed negative feedback loop topology (figure 3.1). The logical configuration of this topology is ensured by the presence of upstream regulatory elements that act as transcription factor binding sites, subsequently leading to up- or down-regulation of the corresponding downstream gene. Figure 4.1 illustrates the upstream regulatory element structure of each system module.

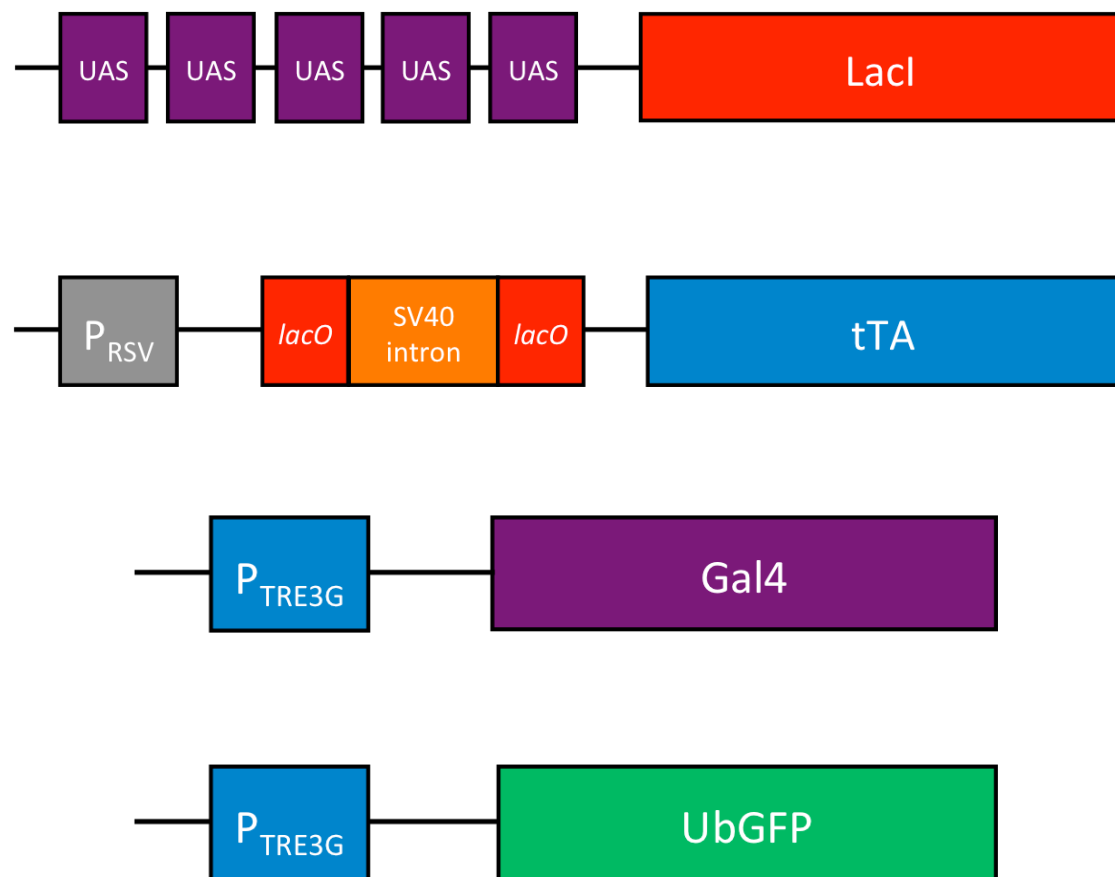


Figure 4.1. Diagrammatic representation of the *lacI*, *tTA*, *gal4* and *gfp* modules. The *lacI* module contains upstream regulatory elements that are responsive to Gal4-mediated transactivation (UAS). The *tTA* module has an upstream RSV promoter followed by a SV40 chimeric intron that is flanked by *lacO* sites. The *gal4* and *gfp* modules are regulated by tetracycline-responsive element promoters, which allow for tTA-mediated up-regulation.

The *lacI* module is controlled by five Upstream Activation Sequences (UAS), which allow for Gal4-mediated up-regulation of *lacI*. These sequences have previously been shown to provide this biological function in eukaryotic hosts, including mammalian cells (Carey, Lin, Green, & Ptashne, 1990; Distel, Wullimann, & Köster, 2009; Sadowski, Ma, Triezenberg, & Ptashne, 1988; Scheer & Campos-Ortega, 1999; Webster et al., 1988). Furthermore, the optimal number of UAS sites Gal4-mediated expression has been shown to be five, thus validating the regulatory organization design in this module (Potter, Tasic, Russler, Liang, & Luo, 2010; Wei, Potter, Luo, & Shen, 2012). The *lacI* gene sequence considered in this design has also been optimized for mammalian cell expression, i.e. *lacI^q* (Agilent Technologies, 2015).

The upstream regulatory structure of the *tTA* module is responsive to LacI-mediated down-regulation. This is due to the presence of two *lacO* operator sites immediately upstream of the *tTA* gene ORF. These *lacO* sites flank a chimeric SV40 intron, which has been shown to increase the strength of LacI repression in mammalian cells (Luo, Celler, & Berndt, 1999). An element capable of driving constitutive expression of *tTA* is required if the proposed system topology is to be followed. In this case, a RSV promoter was chosen to drive constitutive expression of *tTA*, (i) given its reliability when coupled to *lacO* operator sites (Luo et al., 1999), and (ii) its similarity to CMV promoters in the context of transcription initiation strength (Aubrecht et al., 1996; Boulos, Meloni, Arthur, Bojarski, & Knuckey, 2006; Qin et al., 2010; Zarrin et al., 1999)

For the delayed negative feedback loop topology to be completed, the *gal4* module requires a tTA-responsive upstream regulatory element: P_{TRE3G}. This promoter contains seven *tetO* operator sites and ensures tight regulation of the downstream ORF by minimizing the basal level of expression (Becskei & Serrano, 2000; Gossen & Bujard, 1992; Gossen et al., 1995; Kis et al., 2015). Thus, it is also responsible for regulating expression of the ubiquitin-tagged *gfp* readout. The *gal4* ORF sequence considered in this design has also been optimized for mammalian cell expression (Potter et al., 2010).

Three synthetic gene circuits were constructed in order to validate the activity of each individual module *in vivo* (figure 4.2). The *lacI* validation circuit is based on LacI-mediated down-regulation of a *gfp* gene containing upstream *lacO* operator sites that flank a chimeric SV40 intron. Accumulation of LacI occurs due to constitutive expression from an upstream CMV promoter, which provides reliable and strong expression of the downstream *lacI* gene (Aubrecht et al., 1996; Qin et al., 2010). The interaction of LacI with its target binding sites can be inhibited by IPTG. In the *tTA* validation circuit, constitutive expression from a CMV promoter leads to accumulation of tTA, which in turn up-regulates *gfp* expression by binding to seven upstream *tetO* sites in a P_{TRE3G} promoter. The *gal4* validation circuit is also based on constitutive expression from a CMV promoter leading to accumulation of Gal4, which in turn up-regulates *gfp* expression by binding to five UAS sites.

Implementing these three synthetic gene circuits *in vivo* provides a comprehensive validation platform for the components of the proposed synthetic gene oscillator. In the *lacI* validation circuit, accumulation of LacI is expected to inhibit *gfp* expression in the absence of IPTG. Thus, if the *lacI* gene ORF and the upstream regulatory landscape of the *gfp* gene are functional, no fluorescence should be observed. Conversely, fluorescence should be observed in the presence of IPTG due to inhibition of LacI-mediated down-regulation of *gfp*. In the *tTA* validation circuit, accumulation of tTA is expected to up-regulate expression of the *gfp* gene, in the absence of tetracycline. Thus, if the *tTA* gene ORF and the P_{TRE3G} upstream of *gfp* are functional, fluorescence should be observed. Fluorescence should be inhibited in the presence of tetracycline, in a dose-dependent manner, due to inhibition of tTA-mediated up-regulation of *gfp*. In the *gal4* validation circuit, a CMV promoter drives constitutive expression of *gal4*, leading to subsequent accumulation of Gal4 and up-regulation of the *gfp* gene. Thus, if the UAS components upstream of *gfp* and the *gal4* gene ORF are functional, fluorescence should be observed.

The next section provides a detailed description of the methodologies used to construct and implement these three synthetic gene circuits in HEK293T cells.

Construction and *in vivo* validation of system parts

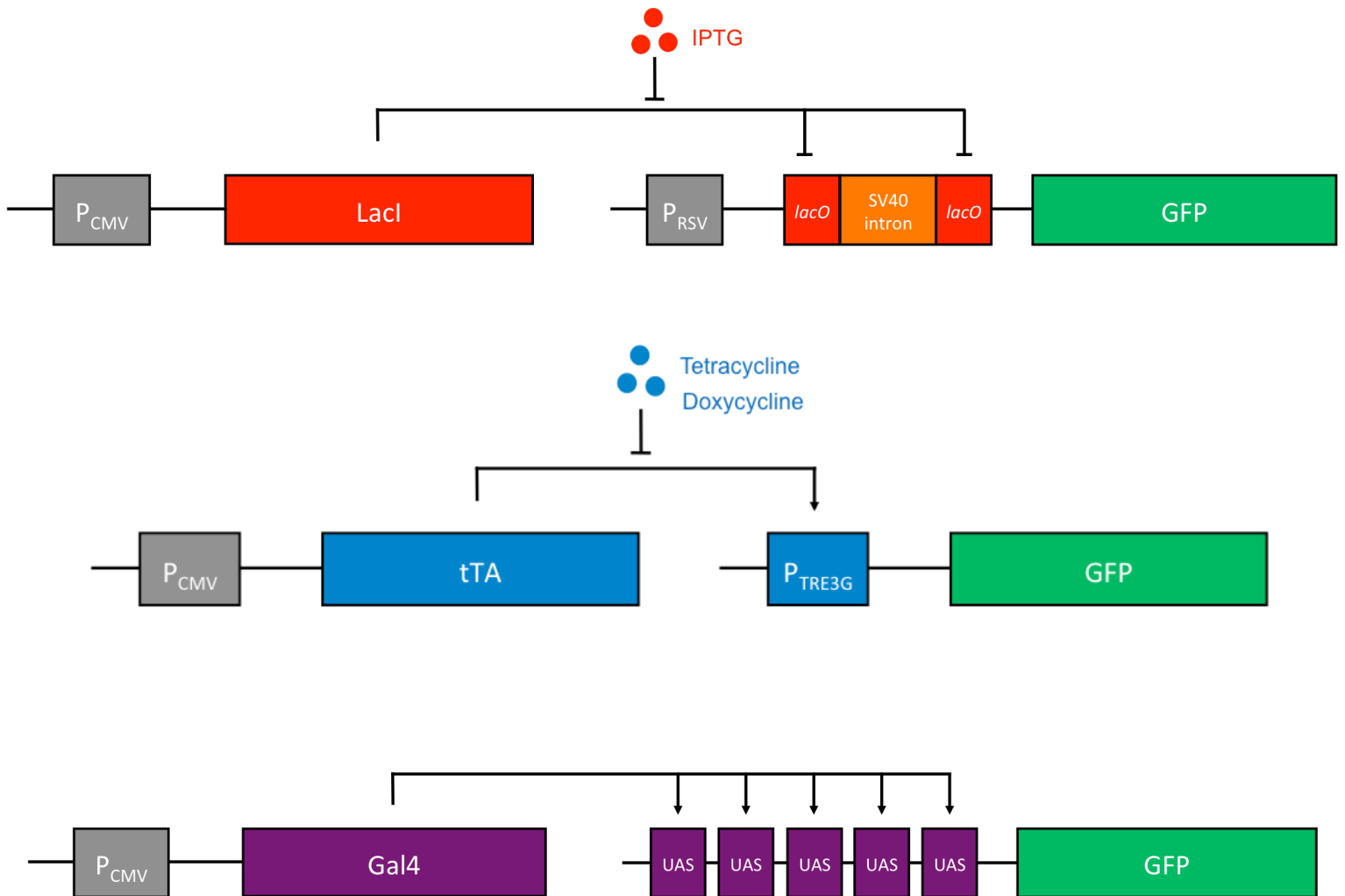


Figure 4.2. Diagrammatic representation of *lacI*-, *tTA*-, and *gal4*-based synthetic gene circuits for individual module *in vivo* validation. All circuits are dependent on constitutive expression from a CMV promoter. In the *lacI*-based circuit, constitutive expression leads to accumulation of *Lacl*, which in turn down-regulates expression of a *gfp* gene. This interaction can be inhibited by IPTG. In the *tTA*-based circuit, constitutive expression leads to accumulation of *tTA*, which up-regulates *gfp* expression from a P_{TRE3G} . This interaction can be inhibited by tetracycline or doxycycline. In the *gal4*-based circuit, constitutive expression leads to *Gal4* accumulation and subsequent *UAS*-mediated up-regulation of *gfp*.

4.3. Methods

4.3.1. Bacterial cell culture conditions

Stellar™ Competent Cells (*E. coli* with genotype F⁻, endA1, supE44, thi-1, recA1, relA1, gyrA96, phoA, Φ80d lacZΔ M15, Δ (lacZYA - argF) U169, Δ (mrr - hsdRMS - mcrBC), ΔmcrA, λ⁻) were grown in Luria-Bertani (LB) broth (1% weight/volume (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) sodium chloride) at 37°C and 200 RPM in a LabNet 311DS shaking incubator. Bacterial colonies were grown on Petri dishes with LB-Agar (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) sodium chloride, 1.5% (w/v) agar) supplemented with the appropriate selection antibiotics. Ampicillin, kanamycin, and zeocin were added to LB broth and LB-Agar media after autoclaving, at 50 µg/mL, 50 µg/mL, and 25 µg/mL, respectively. Stellar™ Competent Cells and Super Optimal Broth with Catabolite Repression (SOC) were purchased from Clontech, Inc. LB-agar, LB broth, and ampicillin sodium salt were purchased from Fisher BioReagents. Kanamycin sulfate and zeocin were purchased from Sigma-Aldrich and InvivoGen, respectively. Petri dishes were purchased from Sterilin Ltd. Additional media supplements were diluted in Milli-Q water with 18.2 MΩ.cm resistivity, produced using Pur1te Select EMD Millipore water purification systems. Instruments were sterilized at 135°C for 5 minutes, and liquids at 121°C for 30 minutes in an MLS-3751L Sanyo autoclave.

4.3.2. *In silico* plasmid design

DNA sequences and plasmid maps were assembled *in silico* with SnapGene® Viewer. All sequences used to assemble the gene constructs referred to in 3.1 were based on the consensus sequences provided by the National Center for Biotechnology Information (NCBI), Agilent Technologies, and Addgene. Basic Local Alignment Search Tool (BLAST) searches were used to confirm this.

4.3.3. Bacterial transformations

A heat shock protocol was followed in order to transform Stellar™ Competent Cells (*E. coli*). In this protocol, 0.1 µL of 1µg/mL plasmid solution was added to sterile 100 µL aliquots of *E. coli* cells and placed on ice for 40 minutes. The cells were then heat shocked at 42°C for 45 seconds and returned to ice for 2 minutes. 450 µL of SOC medium were subsequently added to the mixture, which was placed on a shaking incubator at 200 RPM and 37°C for 1 hour. 50 µL aliquots were then plated on LB-agar Petri dishes supplemented with the appropriate selection antibiotic and incubated at 37°C overnight. Successful transformant single colonies were picked the following day, and plates were stored and maintained at 4°C. Additionally, successful transformants were picked and grown in a 3 mL LB medium overnight culture. 1.4 mL of this culture was supplied to 2 mL cryovials (Thermo Fisher Scientific) already containing 0.5 mL 80% (v/v) glycerol (Sigma-Aldrich) and stored in an -80°C New Brunswick Scientific ultra-low temperature U 725-86 freezer.

4.3.4. Plasmids for system part validation

The Gal4 synthetic gene circuit validation plasmids, pCMVGal4 and pUAS-GG, were purchased from Addgene. The plasmid containing the constitutive *lacI* expression module, pCMVLacI, was purchased from Agilent Technologies as part of the LacSwitch II Mammalian Expression System. Dr Zoltan Kis (Department of Bioengineering, Imperial College London) kindly provided the reporter plasmid in the tTA validation circuit, pZK14. The constitutive expressing module in the tTA validation circuit was constructed using InFusion® cloning (Clontech): pBudCE4.1 (Thermo Fisher Scientific) was used as a backbone and pZK13 (also provided by Dr Zoltan Kis) was used as the template (see 4.4.1). The EF-1 promoter in pBudCE4.1 was removed with NheI/NotI double digestion followed by InFusion®-mediated annealing, with 5'-CCCTCGTCGAGCTAGCAAAGAACCAGCTGG-3' and 5'-ACCTTCGAAGCGGCCGCTTATCATCCGCCAT-3' as the forward and reverse primers, respectively. The corresponding forward and reverse InFusion® primers were 5'-

AAGCAGTACTTCTAGAATGAGATTAGATAAAAAGTAAAGTGATTAAC-3' and 5'-GTTTTTGTTCGGATCCCTACCCACCGTACTCGT-3', and the template 5' and 3' restriction sites were BamHI and XbaI, respectively. The PCR extraction, gel purification and cloning reaction procedures were followed according to the InFusion® Cloning Manual (Clontech). All restriction enzymes and buffers used for this were purchased from New England Biolabs. InFusion® primers were *de novo* synthesized at Sigma-Aldrich with RP1 purification. The reporter plasmid in the LacI validation circuit was constructed by GenScript (Sub-Cloning Services) using pOPRS (Agilent Technologies) as the backbone and pZK14 as the template (see 3.3.1), with the KpnI/NotI restriction site pair.

4.3.5. Plasmid preparation and DNA sequencing

DNA plasmids were extracted from single transformed colonies and purified using the QIAprep Spin Miniprep Kit (Qiagen). For this, single colonies from a freshly streaked selective LB-agar plate were picked and inoculated in 14 mL polypropylene BD Falcon™ tubes containing 5 mL selective LB medium. Liquid cultures were grown overnight on a shaking incubator at 200 RPM and 37°C. Bacterial cells were pelleted after 12-16 hours by centrifugation at 4000 g for 10 minutes in a Hettich Rotanta 460R bench centrifuge at room temperature. Cell pellets were re-suspended using a IKA Lab Dancer S2 vortex shaker in alkaline conditions following the QIAprep Spin Miniprep Kit protocol. The lysate was then neutralized, adjusted to high-salt conditions, and dispensed into a QIAprep spin column. The column was consecutively washed with the appropriate Qiagen buffers and finally eluted with 50 µL of Mili-Q water thus releasing the purified plasmid DNA. Plasmid DNA concentrations and purity were assessed using a Nanodrop™ 2000c Spectrophotometer (Thermo Fisher Scientific). Before measuring samples, the instrument was blanked with 2 µL of Mili-Q water; 2 µL of purified plasmid samples were then loaded into the instrument. DNA purity was assessed by measuring the absorbance ratio at 260 nm and 280 nm: DNA samples were only considered to be pure if the value of this ratio was between 1.8 and 2.0. Likewise, the quality of DNA

samples and presence or absence of contaminants was verified by measuring the absorbance ratio at 260 nm and 230 nm: DNA samples were considered to have an acceptable quality level if this value was between 2.0 and 2.2. DNA samples were Sanger sequenced at GATC Biotech on Sanger ABI 3730xl instruments. Sequencing primers were designed *in silico* using SnapGene® Viewer and *de novo* synthesized at Sigma-Aldrich with RP1 purification. Synthesized primers were re-suspended in Milli-Q water to make up 100 µM stock solutions and stored at -20°C. 20 µL aliquots of 10 µM primer solutions, and 20 µL aliquots of 80 ng/ µL plasmid solutions were sent to GATC for the sequencing reactions to be carried out. The resulting DNA sequences were analyzed with SnapGene® Viewer.

4.3.6. Endotoxin-free plasmid preparation

Plasmids to be transfected in mammalian cells were prepared using the EndoFree® Plasmid Mega Kit from Qiagen, according to the standardized protocol specifications. The initial step in this protocol involved growing a 5 mL bacterial liquid culture in LB medium supplemented with the appropriate antibiotic for 8 hours, in a shaking incubator at 200 RPM and 37°C. 1 mL of this starter culture was then added to 1 L of selective LB medium in a 2.5 L Erlenmeyer polycarbonate flask, and grown overnight in an IOC400 Gallenkamp orbital shaking incubator at 200 RPM and 37°C. Cultures were pelleted after 14 hours by centrifugation in 50 mL polypropylene BD Falcon™ tubes at 4000 g for 15 minutes, using a Hettich Rotanta 460R centrifuge. Pellets were re-suspended using an IKA Lab Dancer S2 vortex shaker and lysed under alkaline conditions. The lysate was then filtered through a QIAfilter Mega Cartridge using a vacuum pump (Cole Parmer Instruments). This was followed by addition of the endotoxin removal buffer and loading of the mixture in a Qiagen Anion-Exchange column in order to separate plasmid DNA from RNA, protein and other impurities. Plasmid DNA was eluted from this column into 50 mL polypropylene BD Falcon™ tubes using a high salt buffer. Isopropanol was added (at 0.7 volume) to the purified DNA with the objective of further concentrating and desalting the solution. Flasks were violently shaken immediately after the addition of isopropanol, and centrifuged

at 4500 g for 1 hour at 4°C in a Hettich Rotanta 460R centrifuge. The resulting DNA pellet was further washed with room temperature 70% ethanol and centrifuged at 4500 g for 1 hour at 4°C in a Hettich Rotanta 460R centrifuge. The final DNA pellets were air dried at room temperature for 30 minutes and re-suspended in TE buffer or endotoxin-free water.

4.3.7. Agarose gel electrophoresis

The size and structural integrity of plasmids was verified by 1% agarose gel electrophoresis using an Owl™ Easycast™ Mini Gel Electrophoresis System (Thermo Fisher Scientific). 1% agarose solutions were prepared by dissolving 1 g of Agarose (Sigma-Aldrich) in 100 mL of Tris-Borate-EDTA (TBE) buffer (Sigma-Aldrich), and further heated in a Russel Hobbs GTS23 microwave for approximately 3 minutes. The resulting clear agarose solution was kept in Grant Instruments GD120 water bath at 55°C for 30 minutes. 10 µL of SYBR® Safe DNA Gel Stain (Life Technologies) were then added to the solution, which was subsequently poured into an appropriate gel tray with combs in place. The gel tray was covered in tin foil due to the photosensitivity of the SYBR® Safe DNA Gel Stain and left at room temperature for 45 minutes to allow the gel to solidify. The gel tray was then loaded in its corresponding electrophoresis chamber, which was appropriately filled with 0.5x TBE buffer. Restriction enzyme-digested plasmid samples were supplied with 6x Purple Loading Dye (NEB) at a 2 µL dye to 50 µL sample ratio; 10 µL of each sample and 5 µL of a 1 kbp DNA ladder were loaded per comb. Electrophoresis was then initiated at constant voltage (180 V) for 45 minutes. Gels were visualized under low intensity 254 nm UV light using a UVP BioSpectrum 500 Imaging System.

4.3.8. Mammalian cell culture conditions

The Human Embryonic Kidney cell line (HEK293T) was kindly provided by Dr James Pease (National Heart and Lung Institute, Imperial College London). HEK cells were

grown in DMEM supplemented with final concentrations of 5 mM L-Glutamine, 10 mM HEPES, 10% (v/v) fetal bovine serum (FBS), 100 µg/mL streptomycin and 100 units/mL penicillin. Cells were cultured in Corning® polystyrene flasks with 25 cm² or 75 cm² CellBIND® cell growth surface area; culture flasks were incubated in MCO-18AC, MCO-18AIC or MCO-18M Sanyo incubators at 37°C with saturated humidity and 5% (v/v) CO₂. Cells were passaged and handled in a Euroclone Bioair Top Safe 1.5 safety cabinet. To passage the cells, culture flasks were washed with a 37°C 1x PBS solution twice (1 mL per 10 cm² surface growth area) after disposing the complete DMEM inside. Cells were detached from the flask surface by trypsinization with 0.4 mL of 0.25% (w/v) trypsin, 0.02% EDTA solution per 10 cm² cell growth area and subsequent incubation at 37°C for 3 minutes. Cell detachment was confirmed by visualizing the cells under a Leica DM IL LED-DFC295 inverted phase contrast microscope. The trypsinization reaction was subsequently inhibited by adding 10 mL of complete DMEM per 1 mL of trypsin solution; the mixture was then transferred to 14 mL CentriStar™ capped tubes and pelleted by centrifugation for 5 minutes at 220 g in a MSE Mistral 2000 centrifuge. Pellets were re-suspended in complete DMEM and transferred to new culture flasks. Liquid volumes from 1 mL to 25 mL were handled with polystyrene Stripettes®; volumes of 1 µL to 1 mL were handled with StarLab TipOne® Filter pipette tips. All working area surfaces were sterilized throughout the protocol with 70% (v/v) ethanol (NORMAPUR®). DMEM and all medium additives were purchased from Life Technologies.

4.3.9. Turbofect™ transfection and *in vivo* implementation of system components

HEK cells were transfected using the Turbofect™ Transfection reagent (Thermo Fisher Scientific). For this, cells were seeded 24 hours prior transfection in 48-well plates containing 500 µL complete DMEM, at 40-50% confluency thus achieving 70-80% confluency at the time of transfection. 1 µg of plasmid was added to 2 µL of transfection reagent and further diluted in 100 µL of serum-free Opti-MEM® medium (Life Technologies). This solution was incubated at room temperature for 45 minutes

and 100 μ L were added to each well, in a drop-wise manner. The 48-well plate was then gently rocked to ensure even distribution of DNA complexes and placed in a MCO-18AC, MCO-18AIC or MCO-18M Sanyo incubator at 37°C for 24 hours. The effects of tetracycline and IPTG on the tTA and LacI validation circuits were respectively assessed by supplying the corresponding wells with doxycycline (Sigma-Aldrich) and IPTG (Agilent Technologies) sterile-filtered solutions at the appropriate concentrations. These were supplemented to the wells 1 hour before the time-lapse image acquisition protocol. The 48-well plates were then moved to the FILM suite at Imperial College London and were kept at 37°C, constant humidity and 5% CO₂ in an imaging chamber throughout the image acquisition protocol. Positive controls were created by transfecting cells with a constitutively expressing *gfp* plasmid, i.e. pZK11.

4.3.10. Microscopy imaging and image processing

Cells were visualized under a Leica DM IL LED-DFC295 inverted phase contrast microscope prior to the image acquisition protocol to confirm optimal confluency and that no contamination was present. Cells were then imaged using a Hamamatsu ORCA-ER camera coupled to a Zeiss Axiovert 200 inverted fluorescent microscope with a fully motorized stage, controlled by Improvion Volocity acquisition software. The image acquisition protocol was initiated by adjusting the brightfield Köhler illumination and setting up the Zeiss EC Plan-Neofluar 10x 0.30 Ph1 objective (Phase Contrast 1 channel – Ph1) and Zeiss Filter Set 10 (FITC) with a mercury fluorescence lamp (Fluorescent channel 1). The motorized stage was then calibrated, and one randomly selected region of interest (ROI) was selected for each well. Images were then automatically acquired for 22 hours at a rate of 1 image / hour / well in the brightfield and fluorescent channels. Brightfield channel image acquisition was performed with an exposure time of 5 milliseconds (ms); fluorescent channel image acquisition was performed with an exposure time of 15 ms. Images were analyzed using the Fluorescence Measurement Thresholder in Volocity, thus allowing for automated detection of fluorescent regions and measurement of fluorescence intensity in Relative Fluorescence Units (RFU). Fluorescent cells were analysed on a

single-cell basis. The fluorescence intensity time evolution data shown in this chapter was generated by summing the mean GFP intensity of each cell per replicate and per time point, allowing for a representation of both fluorescence intensity and number of fluorescent cells. Data from different replicates was then averaged, and the standard deviation was subsequently calculated.

4.4. Results and discussion

4.4.1. Construction of synthetic gene circuits for validation of system components

The *in vitro* validation of system components was initiated by constructing three synthetic gene circuits, each made up of two plasmids: one responsible for constitutive expression of a system core component (i.e. LacI, tTA, Gal4), and another responsible for providing a corresponding inducible readout.

The LacI validation synthetic gene circuit was constructed by coupling pCMVLacI (Addgene), a plasmid constitutively expressing the LacI protein, to pOPRS_GFP, a plasmid containing two *lacO* operons upstream of a GFP ORF. pOPRS_GFP was constructed using pOPRS (Agilent Technologies) as the backbone and the GFP ORF in pZK14 (provided by Dr Zoltan Kis) as the template. The GFP ORF was cloned in pOPRS using the restriction sites KpnI and NotI (Genscript, Inc.). Plasmid maps of pCMVLacI and pOPRS_GFP are provided in figures 4.3 and 4.4, respectively. The Gal4 validation synthetic gene circuit was constructed by coupling pCMVGal4 (Addgene), a plasmid constitutively expressing the Gal4 protein, to pUAS-GG (Addgene), a plasmid containing a UAS-regulated *gfp* gene. The structure of these plasmids is illustrated in figures 4.5 and 4.6. The DNA sequences for pCMVLacI, pOPRS_GFP, pCMVGal4, pUAS-GG and pOPRS are presented in *Appendix A.4*.

The construction of pCMVLacI, pOPRS_GFP, pCMVGal4 and pUAS-GG was validated *in vitro* through restriction enzyme digestion followed by agarose gel electrophoresis. To this end, these four plasmids were individually restriction enzyme digested with XbaI, NheI, HindIII, and NotI, respectively, and ran on an agarose gel according to the protocol specified in 4.3.7. The results of this procedure corroborate the expected plasmid sizes (7995 base pairs (bp), 6334 bp, 8100 bp, and 4467 bp, respectively), and are illustrated in figure 4.7.

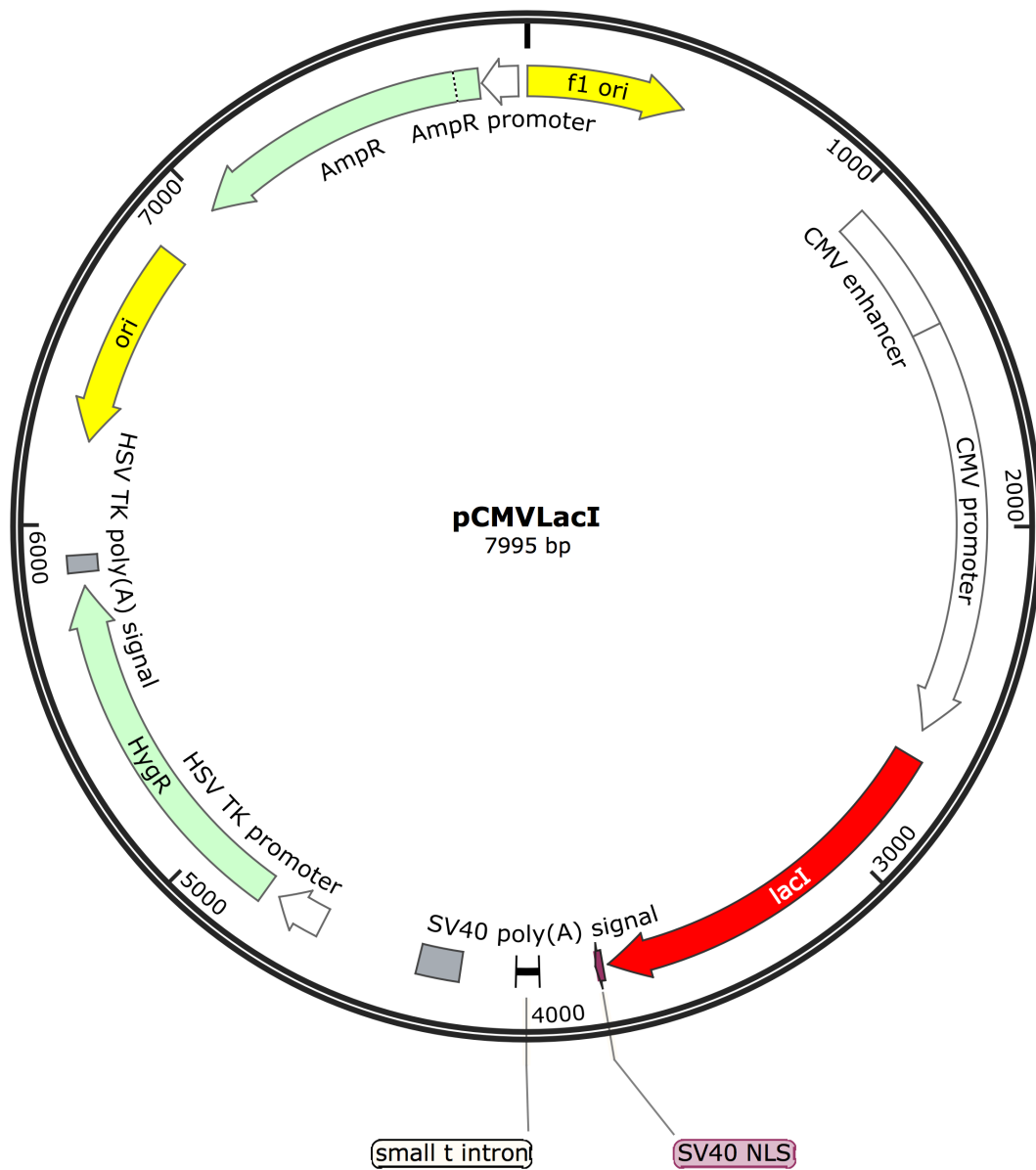


Figure 4.3. pCMVLacI plasmid map. Constitutive expression of the *lacI* gene is driven by a CMV promoter / enhancer region located upstream of the *lacI* ORF. Additionally, SV40 nuclear localization and poly-adenylation signals ensure nuclear translocation and mRNA maturation, respectively.

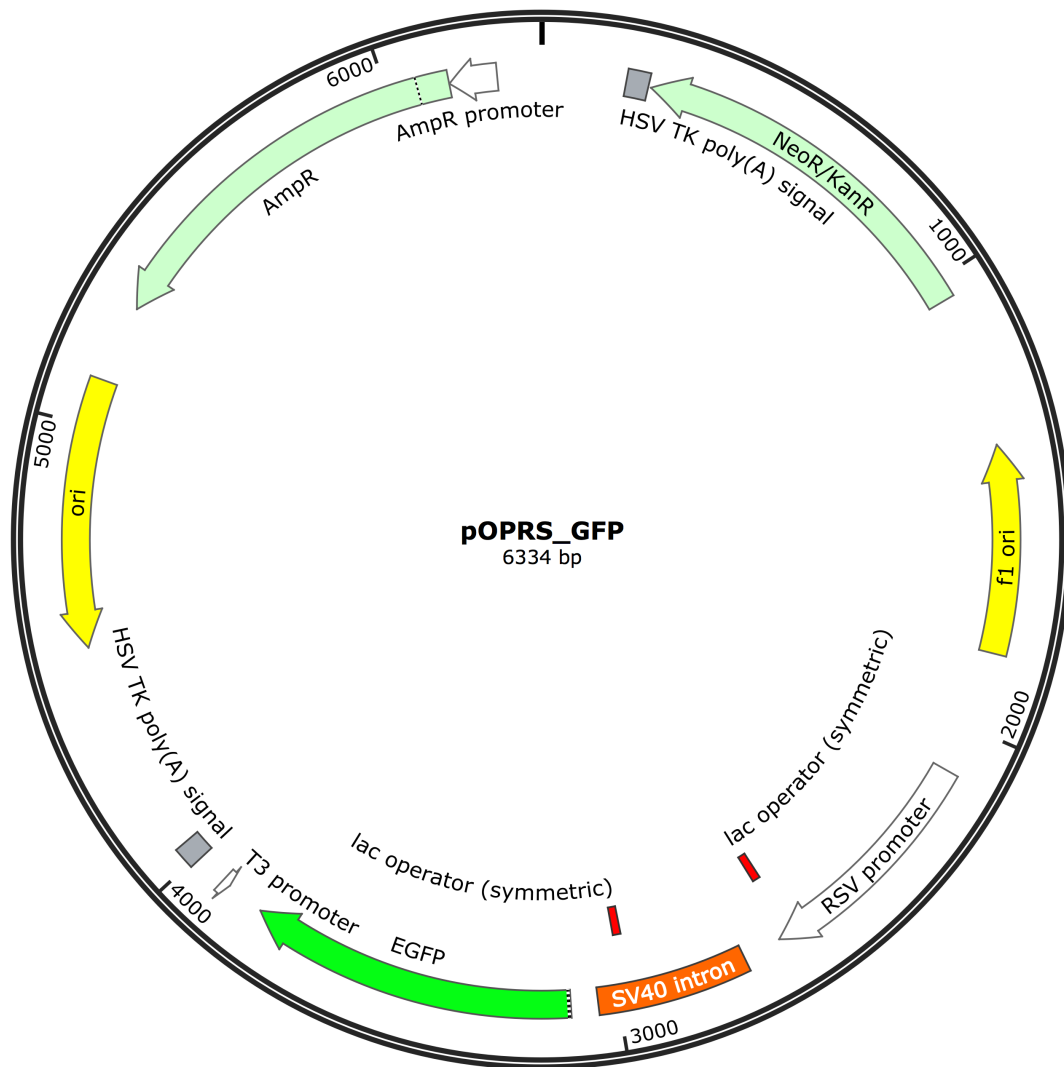


Figure 4.4. pOPRS_GFP plasmid map. Constitutive expression of the *gfp* gene is driven by an RSV promoter located upstream of the *gfp* ORF. Additionally, an HSV poly-adenylation signal ensures mRNA maturation. Two *lacO* operator sites are located between the *gfp* ORF and the RSV promoter. These sites allow for LacI binding and subsequent expression inhibition.

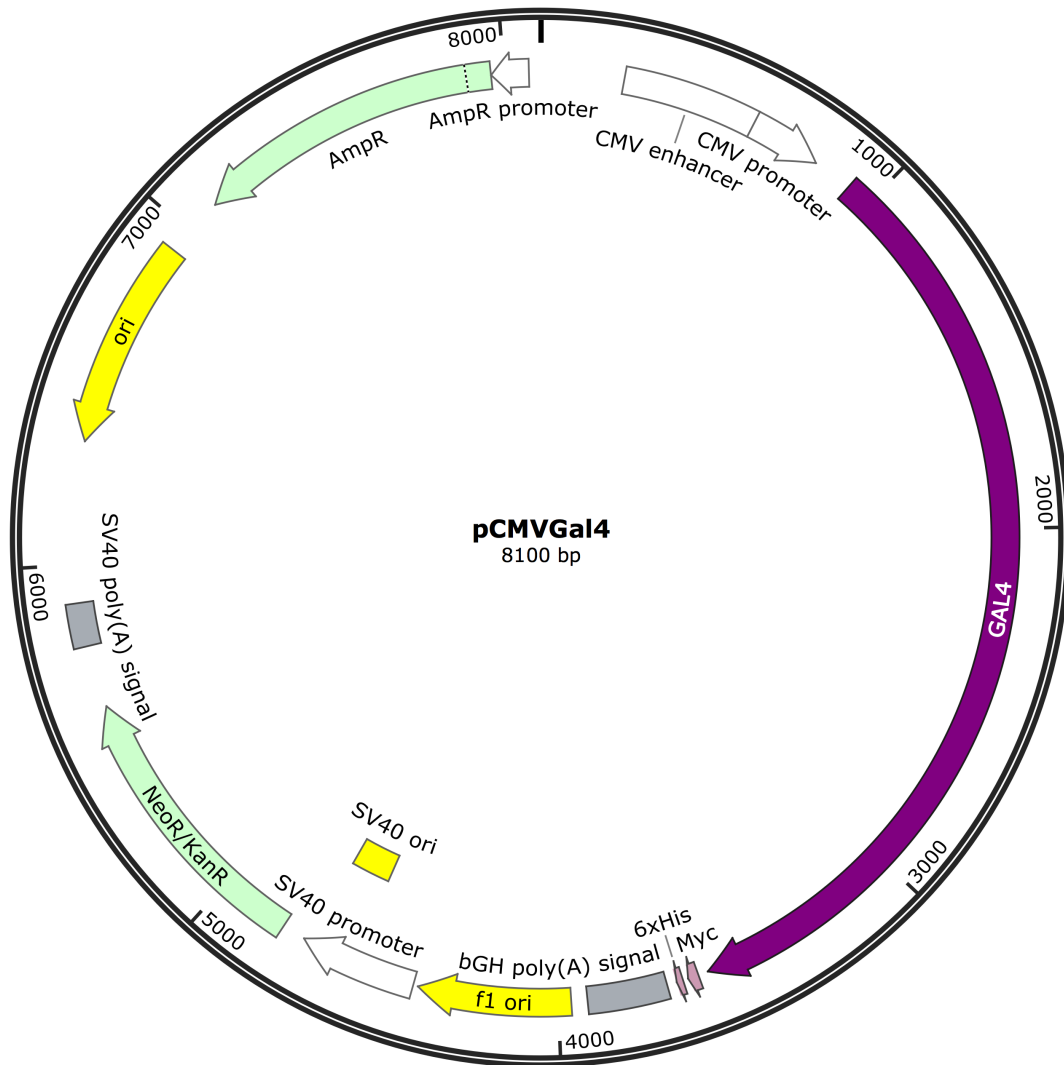


Figure 4.5. pCMVGal4 plasmid map. Constitutive expression of the *gal4* gene is driven by a CMV promoter / enhancer region located upstream of the *gal4* ORF. Additionally, a bGH poly-adenylation signal ensures mRNA maturation.

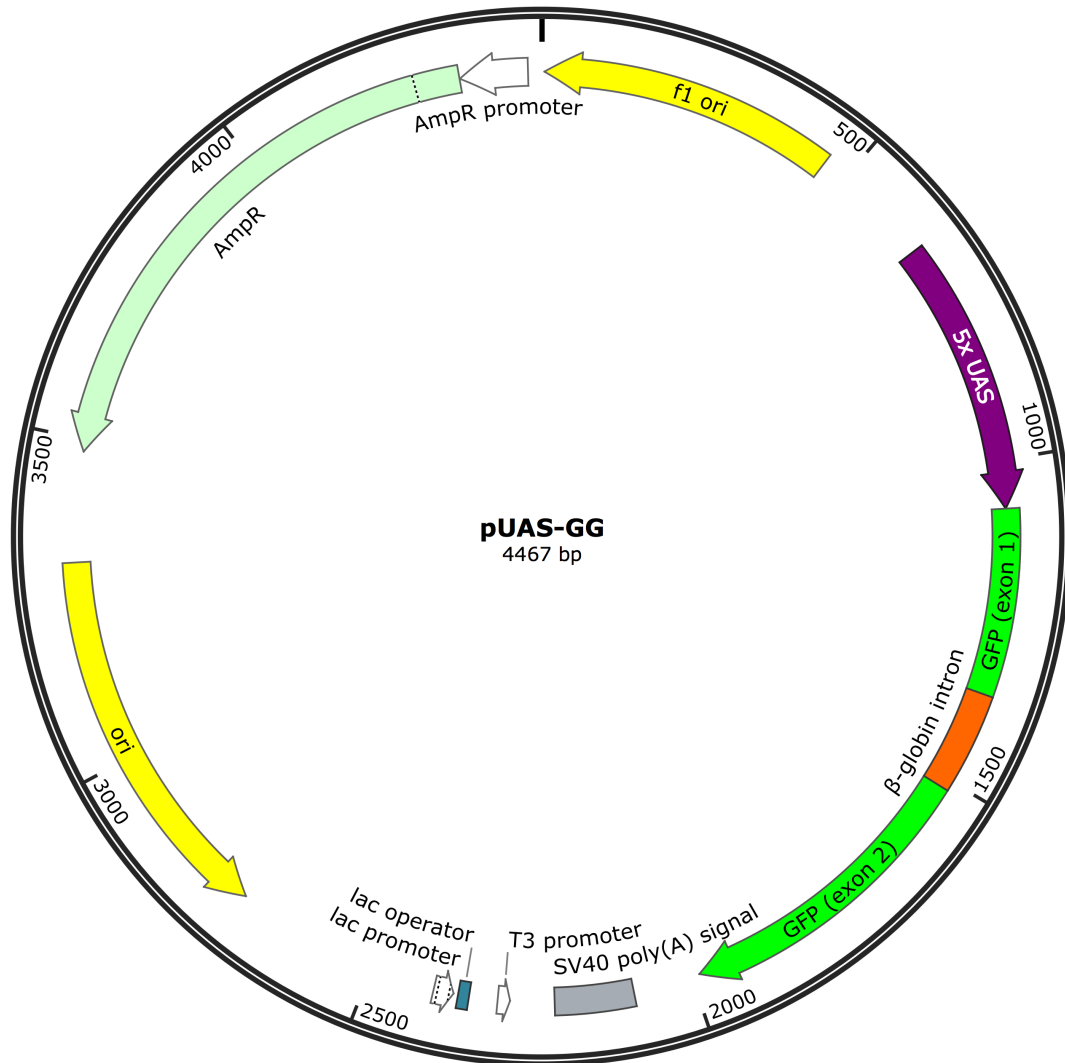


Figure 4.6. pUAS-GG plasmid map. Constitutive expression of the *gfp* gene is driven by a Gal4-responsive UAS region located upstream of the *gfp* ORF. Additionally, an SV40 nuclear poly-adenylation signal ensures nuclear mRNA maturation.

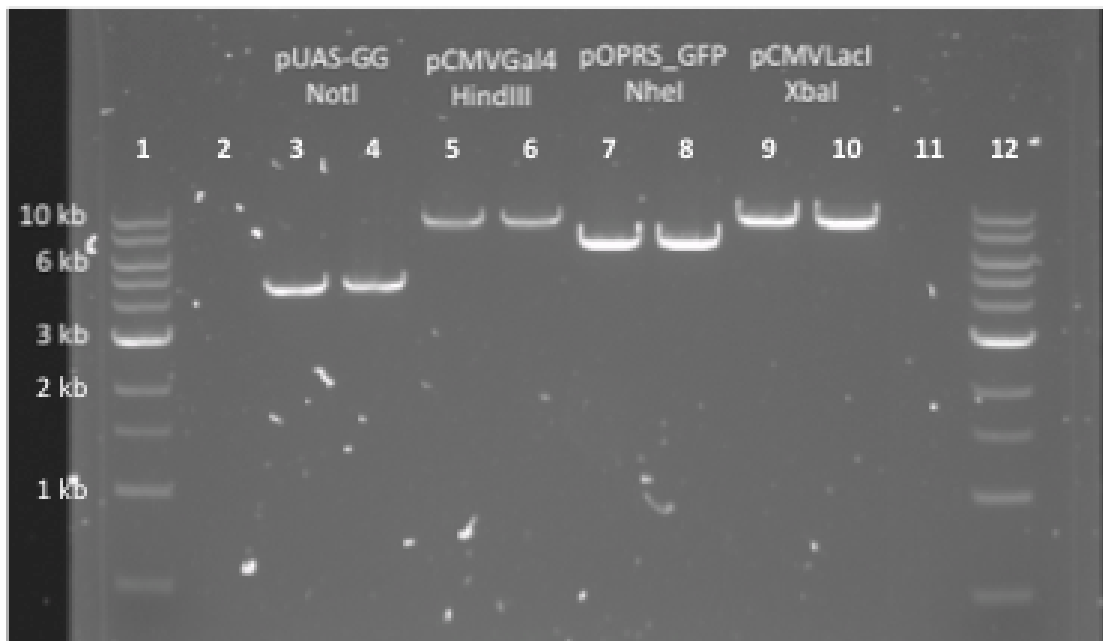


Figure 4.7. Agarose gel electrophoresis of pUAS-GG (third and fourth lanes), pCMVGal4 (fifth and sixth lanes), pOPRS_GFP (seventh and eighth lanes) and pCMVLacI (ninth and tenth lanes), after single restriction enzyme digestion. The corresponding plasmid sizes are 4467 bp, 8100 bp, 6334 bp, and 7995 bp, respectively. The molecular weight indicator (i.e. 1 kb DNA Ladder) is located in lanes 1 and 12.

The tTA validation synthetic gene circuit was constructed by coupling pCMVtTA, a plasmid constitutively expressing the tTA protein, with pZK14, a reporter plasmid containing a P_{TRE3G} promoter regulating expression of a downstream *gfp* gene. The pCMVtTA and pZK14 plasmid maps are shown in figures 4.8 and 4.9, respectively. The complete DNA sequences of these plasmids are presented in *Appendix A.4*. pCMVtTA was constructed using a pBudCE4.1 mammalian expression vector (Thermo Fisher Scientific) without the EF-1 promoter as the backbone (i.e. 3429 bp) and the *tTA* ORF from pZK13 (provided by Dr Zoltan Kis) as the template.

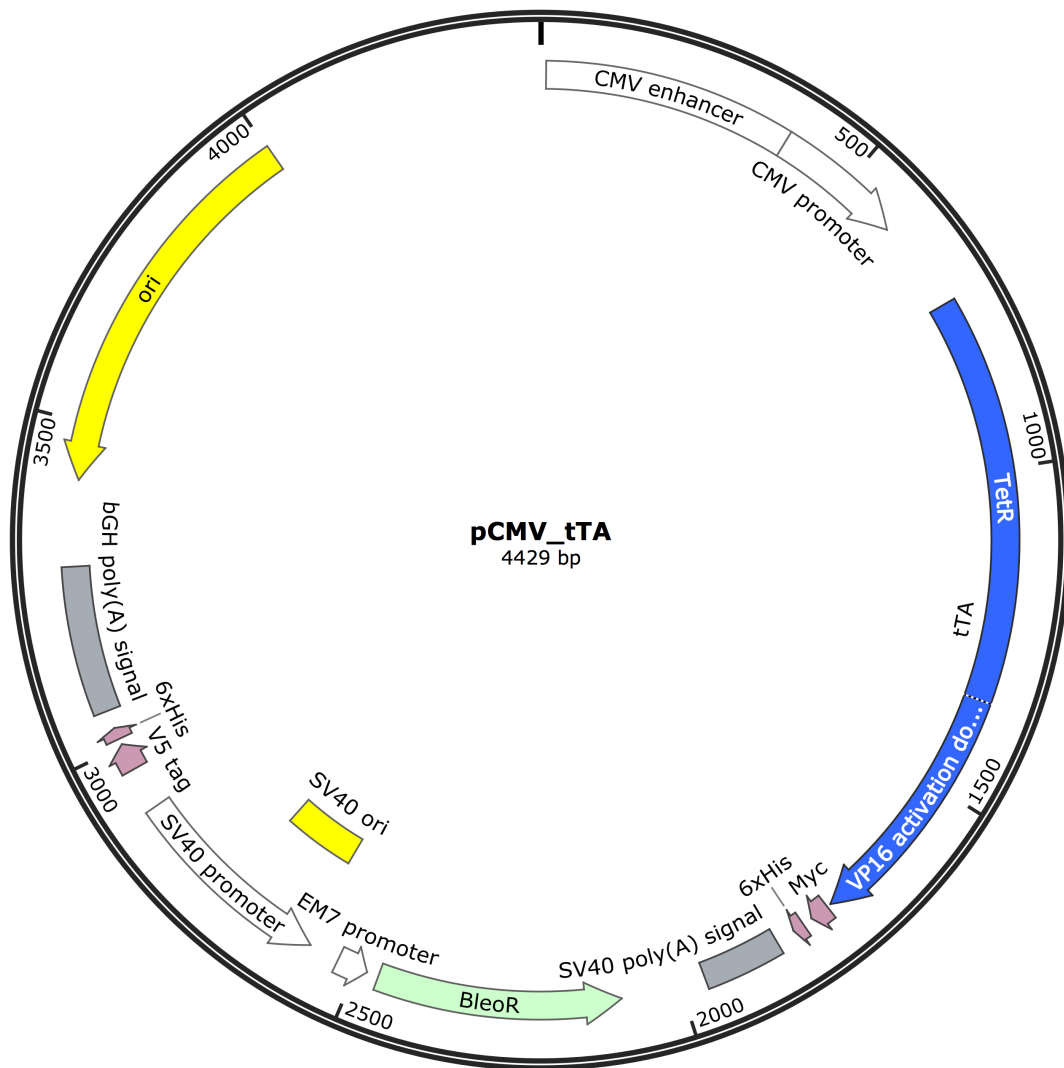


Figure 4.8. pCMVtTA plasmid map. Constitutive expression of the *tTA* gene is driven by a CMV promoter / enhancer region located upstream of the *tTA* ORF. Additionally, an SV40 poly-adenylation signal ensures mRNA maturation.

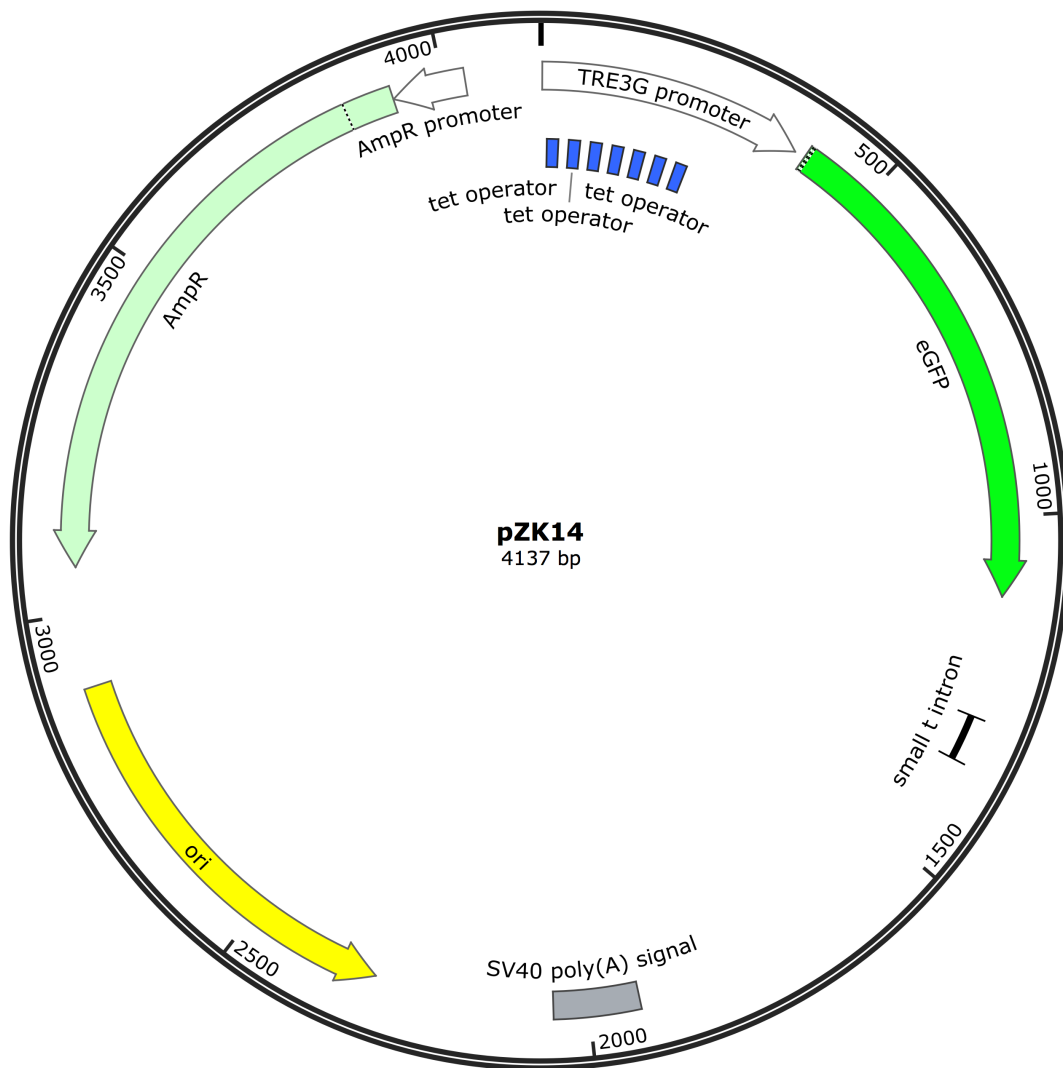


Figure 4.9. pZK14 plasmid map. A TRE3G promoter made up of seven tetracycline responsive elements upstream of the *gfp* ORF regulates expression of GFP in a tTA-responsive manner. A downstream SV40 poly-adenylation signal ensures mRNA maturation.

Construction and in vivo validation of system parts

To this end, the *tTA* ORF in pZK13 was amplified through PCR and subsequently inserted in pBudCE4.1 through InFusion® cloning technology, using the XbaI and BamHI restriction sites. Figure 4.10 illustrates the PCR extraction of the *tTA* ORF and the size validation of pCMVtTA on an agarose gel. It is clear from figure 4.10 that both the InFusion® PCR product (i.e. *tTA* ORF) and pCMVtTA present the expected sizes, i.e. c. 1 kbp and 4.4 kbp, respectively. The success of the cloning reaction was further confirmed by double-digesting pCMVtTA with XbaI and BamHI, the two restriction sites flanking the insert in the target vector.

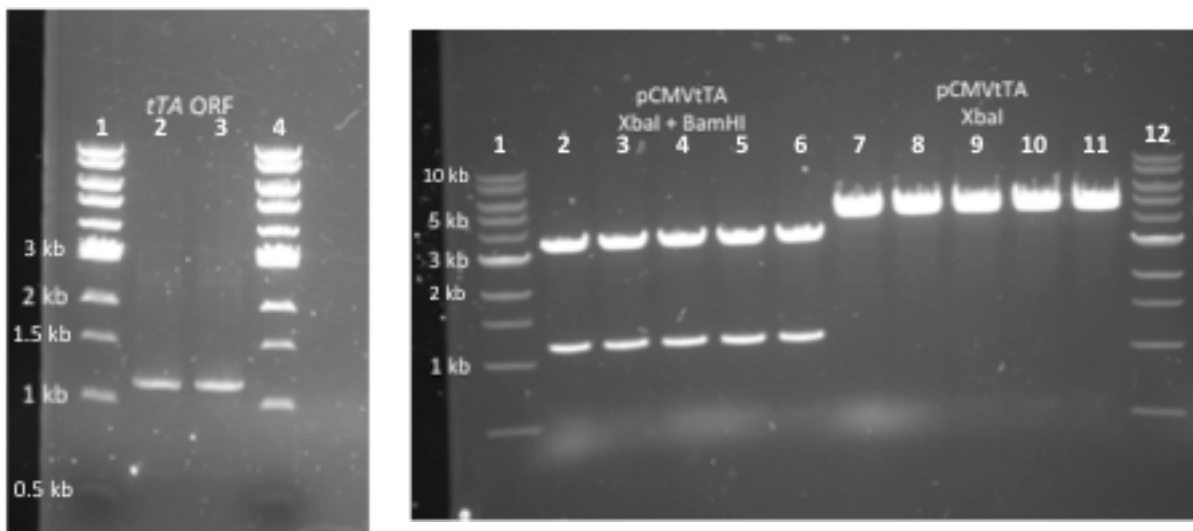


Figure 4.10. (Left) InFusion® PCR-mediated extraction of the *tTA* ORF from pZK13. The first and fourth lanes contain the molecular weight indicator (i.e. 1 kb DNA Ladder). The DNA fragments corresponding to the *tTA* ORF are located in lanes 2 and 3, and have an approximate size of 1000 bp. (Right) pCMVtTA size validation agarose gel. Lanes 1 and 12 contain the molecular weight indicator. Lanes 2 to 6 contain the product of a pCMVtTA digested with XbaI and BamHI, thus releasing the *tTA* ORF insert. Lanes 7 to 11 contain the result of a pCMVtTA digestion with XbaI, thus highlighting the size of pCMVtTA in its linearized form (i.e. c. 4.4 kbp).

The reporter module in the tTA validation synthetic gene circuit is encoded in pZK14. This plasmid was provided by Dr Zoltan Kis, and its size was also validated using agarose gel electrophoresis after a restriction enzyme digestion with BamHI (figure 4.11). As indicated in figure 4.9, this plasmid was expected to have a size of *c.* 4.1 kbp. The results in figure 4.11 corroborate this: a single band containing the linearized pZK14 vector is present immediately after the 4 kbp region. Figure 4.11 illustrates the size of pZK14 together with that of pZK11 (also provided by Dr Zoltan Kis), a constitutively expressing *gfp* plasmid with a size of *c.* 4.7 kbp.

The six plasmids used for the *in vivo* validation experiments (i.e. pCMVGal4, pUAS-GG, pCMVLacI, pOPRS_GFP, pCMVtTA and pZK14) were further amplified in *E.coli* and purified in an endotoxin-free manner before their transfection in HEK293T cells. The obtained concentration and purity (using a Nanodrop™ 2000c Spectrophotometer) of each of these plasmids is outlined in table 4.1.



Figure 4.11. Agarose gel electrophoresis result of a pZK11 and pZK14 restriction enzyme digestions with HindIII and BamHI, respectively. Lanes 1 and 12 contain the molecular weight indicator; lanes 2 to 6 contain the linearized pZK11; lanes 7 to 11 contain the linearized pZK14.

Construction and *in vivo* validation of system parts

All plasmids showed an acceptable concentration in the context of mammalian cell transfection. Indeed, the lowest concentration obtained was that of pCMVGal4 (531 ng/uL). Furthermore, the obtained absorbance ratios at 260 nm / 280 nm and 260 nm / 230 nm were all in the range of 1.8-1.9 and 2.0-2.11. This indicates that both the amplification and purification protocols were successful.

Table 4.1. Concentration and purity of plasmids encoding system parts to be validated through synthetic gene circuit *in vivo* implementation.

Plasmid	Concentration (ug/uL)	A260 / A280	A260/A230
pCMVLacI	1.1	1.89	2.08
pOPRS_GFP	0.816	1.88	2.11
pCMVGal4	0.531	1.85	2.04
pUAS-GG	0.793	1.86	2.07
pCMVtTA	2.3	1.83	2.02
pZK14	3.9	1.89	2.04

4.4.2. *In vivo* implementation of the LacI validation synthetic gene circuit

The *in vivo* validation of both the *lacI* ORF and the *lacO* operator site structure proposed in 4.2 was carried out by transfecting the LacI validation synthetic gene circuit in HEK cells. To this end, cells were subject to three different experimental protocols: (i) co-transfection with pCMVLacI and pOPRS_GFP, (ii) transfection with pOPRS_GFP, and (iii) co-transfection with pCMVLacI and pOPRS_GFP in cells supplemented with 1 mM IPTG. In this experimental setup pOPRS_GFP is expected to drive constitutive expression of the *gfp* gene from an RSV promoter, resulting in

the emission of fluorescence. Co-transfection with pCMVLacI and pOPRS_GFP in the absence of IPTG is expected to abolish any fluorescence due to accumulation of LacI, constitutively expressed from a CMV promoter in pCMVLacI, and subsequent repression of the *gfp* gene in pOPRS_GFP through binding of the two upstream *lacO* operator sites. Conversely, cells co-transfected with pCMVLacI and pOPRS_GFP in the presence of IPTG are expected to emit fluorescence due IPTG-mediated inhibition of accumulated LacI, thus alleviating *gfp* repression at the upstream *lacO* operator sites.

The fluorescence images acquired during the *in vivo* implementation of this synthetic gene circuit are shown in figure 4.12. Images were acquired with a sampling frequency of 1 image / hour, for 24 hours; the first image was acquired 24 hours after transfection. Figure 4.12 shows the fluorescence intensity in HEK cells (24, 32 and 46 hours after transfection) resulting from (i) co-transfection with pCMVLacI and pOPRS_GFP, (ii) transfection with pOPRS_GFP, and (iii) co-transfection with pCMVLacI and pOPRS_GFP in cells supplemented with 1 mM IPTG. Images were also acquired using the brightfield channel; the superimposed brightfield-fluorescence images are shown in *Appendix A.4*. The fluorescence time evolution profile of cells co-transfected with pCMVLacI and pOPRS in DMEM medium supplemented with 1 mM IPTG is presented in figure 4.13.

A clear progressive increase in fluorescence intensity can be observed in cells transfected with pOPRS_GFP, thus corroborating the expected result for this experimental condition. RSV-driven expression of the *gfp* gene is constitutive and uninhibited by LacI, which explains the early fluorescence emission. Moreover, a progressive increase in the number of fluorescence cells can be observed, together with an overall increase in fluorescence intensity. The time discrepancy in the start of fluorescence emission for different cells can be explained by (i) uneven plasmid uptake rates across the cellular monolayer, and (ii) the fact that morphological and cell cycle differences across the cellular monolayer lead to different plasmid expression rates. In contrast with the fluorescence profile observed for cells transfected with pOPRS_GFP, cells that were co-transfected with pCMVLacI and pOPRS_GFP do not show any fluorescence during the 24 hours of image acquisition,

also corroborating the expected results for this experimental condition. Indeed, the constitutive expression of *lacI*, driven by a CMV promoter, and the subsequent binding of the LacI repressor to the *lacO* operator sites upstream of the *gfp* gene in

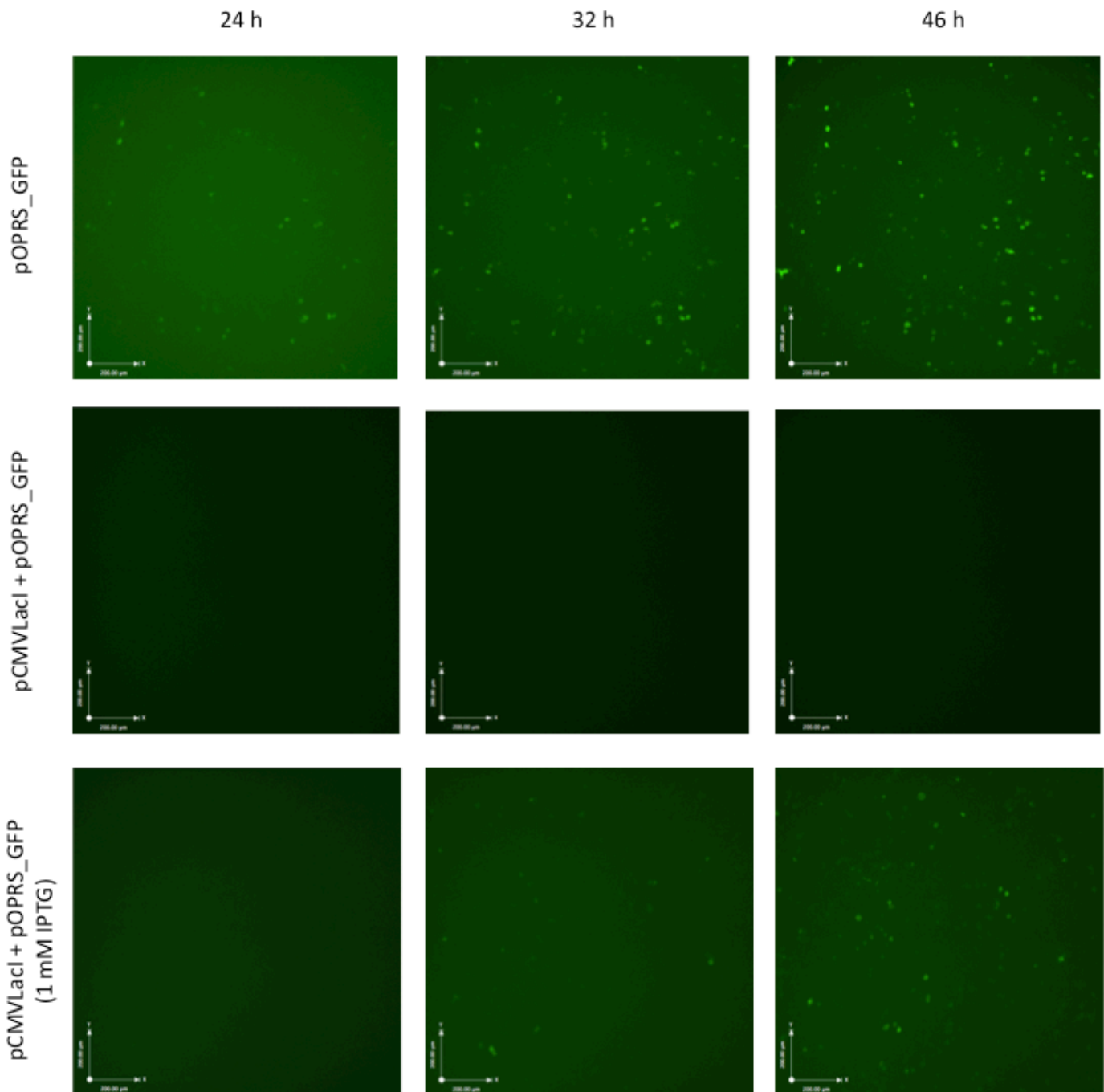


Figure 4.12. Fluorescence channel images of HEK cells transfected with LacI validation synthetic gene circuit components. White arrows on the bottom left of each image represent a 200 μm distance (Top) Cells transfected with pOPRS_GFP 24, 32 and 46 hours after transfection. (Middle) Cells co-transfected with pCMVLacI and pOPRS_GFP 24, 32 and 46 hours after transfection. (Bottom) Cells co-transfected with pCMVLacI and pOPRS_GFP in DMEM medium supplemented with a final concentration of 1 mM IPTG.

pOPRS_GFP result in a complete inhibition of fluorescence emission. Remarkably, no basal expression was observed in this experimental condition, thus confirming that the LacI-*lacO* repressor system is capable of tight gene expression regulation in HEK cells (Aubrecht et al., 1996; Cronin et al., 2001; Luo et al., 1999).

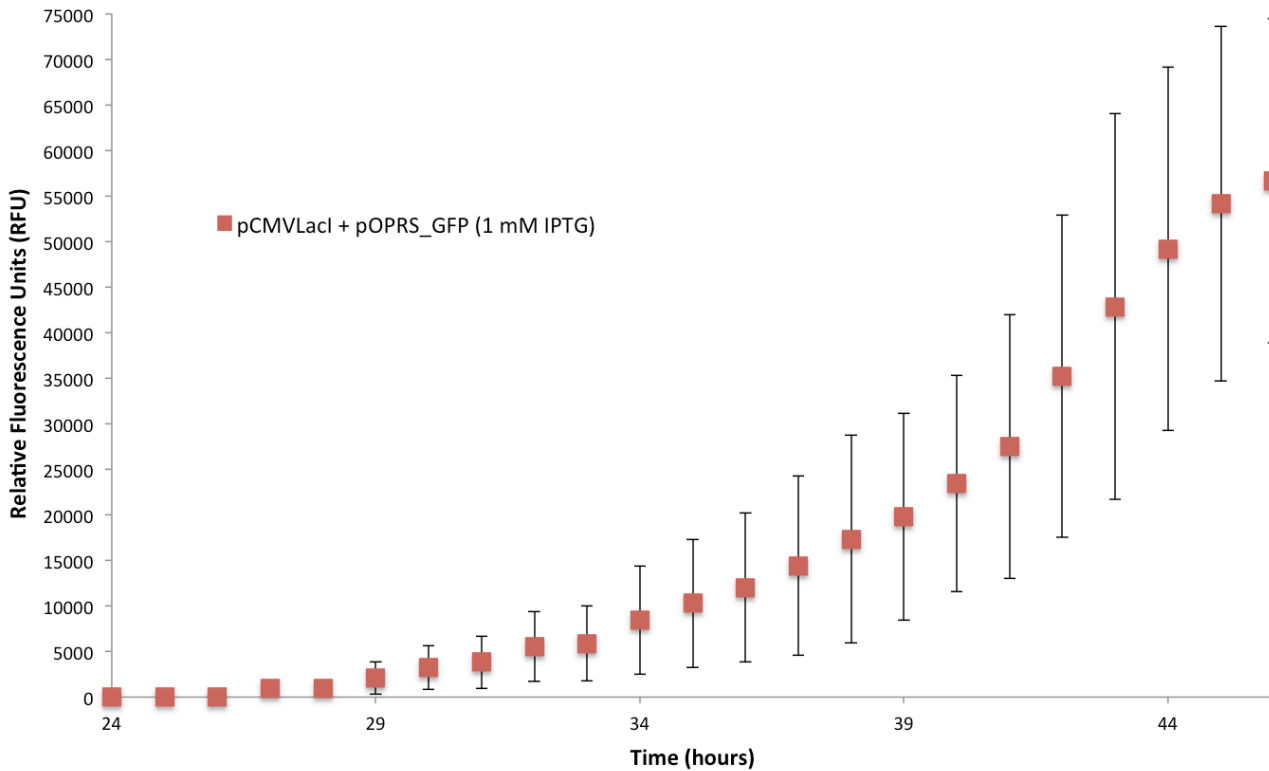


Figure 4.13. Fluorescence intensity time evolution of HEK293T cells co-transfected with pCMVLacI and pOPRS_GFP, in DMEM medium supplemented with 1 mM IPTG. The data presented herein is representative of 185 fluorescent cells (across 4 replicate wells). The error bars represent the standard deviation amongst replicates.

Furthermore, a progressive increase in the number of fluorescent cells as well as in fluorescence intensity was observed in cells that were co-transfected with pCMVLacI and pOPRS_GFP in medium supplemented with 1 mM IPTG (figure 4.13). In this experimental condition, the LacI repressor accumulates due to constitutive expression driven by a CMV promoter in pCMVLacI. The progressive increase in fluorescence indicates that the presence of IPTG results in inhibition of the LacI repressor, which no longer binds to the two *lacO* operator sites upstream of the *gfp*

gene in pOPRS_GFP. This results in a gradual alleviation of the LacI repressor activity and a consequent increase in fluorescence emission. This is corroborated by the clear time delay in the start of fluorescence emission within this experimental condition in comparison to pOPRS_GFP-transfected cells. Indeed, the number of fluorescent cells and the relative fluorescence intensity observed in this experimental condition at 46 hours after transfection is directly comparable to that of pOPRS_GFP transfected cells at 32 hours after transfection. This confirms a key feature of the LacI validation synthetic gene circuit: the ability of IPTG to alleviate LacI-mediated repression in mammalian cells.

The results shown in figures 4.12 and 4.13 indicate that the *in vivo* implementation of the LacI validation synthetic gene circuit successfully generated the expected biological behaviour of the constructs encoded in pCMVLacI and pOPRS_GFP. The Lac repressor system has previously been shown to be functional in prokaryotes, yeast, zebrafish, mouse models, HeLa cells, CHO cells, mouse L TK⁻ cells, HepG2 cells and mouse Swiss 3T3 fibroblasts (Aubrecht et al., 1996; Cronin et al., 2001; Luo et al., 1999; Lutz & Bujard, 1997; Wyborski & Short, 1991). It is of critical importance to confirm that this system is also functional in HEK cells before proceeding with the *in vivo* implementation of the proposed synthetic gene oscillator in this cell line. This is particularly relevant if one considers the variability in the behaviour of synthetic gene circuits when implemented in different mammalian cell lines (Ausländer et al., 2012; Weber & Fussenegger, 2009). Indeed, the results highlighted in figure 4.12 corroborate that: (i) the consensus *lacI* ORF in pCMVLacI is capable of generating a functional LacI repressor protein; (ii) the *lacO* operator site structure in pOPRS_GFP is targeted by the LacI repressor; (iii) IPTG is capable of alleviating LacI-mediated repression.

4.4.3. *In vivo* implementation of the tTA validation synthetic gene circuit

The *in vivo* validation of both the *tTA* ORF and the *tetO* operator site structure (i.e. P_{TRE3G}) proposed in 4.2 was carried out by transfecting the tTA validation synthetic gene circuit in HEK cells. To this end, cells were subject to three different experimental protocols: (i) co-transfection with pCMVtTA and pZK14, (ii) transfection with pZK14, and (iii) co-transfection with pCMVtTA and pZK14 in cells supplemented with 1 $\mu\text{g}/\text{mL}$ doxycycline. In this experimental setup, pCMVtTA is expected to drive constitutive expression of the *tTA* gene from an upstream CMV promoter, consequently leading to up-regulation of the *gfp* gene in pZK14 through binding of the *tetO* operator sites in P_{TRE3G} , leading to the emission of fluorescence. Conversely, co-transfection with pCMVtTA and pZK14 in the presence of doxycycline is expected to result in a lack of fluorescence emission due to doxycycline-mediated inhibition of tTA. Moreover, transfection with pZK14 is expected to emulate a negative control whereby no tTA protein is present and thus no up-regulation of the *gfp* gene in pZK14 is possible other than that caused by P_{TRE3G} leakiness.

The fluorescence images acquired during the *in vivo* implementation of this synthetic gene circuit are shown in figure 4.13. Images were acquired with a sampling frequency of 1 image / hour, for 24 hours; the first image was acquired 24 hours after transfection. Figure 4.13 shows the fluorescence intensity in HEK cells (24, 32 and 46 hours after transfection) resulting from (i) co-transfection with pCMVtTA and pZK14, (ii) transfection with pZK14, and (iii) co-transfection with pCMVtTA and pZK14 in cells supplemented with 1 $\mu\text{g}/\text{mL}$ doxycycline. Images were also acquired using the brightfield channel; the superimposed brightfield-fluorescence images are shown in *Appendix A.4*. The fluorescence time evolution profiles of cells co-transfected with pCMVtTA and pZK14 in DMEM medium supplemented with and without 1 $\mu\text{g}/\text{mL}$ doxycycline are presented in figure 4.15.

A strong progressive increase in fluorescence intensity and number of fluorescent cells can be observed in cells co-transfected with pCMVtTA and pZK14, thus corroborating the expected result for this experimental condition. Conversely, a

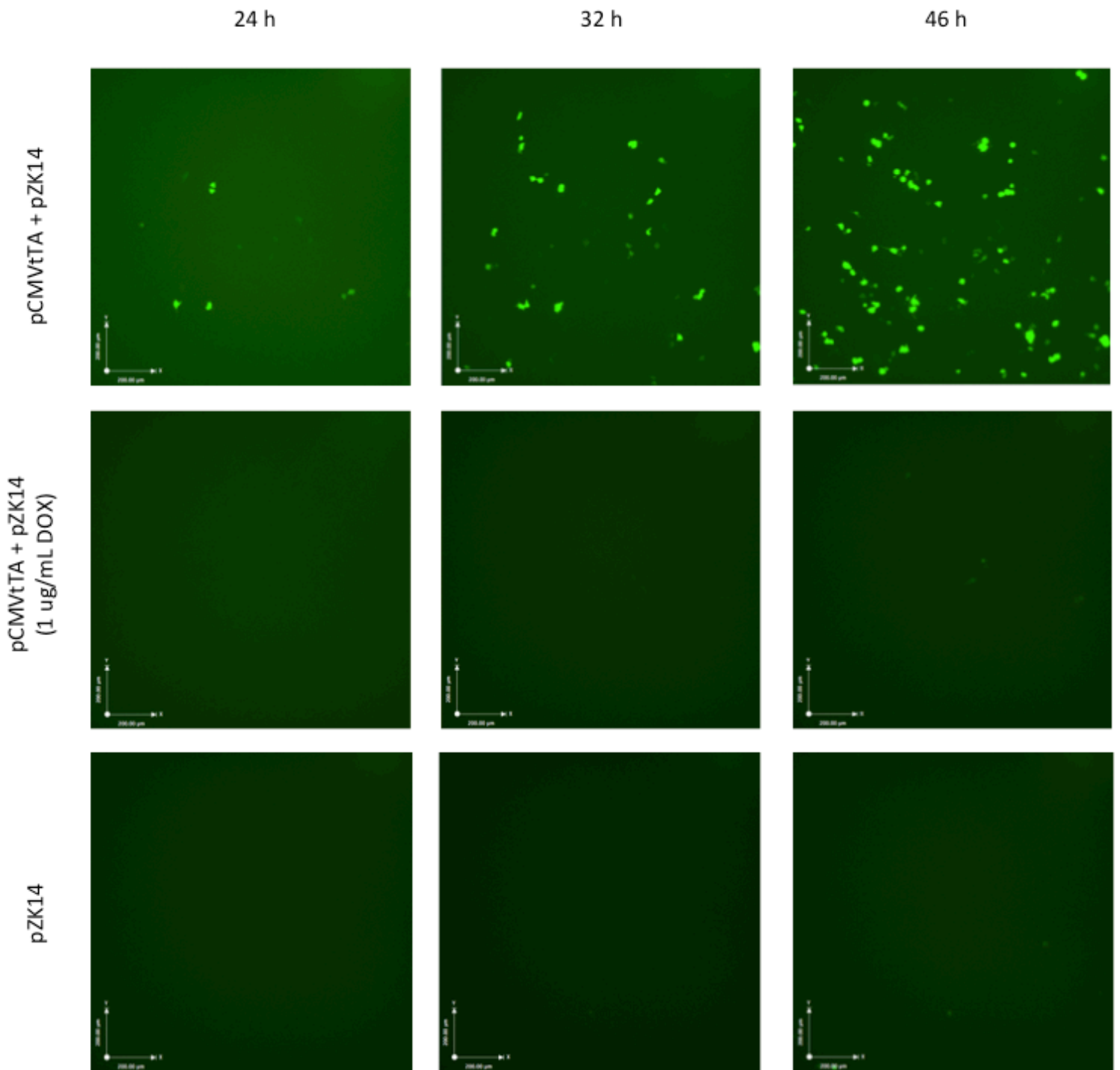


Figure 4.14. Fluorescence channel images of HEK cells transfected with tTA validation synthetic gene circuit components. White arrows on the bottom left of each image represent a 200 µm distance (Top) Cells co-transfected with pCMVtTA and pZK14 24, 32 and 46 hours after transfection. (Middle) Cells co-transfected with pCMVtTA and pZK14 24, 32 and 46 hours after transfection in DMEM medium supplemented with a final concentration of 1 µg/mL doxycycline. (Bottom) Cells transfected with pZK14 24, 32 and 46 hours after transfection.

weak progressive increase in fluorescence was observed in cells co-transfected with

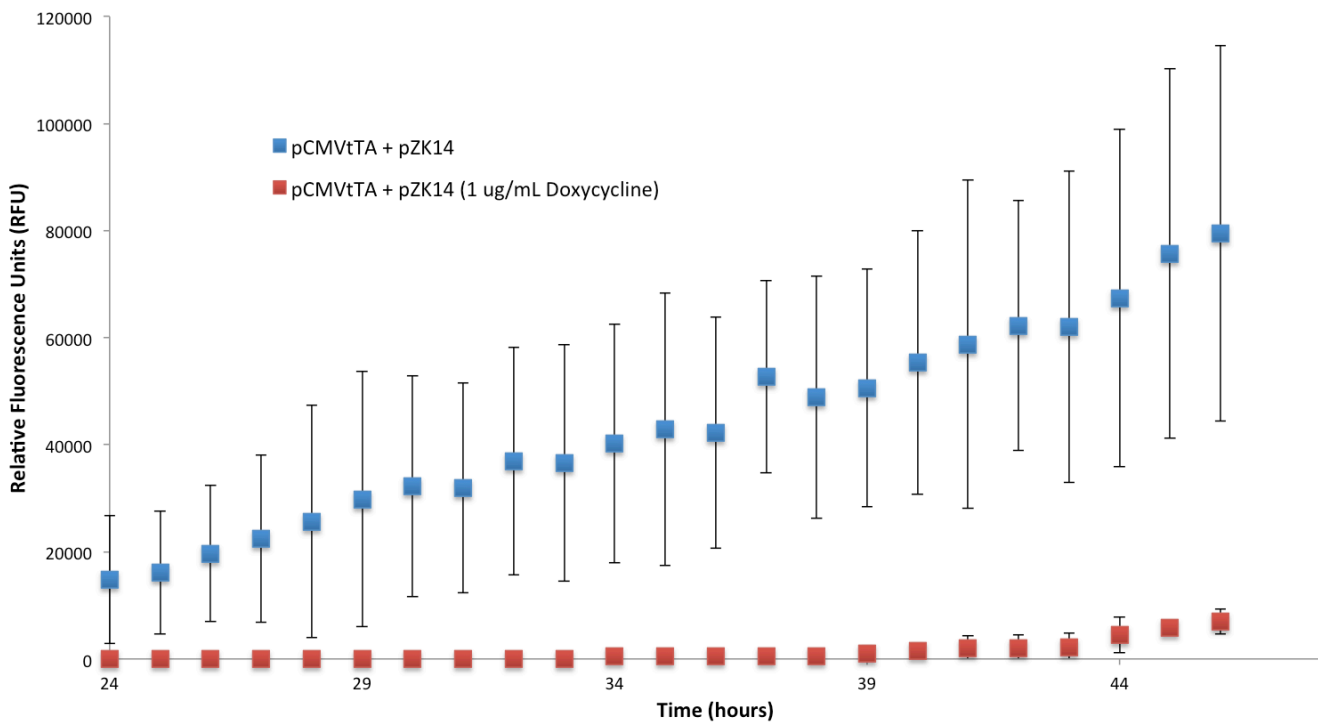


Figure 4.15. Fluorescence intensity time evolution of HEK293T cells co-transfected with pCMVtTA and pZK14, in DMEM medium supplemented with and without 1 ug/mL doxycycline. The data presented herein is representative of 126 fluorescent cells (pCMVtTA + pZK14) and 14 fluorescent cells (pCMVtTA+ pZK14, with 1 ug/mL doxycycline), across 4 replicate wells for each experimental condition. The error bars represent the standard deviation amongst replicates.

pCMVtTA and pZK14 in the presence of doxycycline (figure 4.15). This finding indicates that doxycycline had an inhibitory effect on tTA activity, as expected. Indeed, no fluorescence could be detected in the images acquired 24 and 32 hours after transfection.

Furthermore, no fluorescence was detected during the first 32 hours after transfection in cells transfected with pZK14. A weak fluorescence signal is shown in the image acquired 46 hours after transfection, thus illustrating a minimal level of P_{TRE3G} -driven leakiness.

The tTA- P_{TRE3G} system (i.e. Tet-Off™) is known to work in a variety of host cells and implementation strategies, ranging from prokaryotes to animal models (Aubrecht et al., 1996; Fussenegger et al., 2000; Gossen & Bujard, 1992; Gossen et al., 1995; Karlsson & Weber, 2012; Lutz & Bujard, 1997; May et al., 2008; Tigges et al., 2010; Yao et al., 1998). However, validating the *in vivo* biological function of the genetic organization constructed in pCMVtTA and pZK14 represents an important

step toward the implementation of the proposed mammalian synthetic gene oscillator. Indeed, the findings shown in figures 4.14 and 4.15 confirm that the *tTA* ORF consensus sequence and its upstream CMV promoter in pCMVtTA are capable of driving constitutive expression of the *tTA* gene, thus leading to accumulation of a functional tTA protein. Moreover, the extent of the doxycycline inhibitory effect on the transactivation potential of tTA, as well as the qualitative level of P_{TRE3G}-driven leakiness, were confirmed to be in line with previous mammalian cell implementations of the Tet-Off™ system (Aubrecht et al., 1996; Gossen et al., 1995; May et al., 2008).

4.4.4. *In vivo* implementation of the Gal4 validation synthetic gene circuit

The *in vivo* validation of both the *gal4* ORF and the UAS site structure proposed in 4.2 was carried out by transfecting the Gal4 validation synthetic gene circuit in HEK cells. To this end, cells were subject to two different experimental protocols: (i) co-transfection with pCMVGal4 and pUAS-GG, (ii) transfection with pUAS-GG. This experimental setup followed the same logic as that of the tTA validation synthetic gene circuit implementation: pCMVGal4 is expected to drive constitutive expression of the *gal4* gene through the activity of its upstream CMV promoter, thus leading to accumulation of Gal4 protein and subsequent transactivation of the *gfp* gene in pUAS-GG through binding of the five UAS sites. Thus, cells co-transfected with pCMVGal4 and pUAS-GG are expected to emit fluorescence due to the transactivation potential of Gal4 when bound to UAS sites; conversely, no fluorescence signal other than the basal expression from the UAS structure is expected to be observed in cells transfected solely with pUAS-GG.

The fluorescence images acquired during the *in vivo* implementation of this synthetic gene circuit are shown in figure 4.16. Images were acquired with a sampling frequency of 1 image / hour, for 24 hours; the first image was acquired 24 hours after transfection. Figure 4.16 shows the fluorescence intensity in HEK cells (24, 32 and 46 hours after transfection) resulting from (i) co-transfection with

pCMVGal4 and pUAS-GG, (ii) transfection with pUAS-GG. Like in the image acquisition protocols for the LacI and tTA validation synthetic gene circuits, images were also acquired using the brightfield channel; the superimposed brightfield-fluorescence images are shown in *Appendix A.4*.

A strong gradual increase in the number of fluorescent cells and overall fluorescence intensity was observed in cells co-transfected with pCMVGal4 and pUAS-GG, thus corroborating the expected results in this experimental condition

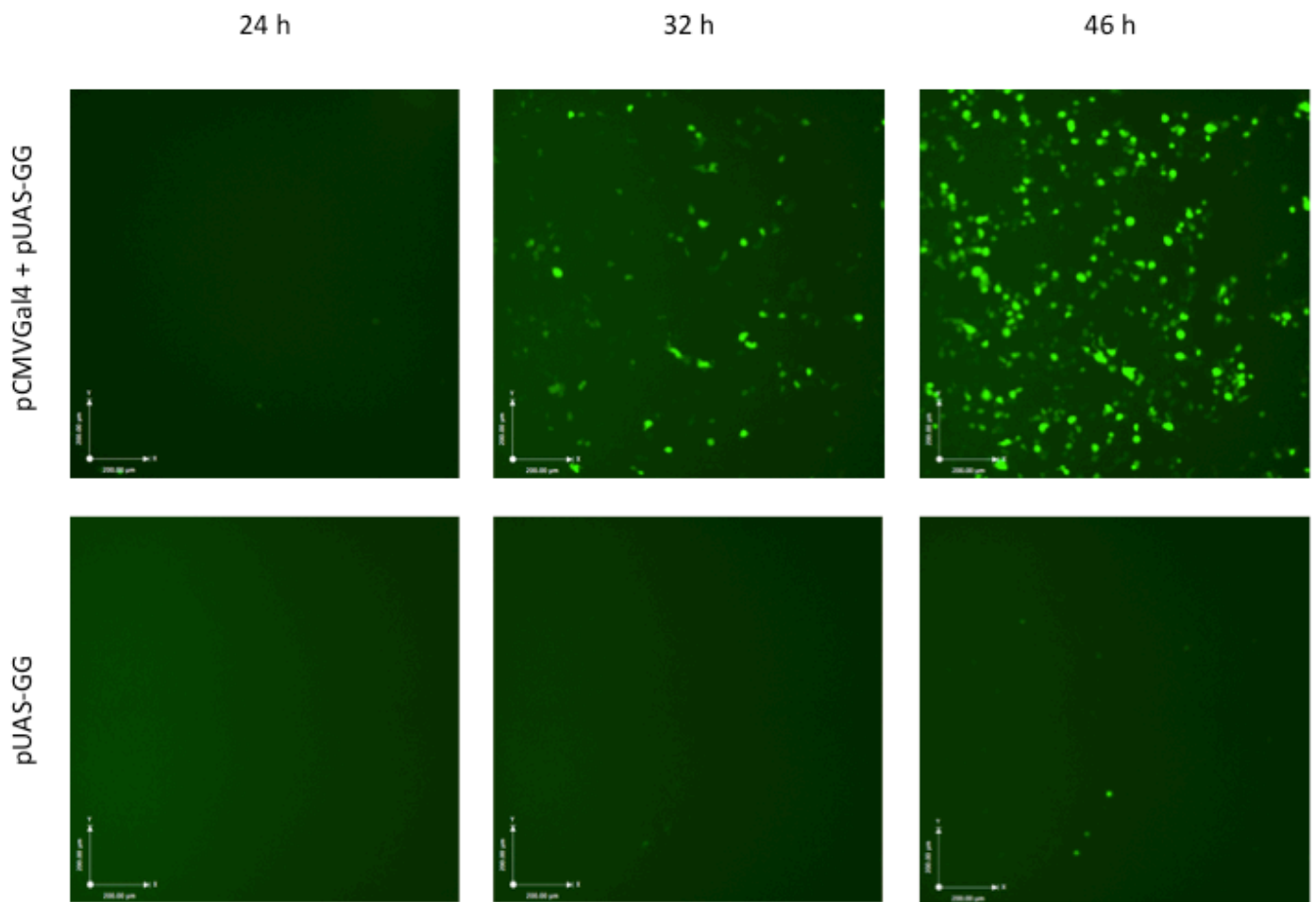


Figure 4.16. Fluorescence channel images of HEK cells transfected with Gal4 validation synthetic gene circuit components. White arrows on the bottom left of each image represent a 200 μm distance (Top) Cells co-transfected with pCMVGal4 and pUAS-GG 24, 32 and 46 hours after transfection. (Bottom) Cells transfected with pUAS-GG 24, 32 and 46 hours after transfection.

(figures 4.16 and 4.17). Moreover, Gal4-mediated up-regulation of the *gfp* gene in pUAS-GG was found to occur rapidly: the first fluorescence signals were detected immediately upon the beginning of the image acquisition protocol (i.e. 24 hours

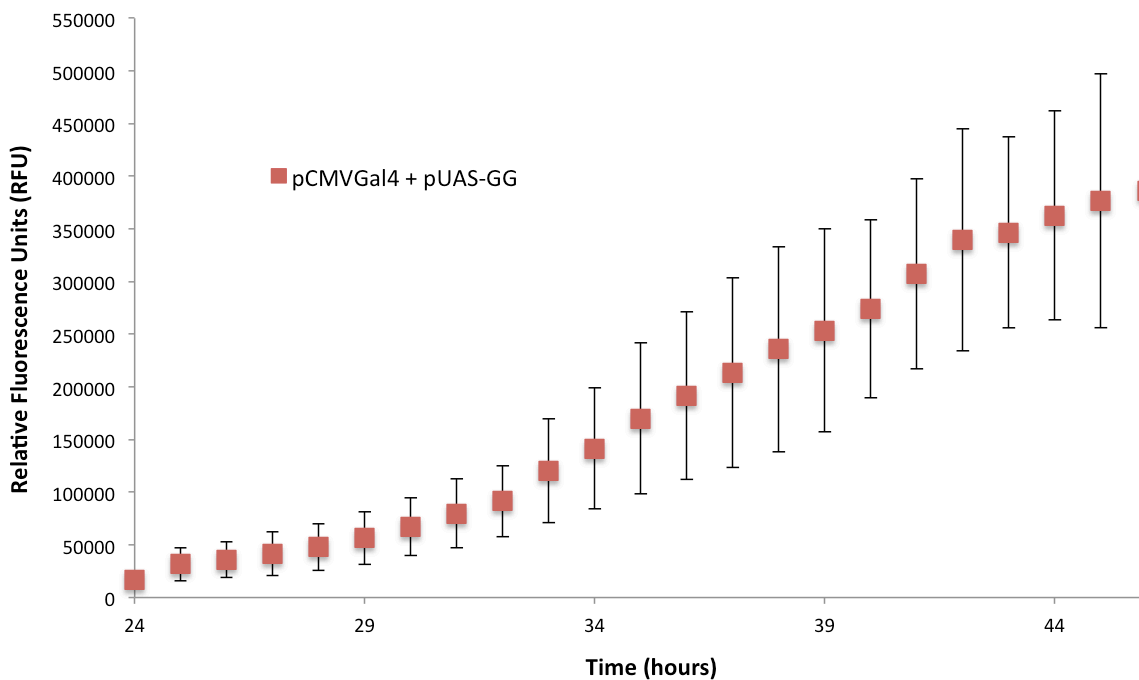


Figure 4.17. Fluorescence intensity time evolution of HEK293T cells co-transfected with pCMVGal4 and pUAS-GG. The data presented herein is representative of 940 fluorescent cells (across 4 replicate wells). The error bars represent the standard deviation amongst replicates.

after transfection). A weak, progressively increasing fluorescence signal was detected in cells transfected with pUAS-GG. Although being proportionally minute when compared to the overall fluorescence signal emitted by cells co-transfected with pCMVGal4 and pUAS-GG (i.e. < 1%), the fluorescence emitted by cells transfected solely with pUAS-GG indicates that the constructed UAS structure does present a basal expression level.

Like the Tet-Off™ system, the Gal4-UAS system has been widely implemented in a variety of host organisms, including mammalian cells (Carey et al., 1990; Kakidani & Ptashne, 1988; Phelps & Brand, 1998; Sadowski et al., 1988; Scheer & Campos-Ortega, 1999; Traven, Jelcic, & Sopta, 2006; Webster et al., 1988). Indeed, previous Gal4-UAS mammalian implementations have also reported both the strong transactivation potential of Gal4 and the existence of a basal expression level from UAS-based promoters (Kakidani & Ptashne, 1988; Webster et al., 1988), thereby corroborating the findings presented in figure 4.16.

Construction and in vivo validation of system parts

Similarly to the corroboration of tTA validation synthetic gene circuit parts, confirmation that the consensus *gal4* ORF sequence in pCMVGal4 and the UAS structure in pUAS-GG are functional in HEK cells represents an important validation step toward the goal of implementing the proposed mammalian synthetic gene oscillator.

4.5. Concluding remarks

The main objective of this chapter was to address and describe the processes leading to the construction and validation of the biological parts to be used in the proposed mammalian synthetic gene oscillator. Furthermore, it addressed *in vitro* and *in vivo* implementation feasibility guidelines that were followed when constructing and validating system parts.

A modular construction and implementation strategy was proposed with the objective of validating transcription factor / operator site pairs in an individual fashion, resorting to validation synthetic gene circuits made up of one constitutive expressing module and one inducible reporter module for each circuit, i.e. LacI, tTA and Gal4.

A comprehensive list of the methodologies used to assemble and implement these validation synthetic gene circuits was provided. The *in vitro* construction of the LacI, tTA and Gal4 validation circuits was based on the assembly of six plasmids: pCMVLacI, pOPRS_GFP, pCMVtTA, pZK14, pCMVGal4, and pUAS-GG. The size and structural integrity of these plasmids was validated through restriction enzyme digestion followed by agarose gel electrophoresis. The subsequent plasmid amplification and purification protocols were shown to be successful using a spectrophotometric method.

The *in vivo* implementation of the LacI, tTA and Gal4 validation synthetic gene circuits was performed in HEK cells. Gene expression activity for each of these circuits was monitored using fluorescence microscopy, which revealed the results were as expected: (i) LacI-mediated inhibition of *gfp* expression occurred when pCMVLacI was co-transfected with pOPRS_GFP – and interaction that could be inhibited by IPTG; (ii) tTA-mediated upregulation of *gfp* occurred when pCMVtTA was co-transfected with pZK14 – a behaviour that could be inhibited by doxycycline; (iii) Gal4 was successfully produced from pCMVGal4 and up-regulated expression of a *gfp* gene in pUAS-GG by binding UAS sites.

It is worth noting that the validation of individual system components could benefit from further experiments based on (i) re-engineering of regulatory

sequences up- and down-stream of component ORFs, and (ii) on the characterisation and optimisation of simultaneous transfection of multiple constructs. This is particularly relevant in the light of having a rigorous approach to the assembly of the whole system, i.e. as described in the next chapter:

- I. Specifically, additional *in vivo* part validation experiments with alternative promoter structures, operator site organizations (in terms of number, spacing, structure of operator sites, and distance to the Transcriptional Start Site (TSS)), and poly-adenylation signals would provide a highly valuable increase in biological design modularity, while allowing for a platform of continuous optimisation of module functionality. A recursive approach to module re-engineering and implementation is needed to accomplish this. Moreover, an approach based on a combination of experiments and model-based parameter estimation, with *in vivo* part implementation results as the input, would allow for a refinement of the kinetic parameters governing system dynamics. Taken in conjunction with the abovementioned additional part validation experiments with alternative regulatory elements, this approach allows for a more recursive and controllable system construction / implementation process.
- II. Moreover, the standardisation and characterisation of the transfection protocol with multiple plasmids can be made more rigorous if the intracellular concentration ratios of different plasmids are calculated. This can be achieved using different fluorescent-tagged versions for each construct. This approach can potentially provide valuable insight into alternative whole-system assembly strategies as it allows for a real-time assessment of relative plasmid concentration ratios during module implementation protocols. Additionally, this approach constitutes a simple framework for verifying the presence of each system component in target cells.

Construction and in vivo validation of system parts

The next chapter will address the construction and *in vitro* validation of the final system modules as well as the *in vivo* implementation of the proposed mammalian synthetic gene oscillator.

Chapter 5

In vivo implementation of a novel mammalian synthetic gene oscillator

5.1. Introduction

This chapter addresses the re-organization and coupling of the system parts presented in *Chapter 4*, as well as the *in vivo* implementation of the proposed mammalian synthetic gene oscillator.

The biological parts needed to implement the 3-component delayed negative feedback topology outlined in *Chapter 3* were shown to be functional in the HEK cell line. However, these system parts were constructed and implemented *in vivo* in an isolated manner, allowing for a modular validation of their respective biological functions. Thus, in order to assemble the proposed mammalian synthetic gene oscillator one must re-organize the validated system parts in a way that reflects the chosen topology (figure 3.1). Specifically, this task implies the replacement of CMV promoters in the *lacI*, *tTA*, and *gal4* constitutive expression modules by upstream regulatory elements that allow for a regulatory coupling framework while simultaneously reflecting the 3-component delayed negative feedback topology. In the case of the proposed mammalian synthetic gene oscillator, this can be accomplished by placing (i) UAS sites upstream of the validated *lacI* ORF, (ii) *lacO* operator sites upstream of the validated *tTA* ORF, and (iii) a P_{TRE3G} upstream of the validated *gal4* ORF. The construction of these system modules is presented and discussed in this chapter.

The dynamic behaviour of the system in different experimental conditions is then presented and discussed. The system is assembled in HEK cells by co-transfecting plasmids that encode the system parts. It is worth noting that this *in vivo* implementation also follows the implementation feasibility guidelines outlined in the beginning of *Chapter 4*.

The next section provides a comprehensive review of the re-organization of system modules. This is followed by a description of the experimental methodology used in the construction and implementation of the system. Lastly, the results regarding the *in vitro* system construction and its subsequent *in vivo* implementation are presented and discussed, followed by a conclusion.

5.2. Re-organization and coupling of system modules

The core structure of the proposed mammalian synthetic gene oscillator relies on the regulatory activity of LacI-, tTA-, and Gal4-responsive upstream elements. Indeed, these have been implemented *in vivo* as reporter regulators of validation synthetic gene circuits in *Chapter 4*: (i) five UAS sites in pUAS-GG, (ii) P_{TRE3G} in pZK14, and (iii) two lacO operator sites in pOPRS_GFP. The final structure of each system module comprises one of these elements placed upstream of their corresponding ORF (figure 4.1): UAS sites are placed upstream of the validated *lacI* ORF in order to allow for Gal4-mediated transactivation of the *lacI* gene; *lacO* operator sites are placed upstream of the validated *tTA* ORF in order to allow for LacI-mediated repression of the *tTA* gene; a P_{TRE3G} is placed upstream of the validated *gal4* ORF in order to allow for tTA-mediated upregulation of the *gal4* gene.

This regulatory structure reflects the 3-component delayed negative feedback presented in *Chapter 3* and is illustrated in figure 5.1. The RSV promoter upstream of the *lacO*-SV40 chimeric intron element allows for constitutive expression of the *tTA* gene, leading to synthesis of the tTA transactivator. Accumulation of the tTA protein leads to transactivation of the *gal4* gene, through binding of its upstream P_{TRE3G} promoter, consequently leading to synthesis of the Gal4 transactivator. Accumulation of the Gal4 protein leads to upregulation of the *lacI* gene through binding of its UAS sites. This leads to synthesis of the LacI repressor which will inhibit *tTA* expression and thus complete the delayed negative feedback. As validated in the previous chapter, the LacI-*lacO* and tTA-P_{TRE3G} interactions can be inhibited by IPTG and tetracycline / doxycycline, respectively.

Moreover, a reporter gene is required for real-time monitoring of system dynamics. As mentioned in *Chapter 2* and *Chapter 3*, a conventional, consensus *gfp* gene cannot be used in this instance due to the long half life of the encoded GFP protein (c. 25 hours) (Andersen et al., 1998; Corish & Tyler-Smith, 1999; Tigges et al., 2009). Thus, an ubiquitin-tagged *gfp* gene controlled by an upstream P_{TRE3G} promoter has been chosen as the reporter module. This GFP half-life reducing strategy (as well as the downstream addition of *ssra* tags for prokaryotic systems) has been used in previous implementations of synthetic gene oscillators, reducing GFP half-life to c. 2

In vivo implementation of a novel mammalian synthetic gene oscillator

hours (Michael B Elowitz & Leibler, 2000; Stricker et al., 2008; Tigges et al., 2010, 2009).

The next section addresses the methodologies used to construct and implement the system modules illustrated in figure 5.1. This is followed by a presentation and discussion of the *in vitro* system assembly and *in vivo* system implementation results.

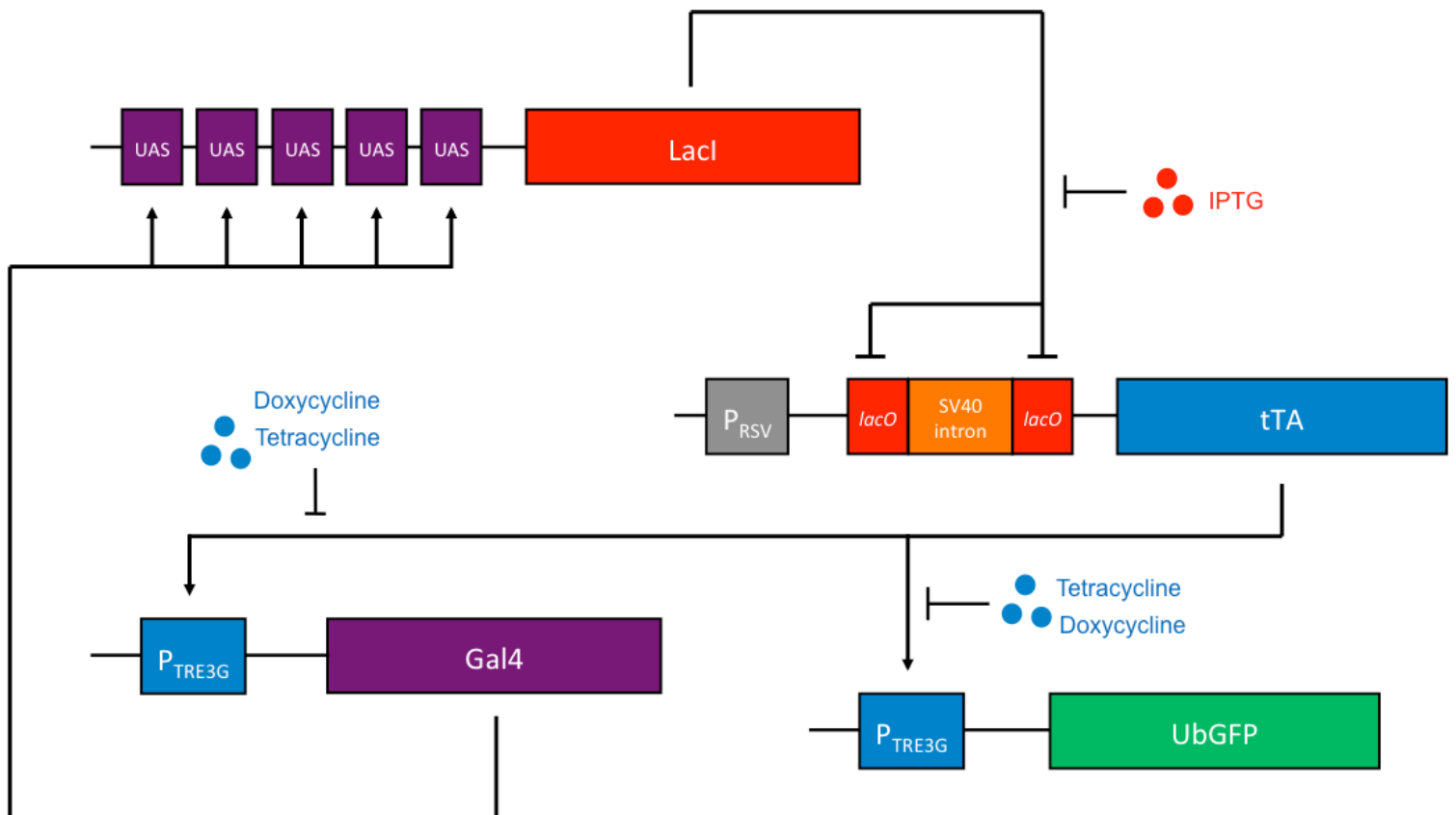


Figure 5.1. Diagrammatic representation of the proposed mammalian synthetic gene oscillator topology and biological organization. Constitutive expression of *tTA*, driven by an upstream RSV promoter, leads to accumulation of the tTA protein which upregulates *gal4* expression. Subsequent accumulation of the Gal4 protein leads to UAS-mediated upregulation of a *lacI* gene and synthesis of the LacI repressor, which completes the negative feedback loop by inhibiting *tTA* expression at *lacO* operator sites. A tTA-responsive ubiquitin-tagged *gfp* gene monitors system dynamics.

5.3. Methods

The methodologies concerning bacterial cell culture conditions, *in silico* plasmid design, bacterial transformations, plasmid preparation, DNA sequencing, endotoxin-free plasmid preparation, agarose gel electrophoresis, mammalian cell culture conditions, and image processing were all carried out according to the specifications provided in 4.3.

5.3.1. Plasmids for system *in vivo* implementation

The UAS / *lacI* ORF sequence was *de novo* synthesized by GeneArt® (Life Technologies) and subsequently cloned into the Multiple Cloning Site (MCS) of a pcDNA3.1(Neo) mammalian expression vector using the HindIII and Acc65I restriction sites (GeneArt® Sub-Cloning Services). The resulting plasmid (containing the *LacI* module) was conveniently named pUAS_*LacI*. The *tTA* module was constructed by cloning the *tTA* ORF in pCMVtTA into the MCS of a pOPRS mammalian expression vector, using the KpnI and NotI restriction sites (see *Chapter 4*) (Genscript Sub-Cloning Services). The resulting plasmid was conveniently named pOPRS_*tTA*. The *Gal4* module was constructed by cloning the *gal4* ORF in pCMVGal4 into pZK14 using the Sall and BamHI restriction sites, thus replacing the original *gfp* ORF in pZK14 by a *gal4* ORF (see *Chapter 4*) (Genscript Sub-Cloning Services). The resulting plasmid was named pTRE3G_*Gal4*.

The ubiquitin-tagged *gfp* reporter module was based on the pUb-G76V-GFP mammalian expression vector (kindly provided by Dr Kirsten Jensen, Imperial College London) (Tigges et al., 2009). The P_{TRE3G} promoter in pZK14 was cloned into pUb-G76V-GFP using the *Asel* and *NheI* restriction sites, i.e. resulting in the replacement of the original pUb-G76V-GFP CMV promoter with a P_{TRE3G} promoter (Genscript Sub-Cloning Services). The resulting plasmid was named pTRE3G_UbGFP.

The DNA sequences of pUAS_*LacI*, pOPRS_*tTA*, pTRE3G_*Gal4*, and pTRE3G_UbGFP are presented in *Appendix A.5*.

5.3.2. Turbofect™ transfection and system *in vivo* implementation protocol

HEK293T cells were transfected using the Turbofect™ Transfection reagent (Thermo Fisher Scientific). For this, cells were seeded 24 hours prior transfection in 48-well plates containing 500 μ L complete DMEM, at 40-50% confluency thus achieving 70-80% confluency at the time of transfection. 2 μ g of a pUAS_LacI/ pOPRS_tTA / pTRE3G_Gal4 / pTRE3G_UbGFP plasmid mixture was added to 4 μ L of transfection reagent and further diluted in 100 μ L of serum-free Opti-MEM® medium (Life Technologies). This solution was incubated at room temperature for 45 minutes and 100 μ L were added to each well, in a drop-wise manner. The 48-well plate was then gently rocked to ensure even distribution of DNA complexes and placed in a MCO-18AC, MCO-18AIC or MCO-18M Sanyo incubator at 37°C for 24 hours.

The tunability potential of the system was assessed in: (i) cells supplemented with 50 ng/mL, 100 ng/mL, and 150 ng/mL doxycycline; (ii) cells supplemented with 1 mM, 2 mM, and 3 mM IPTG. IPTG and doxycycline were supplemented to wells 1 hour before the time-lapse image acquisition protocol. The 48-well plates were then moved to the FILM suite at Imperial College London and were kept at 37°C, constant humidity and 5% CO₂ in an imaging chamber throughout the image acquisition protocol. Negative controls were created by co-transfecting HEK293T cells with pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pZK14.

5.3.3. Microscopy imaging, image processing and single cell analysis

Cells were visualized under a Leica DM IL LED-DFC295 inverted phase contrast microscope prior to the image acquisition protocol to confirm optimal confluency and that no contamination was present. Cells were then imaged using a Hamamatsu ORCA-ER camera coupled to a Zeiss Axiovert 200 inverted fluorescent microscope with a fully motorized stage, controlled by Improvion Volocity acquisition software. The image acquisition protocol was initiated by adjusting the brightfield Köhler

In vivo implementation of a novel mammalian synthetic gene oscillator

illumination and setting up the Zeiss EC Plan-Neofluar 10x 0.30 Ph1 objective (Phase Contrast 1 channel – Ph1) and Zeiss Filter Set 10 (FITC) with a mercury fluorescence lamp (Fluorescent channel 1). The motorized stage was then calibrated, and one randomly selected region of interest (ROI) was selected for each well. Images were then automatically acquired for 48 hours at a rate of 1 image / 15 minutes / well in the brightfield and fluorescent channels. Brightfield channel image acquisition was performed with an exposure time of 5 milliseconds (ms); fluorescent channel image acquisition was performed with an exposure time of 15 ms. Images were analyzed on a single-cell basis using the Fluorescence Measurement Thresholder in Volocity, thus allowing for automated detection of fluorescent regions and measurement of mean fluorescence intensity (in segmented fluorescent cells) in Relative Fluorescence Units (RFU).

5.4. Results and discussion

5.4.1. *In vitro* construction of system modules

The LacI, tTA, Gal4 and ubiquitin-tagged GFP system modules were constructed by modularly assembling four mammalian expression vectors: pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pTRE3G_UbGFP, respectively. Their corresponding plasmid maps are illustrated in figures 5.2 to 5.5.

The LacI system module is based on a consensus *lacI^q* ORF with an upstream Gal4-responsive promoter structure, i.e. six spaced UAS sites (pGene/V5, Life Technologies), and two downstream regulatory elements that ensure mRNA maturation and protein nuclear translocation: a SV40 nuclear localization signal, and a bGH poly-adenylation signal. The mammalian expression vector encoding this module, pUAS_LacI, was constructed by *de novo* synthesizing the UAS/*lacI^q* DNA sequence (GeneArt®) and subsequently cloning it in a pcDNA3.1(Neo) plasmid (Life Technologies), using the HindIII and Acc65I restriction sites (Genscript, Inc.).

The tTA system module is based on a *tTA* ORF with a previously validated LacI-responsive upstream regulatory structure, i.e. an RSV promoter with downstream *lacO* operator sites. A downstream poly-adenylation signal is also present in this module, allowing for mRNA maturation in HEK293T cells. The mammalian expression vector encoding this module, pOPRS_tTA, was constructed by inserting the *tTA* ORF in pCMVtTA (see *Chapter 4*) into a pOPRS plasmid (Agilent Technologies). This was carried out using the KpnI and NotI restriction sites (Genscript, Inc.).

The Gal4 system module is based on a previously validated *gal4* ORF placed downstream of a P_{TRE3G} promoter. Additionally, downstream SV40 poly-adenylation and nuclear localization signals ensure mRNA maturation and protein nuclear translocation. This module is encoded in pTRE3G_Gal4, a mammalian expression vector created by inserting the validated *gal4* ORF of pCMVGal4 into pZK14, using the Sall and BamHI restriction sites (Genscript, Inc.). Indeed, this restriction site pair

In vivo implementation of a novel mammalian synthetic gene oscillator

allowed for a removal of the original *gfp* ORF in pZK14 and its subsequent *in loco* replacement with the *gal4* ORF.

The construction of the reporter plasmid, pTRE3G_UbGFP, was carried out by replacing the upstream CMV promoter / enhancer complex in pUb-G76V-GFP by a P_{TRE3G} promoter. This was accomplished by digesting pUb-G76V-GFP with the AseI and NheI restriction enzymes and subsequently *in loco* inserting the pZK14 P_{TRE3G} promoter.

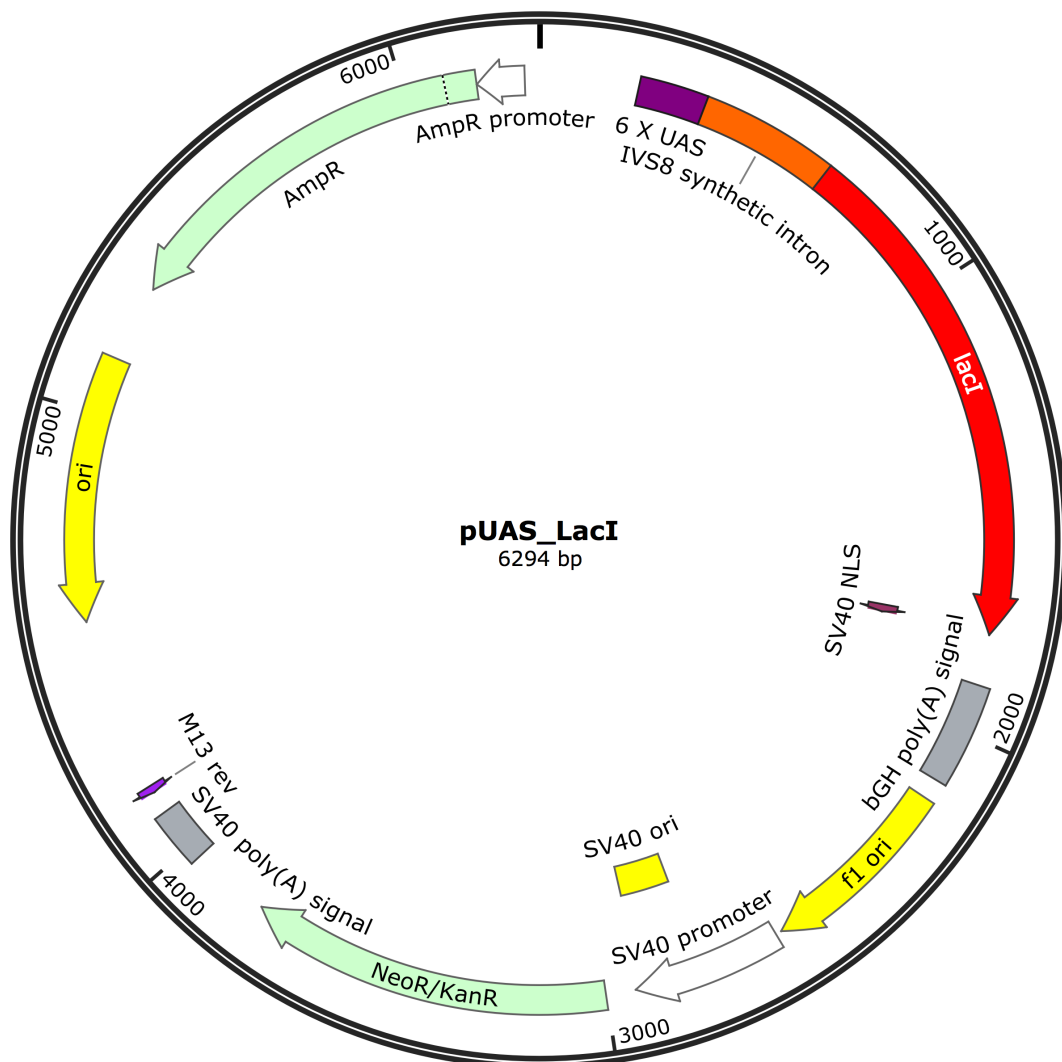


Figure 5.2. pUAS_LacI plasmid map. An upstream regulatory structure made up of six spaced UAS sites and an IVS8 synthetic intron regulates *lacI* expression in a Gal4-dependent manner. Downstream bGH poly-adenylation and SV40 nuclear localization signals ensure mRNA maturation and LacI nuclear translocation.

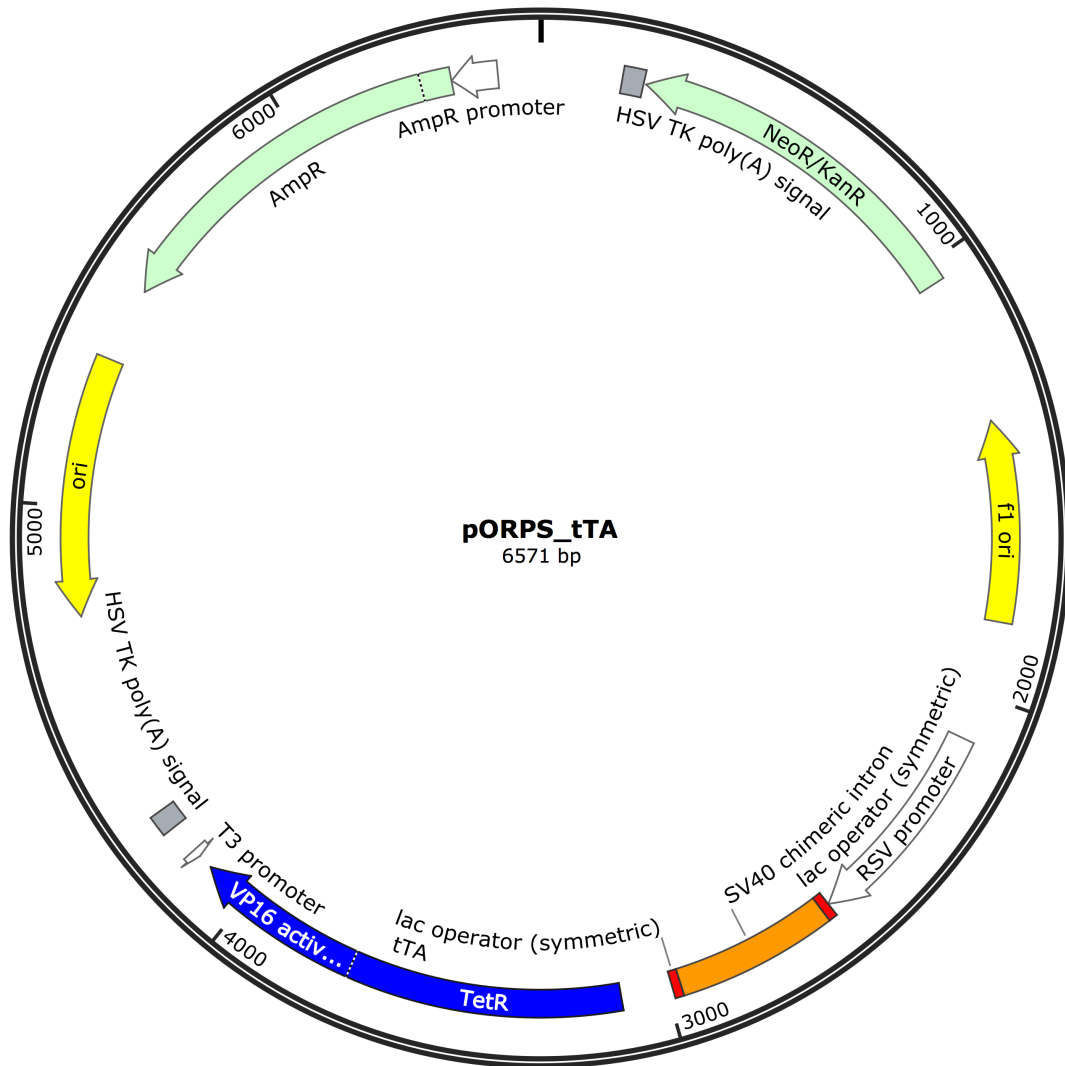


Figure 5.3. pOPRS_tTA plasmid map. An upstream regulatory structure made up of two *lacO* operator sites flanking a SV40 chimeric intron and an RSV promoter ensure LacI-responsive regulation and constitutive expression in the absence of LacI. A downstream HSV poly-adenylation signal ensures mRNA maturation.

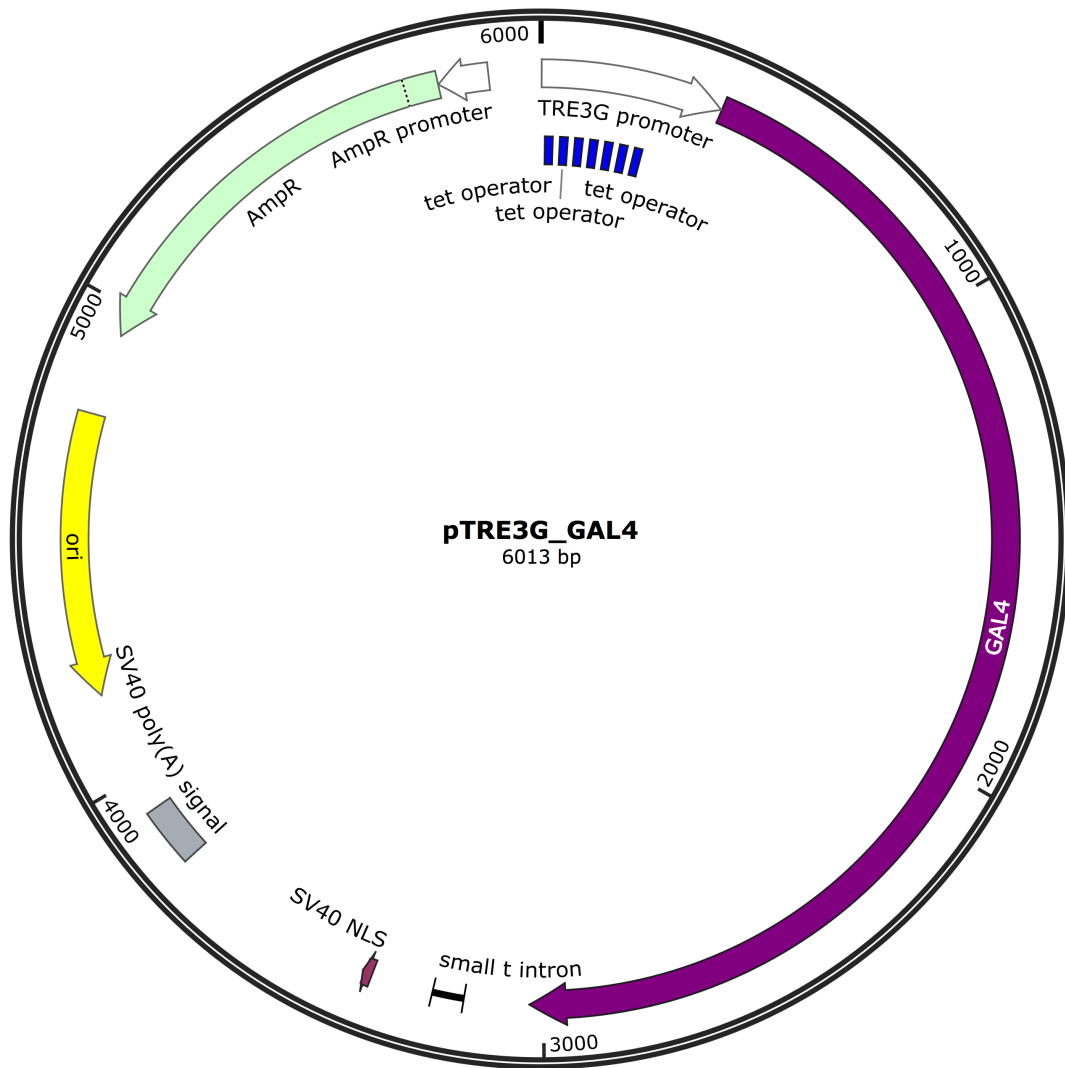


Figure 5.4. pTRE3G_Gal4 plasmid map. A TRE3G promoter made up of seven tetracycline responsive elements upstream of the *gfp* ORF regulates expression of *gal4* in a tTA-responsive manner. Downstream SV40 poly-adenylation and nuclear localization signals ensure mRNA maturation and Gal4 nuclear translocation.

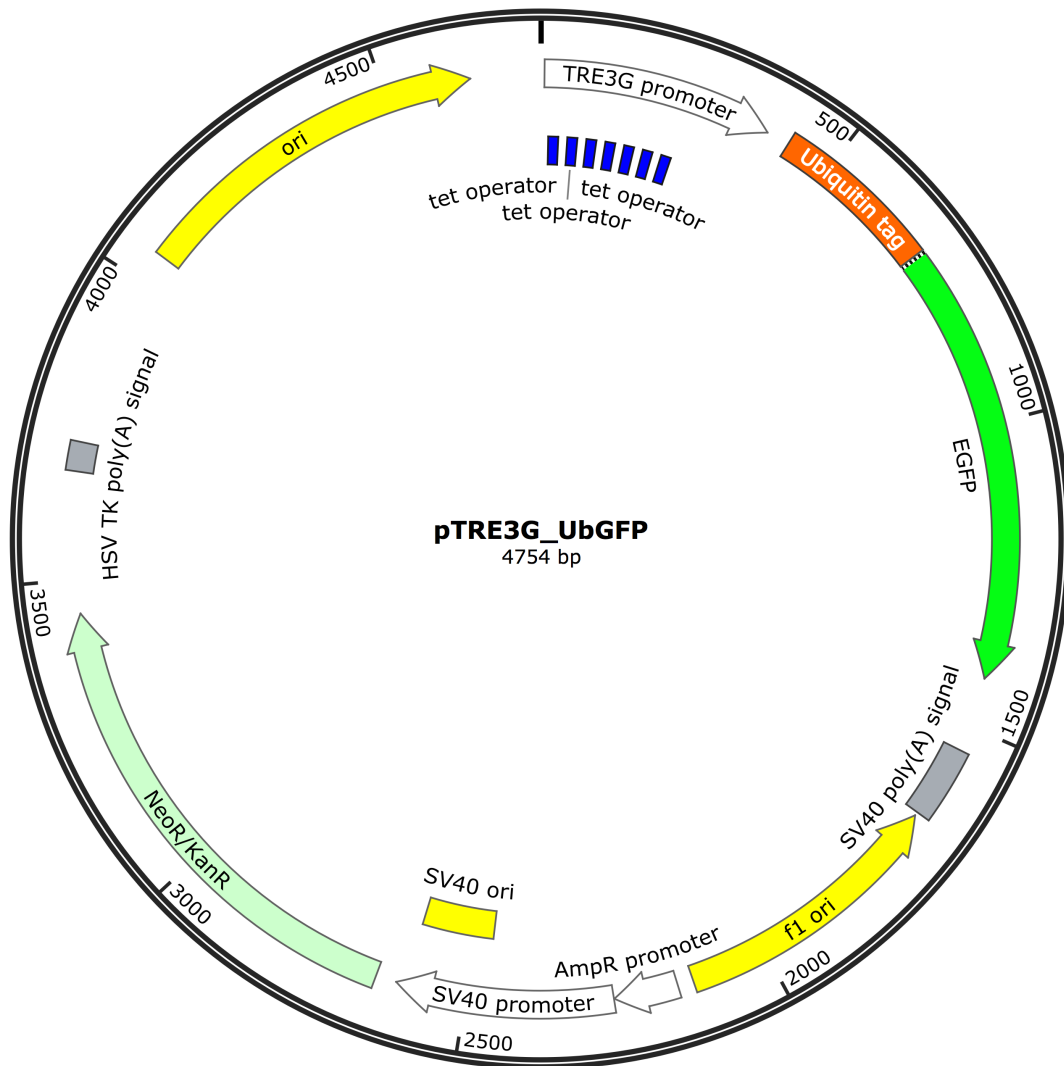


Figure 5.5. pTRE3G_UbGFP plasmid map. A TRE3G promoter made up of seven tetracycline responsive elements upstream of the *ubgfp* ORF regulates expression of GFP in a tTA-responsive manner. A downstream SV40 poly-adenylation signal ensures mRNA maturation.

In vivo implementation of a novel mammalian synthetic gene oscillator

The *in vitro* assembly of pUAS_LacI, pTRE3G_Gal4, pOPRS_tTA, and pTRE3G_UbGFP was validated through restriction enzyme digestion followed by agarose gel electrophoresis. To this end, these four plasmids were individually restriction enzyme digested with BamHI, EcoRI, NheI, and BamHI, respectively, and subsequently ran on an agarose gel according to the protocol specified in 4.3.7. The corresponding results corroborated the expected plasmid sizes (6294 bp, 6571 bp, 6013 bp, 4754 bp) and are respectively presented in figures 5.6 to 5.9. Due to the approximate similarity in plasmid sizes, the bands corresponding to pUAS_LacI and pTRE3G_Gal4 are shown in individual, separate agarose gels (figures 5.6 and 5.7). Moreover, the bands corresponding to pOPRS_tTA and pTRE3G_UbGFP are shown in conjunction with plasmids presented in *Chapter 4*, thus allowing for a comprehensive plasmid size comparison.

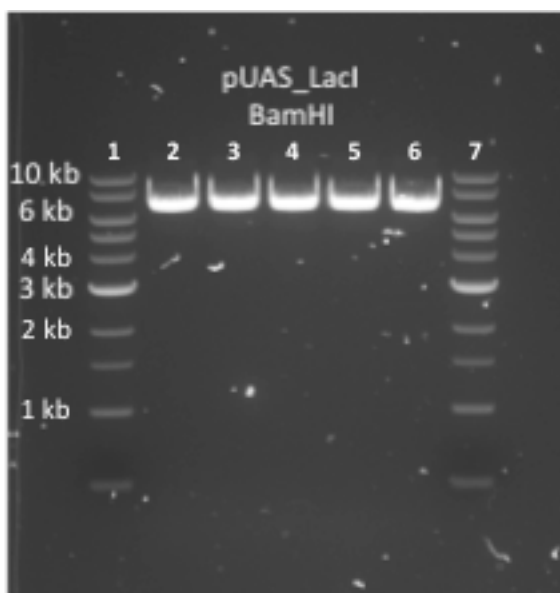


Figure 5.6. Agarose gel electrophoresis of pUAS_LacI (lanes 2 to 6) after BamHI restriction digestion. Clear bands can be observed over the 6 kb mark, thus corroborating the expected plasmid size (6294 bp). The molecular weight indicator is located in lanes 1 and 7.

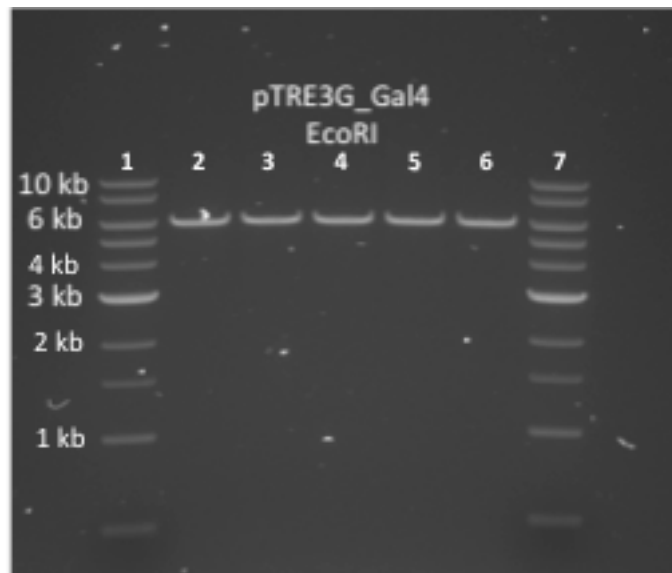


Figure 5.7. Agarose gel electrophoresis of pTRE3G_Gal4 (lanes 2 to 6) after EcoRI restriction digestion. Clear bands can be observed around the 6 kb mark, thus corroborating the expected plasmid size (6013 bp). The molecular weight indicator is located in lanes 1 and 7.

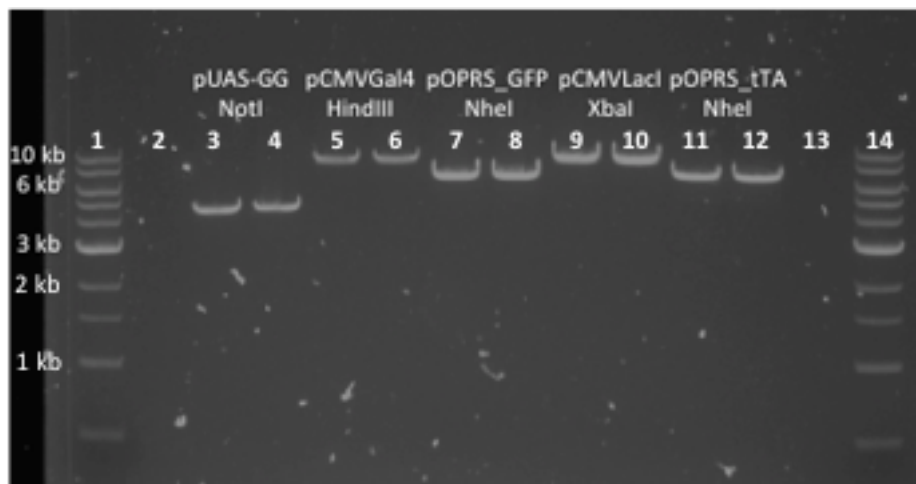


Figure 5.8. Agarose gel electrophoresis of pOPRS_tTA after NheI restriction digestion (lanes 11 and 12), shown in conjunction with pUAS-GG (lanes 3 and 4), pCMVGal4 (lanes 5 and 6), pOPRS_GFP (lanes 7 and 8), and pCMVLacI (lanes 9 and 10). Clear bands corresponding to linearized pOPRS_tTA can be found slightly above the 6 kb mark, corroborating the expected plasmid size (6571 bp). The molecular weight indicator (i.e. 1 kb DNA Ladder) is located in lanes 1 and 14.



Figure 5.9. Agarose gel electrophoresis of pTRE3G_UbGFP after BamHI restriction digestion (lanes 2 to 6), shown in conjunction with pZK14 (lanes 7 to 11). Clear bands corresponding to linearized pTRE3G_UbGFP can be found between the 4 kb and 5 kb marks, thus corroborating the expected plasmid size (4754 bp). The molecular weight indicator (i.e. 1 kb DNA Ladder) is located in lanes 1 and 12.

Like the plasmids presented in *Chapter 4*, the four plasmids encoding the system components (i.e. pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4, and pTRE3G_UbGFP) were further amplified in *E.coli* and purified in an endotoxin-free manner before their transfection in HEK293T cells. The obtained concentration and purity (using a Nanodrop™ 2000c Spectrophotometer) of each of these plasmids is outlined in table 5.1.

All plasmids showed an acceptable concentration in the context of mammalian cell transfection. Indeed, the lowest concentration obtained was that of pTRE3G_UbGFP (1.7 $\mu\text{g}/\mu\text{L}$). Moreover, the obtained absorbance ratios at 260 nm / 280 nm and 260 / 230 nm were all in the range of 1.8-1.9 and 2.0-2.12, thus indicating that both the amplification and purification protocols were successful.

Table 5.1. Concentration and purity of plasmids encoding the finalized system parts, i.e. pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4, and pTRE3G_UbGFP.

Plasmid	Concentration ($\mu\text{g}/\mu\text{L}$)	A260 / A280	A260/A230
pUAS_LacI	1.34	1.86	2.1
pOPRS_tTA	2.1	1.89	2.07
pTRE3G_Gal4	2.4	1.84	2.12
pTRE3G_UbGFP	1.7	1.89	2.09

5.4.2. *In vivo* implementation of the proposed mammalian synthetic gene oscillator

The proposed mammalian synthetic gene oscillator was implemented *in vivo* through co-transfection of HEK293T cells with pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pTRE3G_UbGFP. Like in previous synthetic gene oscillator implementations, the data presented herein relies on single-cell image analysis (Atkinson et al., 2003; M B Elowitz & Leibler, 2000; Stricker et al., 2008; Tigges et al., 2010, 2009).

Unstable transient behaviour with an initial peak showing an average period of c. 6 hours (± 2.81) were observed in HEK293T cells co-transfected with the system-encoding plasmids at equimolar ratios. Indeed, for the vast majority of cells exhibiting this behaviour (i.e. > 95%), one fluorescence peak was observed followed by stable dynamics for the remainder of the image acquisition protocol. A minority of cells (i.e. < 1%) exhibited two peaks (data not shown), which were also followed by stable dynamics throughout the remainder of the image acquisition protocol. Figure 5.10 illustrates the fluorescence signal observed in three cells implemented with the proposed synthetic gene oscillator system. Data concerning fluorescence

In vivo implementation of a novel mammalian synthetic gene oscillator

signals from additional single cells is available upon request, as well as the corresponding time-lapse videos.

The three fluorescence signals presented in figure 5.10 represent the lowest (*Cell 2*), average (*Cell 1*), and highest (*Cell 3*) recorded peak amplitudes, and were chosen for this illustration in order to provide a comprehensive outline of the variability observed in the experimental cell population. Indeed, fluorescence signals were found to be asynchronous and irregular across the cellular population. The fluorescence signals illustrated in figure 5.10 were thus plotted in the same time interval. Moreover, the fluorescence channel images corresponding to the transient behaviour represented in figure 5.10 are shown in figure 5.11.

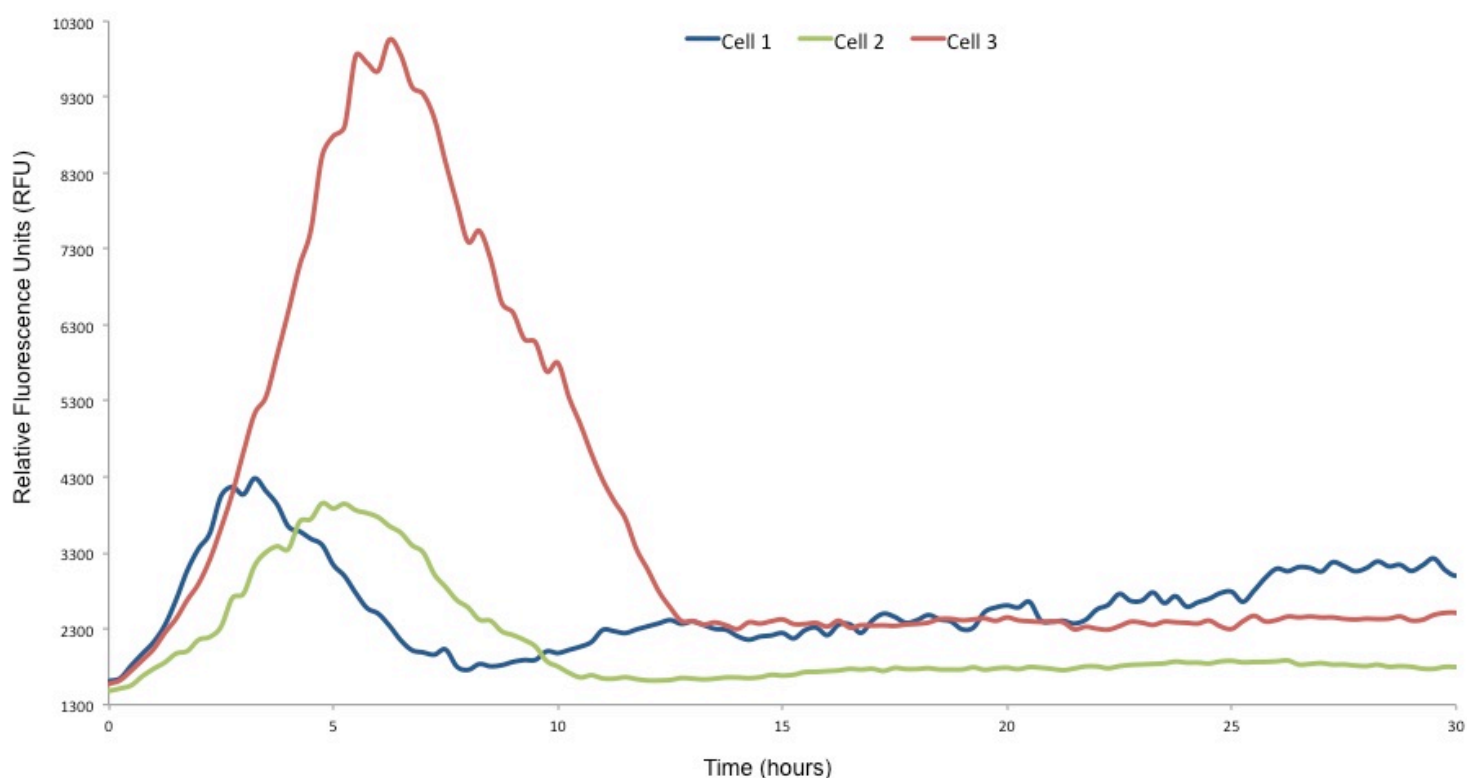


Figure 5.10. Single cell analysis of three cells exhibiting unstable transient behaviour after co-transfection with equimolar ratios of pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pTRE3G_UbGFP. The fluorescence channel images corresponding to the behaviour observed in each of these cells are presented in figure 5.11.

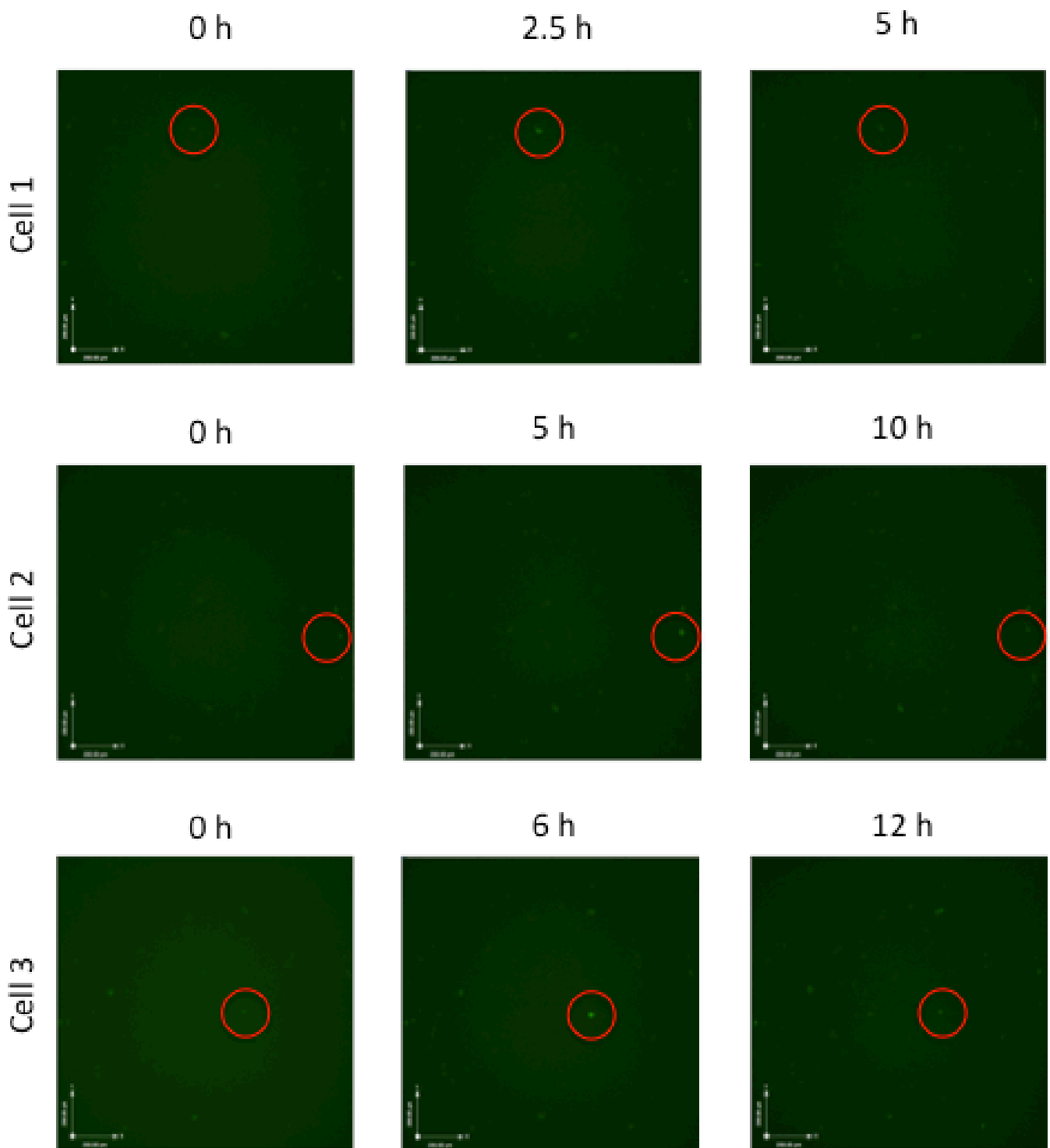


Figure 5.11. Fluorescence channel images corresponding to the single-cell signals illustrated in figure 5.10. Red circles highlight the respective cells in each image. The time frames indicated above each picture are in accordance with the x-axis in figure 5.10.

In vivo implementation of a novel mammalian synthetic gene oscillator

The variability in period and amplitude, as well as the overall irregularity of the fluorescence peaks observed in the *in vivo* implementation of the proposed synthetic gene oscillator is in line with previous implementations reported in the literature, and suggests no functional coupling between cells showing this transient behaviour (Atkinson et al., 2003; M B Elowitz & Leibler, 2000; O'Brien, Itallie, & Bennett, 2012; Purcell, Savery, Grierson, & di Bernardo, 2010; Tigges et al., 2010, 2009). Moreover, only a minority of fluorescent cells exhibited this unstable transient behaviour: the robustness of this system implementation was c. 2.34%. It is hereby hypothesized that this low robustness (as well as the absence of oscillatory behaviour) is related to the modular implementation strategy carried out: system components were encoded in four plasmids thus reducing the probability that one cell would incorporate all components. As a comparative exercise, the original Fussenegger oscillator was comprised of two plasmids containing the core components, and an additional reporter plasmid; it presented a robustness lower than 20% (Purcell et al., 2010; Tigges et al., 2009). In the case of the proposed mammalian synthetic gene oscillator, the core components are encoded in three plasmids, and an additional reporter plasmid is also required, thereby increasing the combinatorial complexity inherent to system uptake by cells (Kis et al., 2015).

Furthermore, the unexpected detection of a transient unstable behaviour instead of sustained oscillatory behaviour, as well as the disagreement in the predicted *versus* observed period in fluorescence peaks, can be explained by a conjugation of two perspectives, one related to specific biological features of the system, the other to design specifications:

- The dynamical properties of synthetic gene oscillators, and synthetic gene networks in general, are known to be highly dependent on the host organism (Karlsson & Weber, 2012; Kis et al., 2015; Weber & Fussenegger, 2009). In particular, the dynamics of mammalian synthetic gene networks have been shown to have great variability across different cell line implementations (Ausländer et al., 2012; Greber & Fussenegger, 2007; Weber & Fussenegger, 2009, 2010). The proposed mammalian synthetic gene oscillator was implemented in the HEK293T cell line and showed an unstable transient

In vivo implementation of a novel mammalian synthetic gene oscillator

behaviour instead of the model-predicted sustained oscillatory behaviour. However, the cell line-dependent variability in system dynamics observed in past mammalian synthetic gene network implementations is indicative that the proposed mammalian synthetic gene oscillator can potentially produce alternative dynamic behaviours, such as sustained oscillations or a stable steady state, if implemented in other mammalian cell lines, e.g. Chinese Hamster Ovary (CHO) cells. Furthermore, although the *in vivo* implementation presented herein was based on equimolar co-transfection with pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4, and pTRE3G_UbGFP, there are currently no methodologies to ensure equimolar plasmid uptake by mammalian host cells, unlike in prokaryotic systems. This uncontrollability of molarity ratios during plasmid uptake suggests that cells showing an unstable transient behaviour have not internalized the system components in equal concentrations. As outlined in the gene dosage bifurcation analysis results (*Chapter 3*), plasmid ratios can (i) theoretically approximate, create and/or abolish oscillatory behaviour, and (ii) alter the period of oscillations, once the system is within a Hopf region. For instance, a *lacI* relative gene dosage lower than 0.379 has been shown to theoretically abolish sustained oscillations, potentially justifying the unexpected observation of an unstable transient behaviour, in the context of plasmid uptake uncontrollability. This feature can be confirmed by simulating the model of the proposed system with different values of parameter g_1 (illustrated in figure 5.12). This argument can also be used to explain the low robustness, in comparison to the results reported by Stricker *et al.* (2008) and Tigges *et al.* (2010), of this *in vivo* implementation: it is possible that cells showing stable fluorescence signals throughout the experimental protocol have internalized system components at molarity ratios that do not dynamically permit oscillatory behaviour;

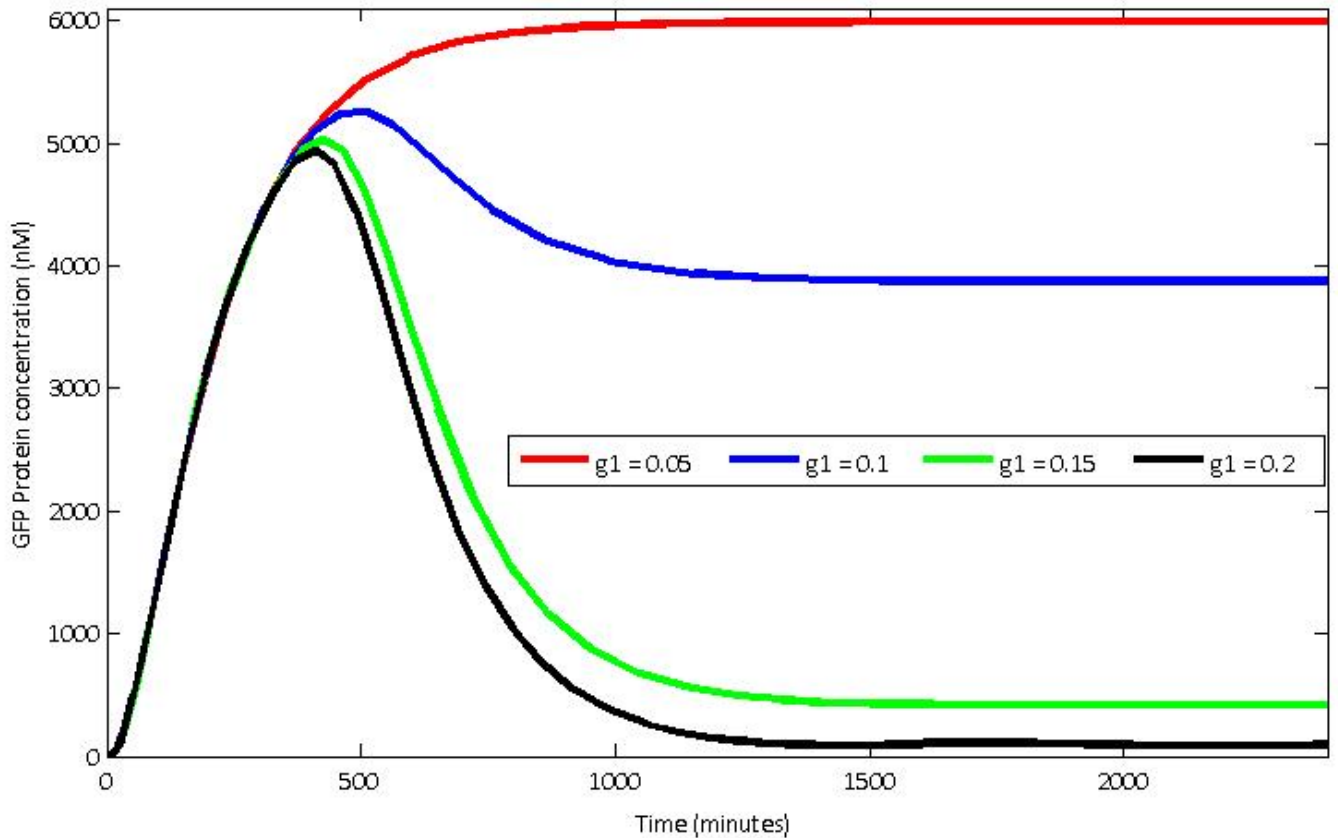


Figure 5.12. Damped GFP oscillatory signal obtained by model simulation assuming 0.05:1:1, 0.1:1:1, 0.15:1:1, and 0.2:1:1 molarity ratios between pUAS_LacI, pOPRS_tTA, and pTRE3G_Gal4, respectively.

- The proposed system topology is known to generate sustained oscillations (Novák & Tyson, 2008; Purcell et al., 2010; Tigges et al., 2009). However, the modelling framework (as well as dynamic mathematical models of previous synthetic gene oscillator implementations) used to theoretically predict system dynamics does not take into account several aspects that contribute to its *in vivo* biological function. Specifically, the distances of promoters and operator elements to their corresponding transcription start sites (TSS), the variability of cell cycle dynamics across cellular populations, the mRNA / protein nuclear translocation processes, and the diffusion of system components throughout the cell were not considered in the design of the system. Moreover, the kinetic parameters governing the dynamics of system processes are subject to an uncertainty that is widespread across the field of mammalian synthetic biology (Kis et al., 2015; Novák & Tyson, 2008). This

In vivo implementation of a novel mammalian synthetic gene oscillator

incomplete capture of the biological context in which the system is implemented has contributed to a theoretical prediction of sustained oscillatory behaviour, which does not fully reflect the biological organization of the system nor the dynamics observed in the experimental results.

Moreover, the differences in relative promoter strengths across the Gal4, tTA, and LacI modules reported in 4.4.2 – 4.4.4 (as a measure of GFP fluorescence intensity) were considered as a potential cause for the absence of oscillatory behaviour in cells co-transfected with pOPRS_tTA, pUAS_LacI, pTRE3G_Gal4, and pTRE3G_UbGFP. To test this, the dynamic model of the proposed synthetic gene oscillator was simulated using the proportionately (i.e. relative fluorescence intensities at 46 hours after transfection) updated relative TRE3G, UAS, and RSV_{lacO}

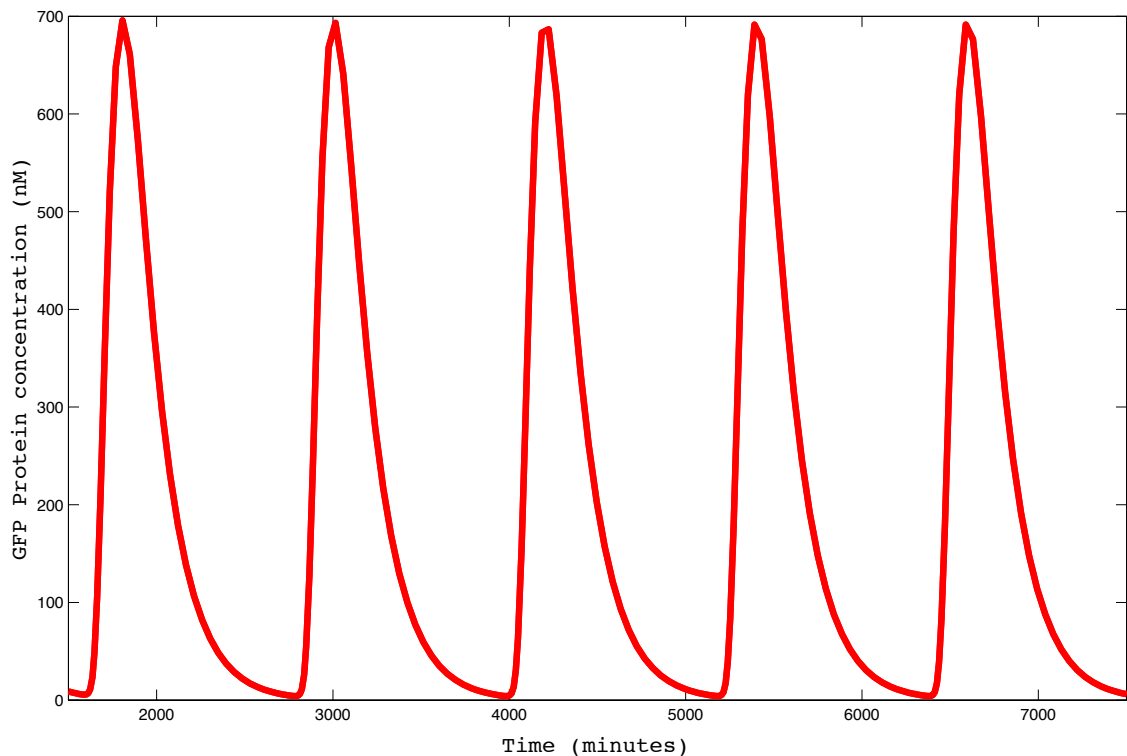


Figure 5.13. Sustained oscillatory signal obtained by simulating the dynamic model of the proposed synthetic gene oscillator with updated maximal transcription rates: 22.8 min^{-1} , 30 min^{-1} , and 153 min^{-1} for tTA, Gal4 and LacI, respectively.

strengths as the Gal4, LacI, and tTA maximal transcription rates, respectively (figure 5.13).

Indeed, figure 5.13 clearly shows that the model of the proposed system is capable of generating sustained oscillatory behaviour when the updated maximal transcription rates for LacI, tTA and Gal4 are assumed. This contradicts the hypothesis that unbalanced promoter strengths (reported in 4.4.2 - 4.4.4) across system modules lead to abolishment of oscillatory behaviour.

Although a transient behaviour was observed instead of oscillatory behaviour, the results concerning the *in vivo* system implementation highlight that the proposed system is capable of generating unstable dynamics. This represents one of the main requirements in the construction of a synthetic gene oscillator.

5.4.3. Exploring system tunability and instability with chemical inhibitors

The next logical step in the context of *in vivo* system implementation was to test how the chemical inhibitors doxycycline and IPTG affected system dynamics. To this end, HEK293T cells co-transfected with pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4, and pTRE3G_UbGFP at equimolar ratios were supplied with concentration gradients of these two chemical inhibitors according to the methodology specified in 5.2.2.

Doxycycline was found to lead to stable dynamics (data not shown). However, HEK293T cells co-transfected with pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pTRE3G_UbGFP, supplied with 1 mM, 2mM, and 3mM IPTG exhibited a transient unstable behaviour. All cells showing this behaviour exhibited one fluorescence peak, followed by stable dynamics. The fluorescence signal time evolutions observed in cells showing this behaviour are shown in figures 5.14 (for cells supplied with 1 mM IPTG), 5.15 (for cells supplied with 2 mM IPTG), and 5.16 (for cells supplied with 3 mM IPTG). Data concerning fluorescence signals from additional single cells is available upon request, as well as the corresponding time-lapse videos.

In vivo implementation of a novel mammalian synthetic gene oscillator

Like in the single cell analysis presented in the previous section, three fluorescence signals representing the highest, average, and lowest recorded peak amplitudes (*Cell 1*, *Cell 2*, *Cell 3*) were chosen for the illustrations in figures 5.14 to 5.16. Moreover, these fluorescence signals were plotted in the same time interval due to the asynchronous and irregular nature of fluorescence signals across the cellular population. The fluorescence channel images corresponding to the behaviour represented in figures 5.14 to 5.16 are shown in *Appendix A.5*.

An unstable transient behaviour exhibiting peaks with average periods of *c.* 5.35 (± 2.12), 3.59 (± 2.44), and 3.39 (± 1.76) hours were detected for cells supplemented with 1 mM, 2 mM, and 3 mM IPTG, respectively. Fluorescence signals were irregular across cellular populations and suggested no functional coupling between cells: different peak amplitudes and periods were recorded within the observed fluorescence signals of each experimental condition, in line with the results presented in the previous section.

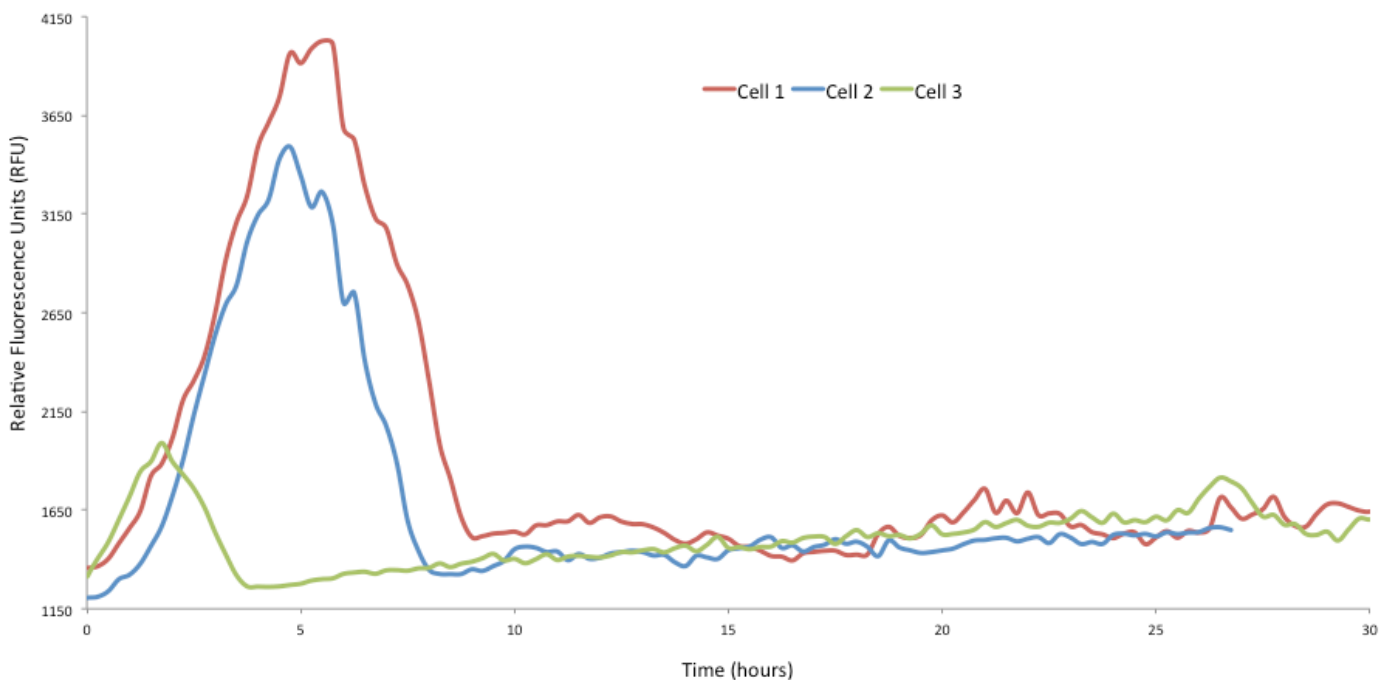


Figure 5.14. Single cell analysis of three cells exhibiting an unstable transient behaviour after co-transfection with equimolar ratios of pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pTRE3G_UbGFP, in complete DMEM medium supplied with 1 mM IPTG. The fluorescence channel images corresponding to the fluorescence peak

In vivo implementation of a novel mammalian synthetic gene oscillator

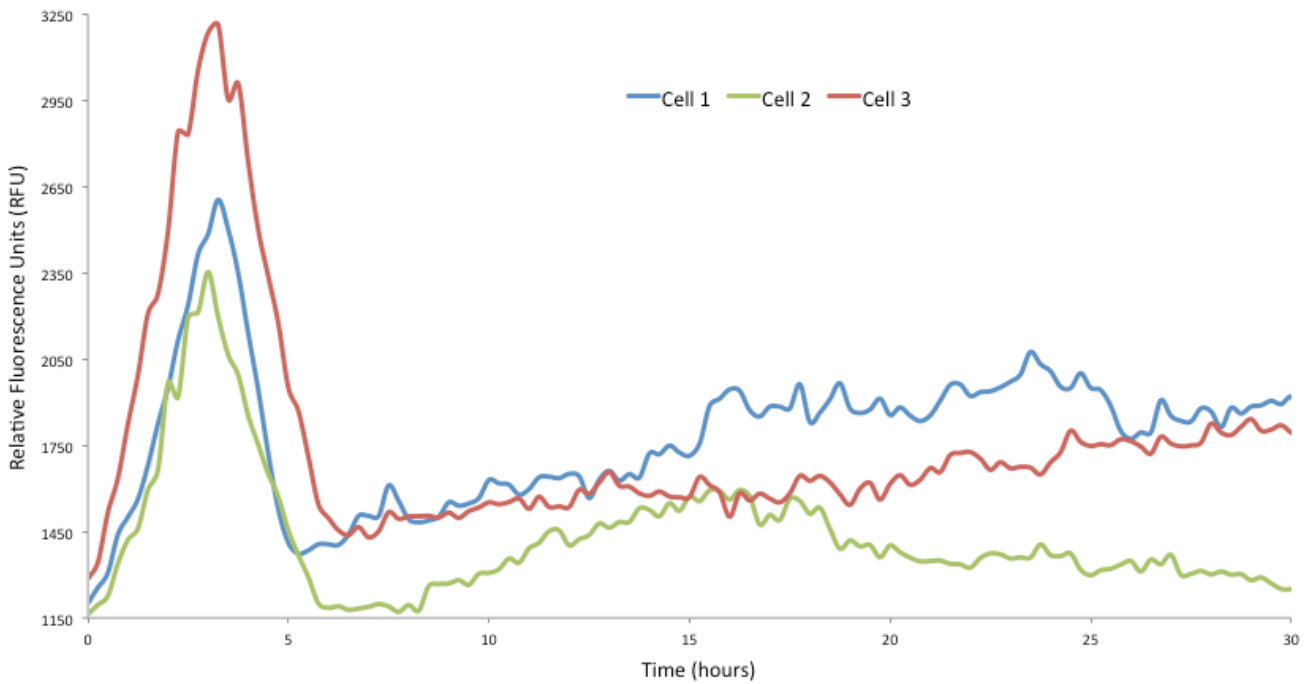


Figure 5.15. Single cell analysis of three cells exhibiting an unstable transient behaviour after co-transfection with equimolar ratios of pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pTRE3G_UbGFP, in complete DMEM medium supplied with 2 mM IPTG. The fluorescence channel images corresponding to the fluorescence peak observed in each of these cells are presented in *Appendix A.5*.

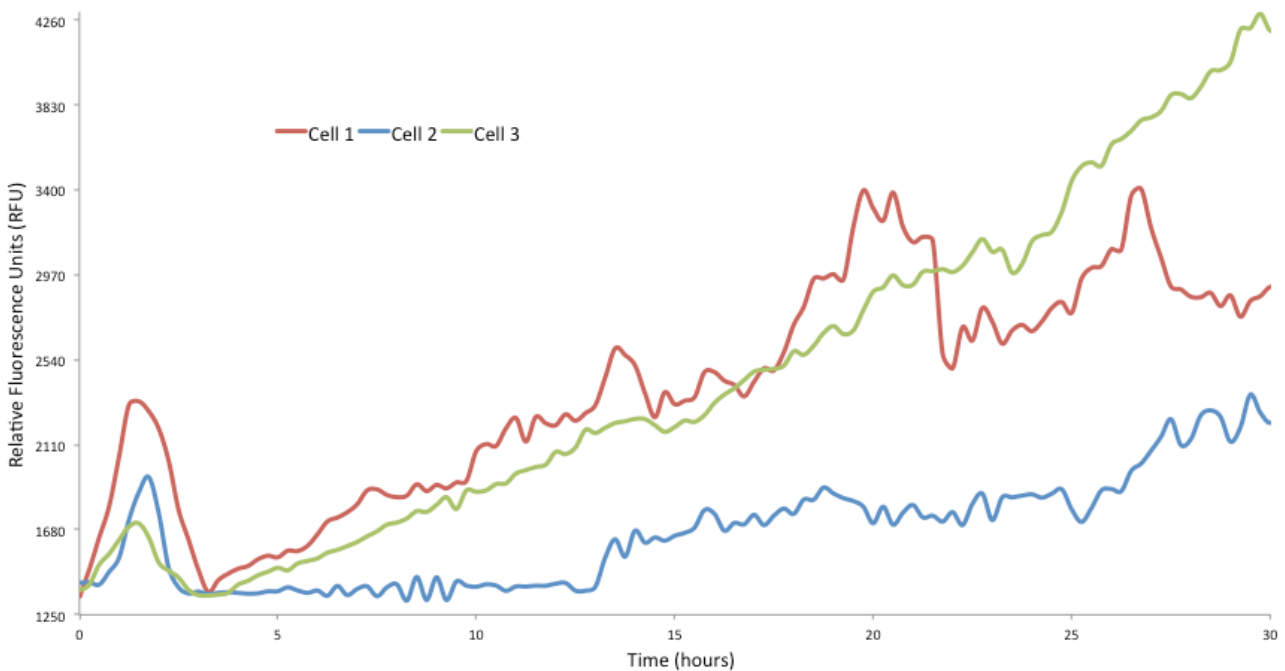


Figure 5.16. Single cell analysis of three cells exhibiting an unstable transient behaviour after co-transfection with equimolar ratios of pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pTRE3G_UbGFP, in complete DMEM medium supplied with 3 mM IPTG. The fluorescence channel images corresponding to the fluorescence peak observed in each of these cells are presented in *Appendix A.5*. 147

In vivo implementation of a novel mammalian synthetic gene oscillator

Different fluorescence time evolution profiles were recorded across the three experimental conditions. HEK293T cells supplemented with 1 mM IPTG and co-transfected with pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pTRE3G_UbGFP at equimolar ratios revealed one fluorescence peak followed by an approximately constant basal fluorescence signal. Likewise, one fluorescence peak was detected in cells supplemented with 2 mM IPTG; however, this transient unstable behaviour was followed by a non-constant signal. Cells supplemented with 3 mM IPTG under the same plasmid co-transfection conditions also revealed one fluorescence peak, but followed by an approximately linear increase in fluorescence intensity that lasted until the end of the image acquisition protocol.

Furthermore, these differences in the general shape of the fluorescence time profiles detected for HEK293T cells supplemented with 1 mM, 2 mM, and 3 mM IPTG was accompanied by differences in the average period of observed fluorescence peaks. For illustration and comparative purposes, the fluorescence peak periods across different experimental conditions are shown in figure 5.17.

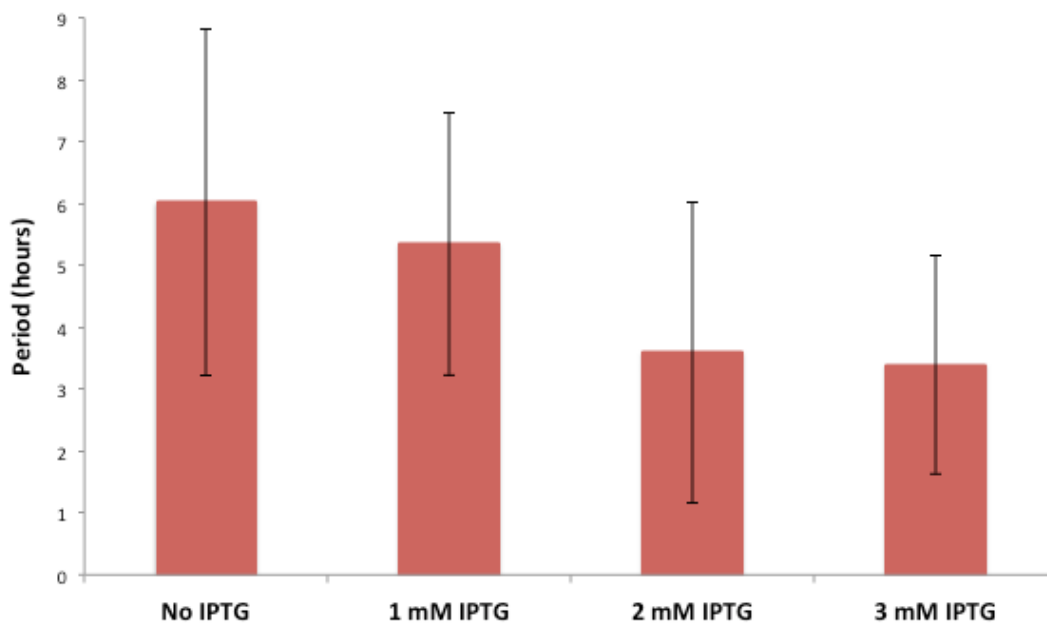


Figure 5.17. Averaged detected single-cell fluorescence peak periods for HEK293T cells co-transfected with pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pTRE3G_UbGFP at equimolar ratios, and supplied with 1 mM, 2mM, and 3 mM IPTG. Average values presented herein are for populations of 29, 13, 8, and 8 cells for each experimental condition, respectively. Error bars represent standard deviation.

In vivo implementation of a novel mammalian synthetic gene oscillator

IPTG was expected to have an effect on system dynamics due to its LacI inhibitory activity (Aubrecht et al., 1996; Cronin et al., 2001; Hu & Davidson, 1987; Lutz & Bujard, 1997; Wyborski & Short, 1991). Indeed, a decrease in the average period of fluorescence peaks was observed: cells that were not supplemented with IPTG showed approximately the same average period in this peak as cells supplemented with 1 mM IPTG; cells supplemented with 2 mM and 3 mM IPTG presented lower average periods. In addition, no statistically significant difference between the average periods in the four experimental conditions outlined in figure 5.17 was detected: the great variability in single-cell fluorescence peak periods across cellular populations is reflected in this statistical feature. IPTG-mediated period reduction has been predicted by simulating the mathematical model presented in *Chapter 3*. According to simulation results, the kinetic parameter governing IPTG concentration has the potential of changing the period of oscillations from c. 875 minutes (0 mM IPTG) to c. 700 minutes (3 mM IPTG). Although the predicted and experimental observations do not agree, the model-predicted tunability potential of IPTG was, in fact, observed in the experimental results. Moreover, the corresponding IPTG-mediated tunability simulations reveal that the period of oscillations is approximately the same if 0 nM and 1000 nM IPTG concentrations are assumed: IPTG-mediated tuning is relevant for IPTG concentrations higher than 1000 nM. This model-predicted feature is also reflected in the results shown in figure 5.17.

In addition to the peak period-shortening observations, it is clear from figures 5.14 to 5.16 that an increase in IPTG concentration leads to a higher accumulation of GFP after fluorescence peaks take place. This can be explained by an increase in IPTG-mediated inhibition: higher IPTG concentrations lead to higher alleviation of LacI-mediated repression. Indeed, cells supplied with 3 mM IPTG present a progressive increase in fluorescence intensity throughout the experimental protocol, after the fluorescence cycle occurs. This suggests that the differences in fluorescence time profiles amongst cells supplied with 1 mM, 2mM and 3 mM IPTG are due to a differential degree of IPTG-mediated LacI inhibition. Moreover, it corroborates the

role of LacI as the negative feedback effector within the topology of the proposed mammalian synthetic gene oscillator.

Doxycycline was found to abolish the unstable transient behaviour. This finding contrasts with model predictions, which highlighted the tunability potential of this chemical inhibitor. However, the model simulations have also predicted the tTA protein to oscillate at a considerably lower concentration than its system counterparts. This represents a possible explanation for the doxycycline-mediated abolishment of transient behaviour: lower intracellular tTA concentrations require less doxycycline in order to ensure tTA inhibition. Thus, the doxycycline concentrations provided in this experimental setup were sufficient to prevent unstable behaviour, instead of allowing for system tuning.

5.5. Concluding remarks

The main objective of this chapter was to address and describe the processes leading to the *in vitro* construction of the proposed mammalian synthetic gene oscillator, as well as its *in vivo* performance. A comprehensive review of the employed module re-organization strategy was also provided, followed by a list of methodologies that were used to assemble each system module.

The assembly of the system consisted, at a first stage, on the construction of four mammalian expression vectors: pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4 and pTRE3G_UbGFP. pUAS_LacI encoded the LacI system module, that is, a UAS-ISV8 intron structure located upstream of a codon-optimized *lacI^f* ORF, with downstream regulatory elements that ensured mRNA maturation and protein nuclear translocation. pOPRS_tTA encoded a *lacO*-SV40 intron structure located upstream of a *tTA* ORF, also with downstream regulatory elements that ensured maturation and nuclear translocation of mRNA and protein. pTRE3G_Gal4 encoded a P_{TRE3G} promoter located upstream of a *gal4* ORF, also containing downstream regulatory elements. The tTA-responsive reporter module consisted on an ubiquitin-tagged *gfp* gene with an upstream P_{TRE3G} promoter, and was encoded in pTRE3G_UbGFP. The assembly,

In vivo implementation of a novel mammalian synthetic gene oscillator

amplification, and purification of these four plasmids were validated *in vitro* using a combination of restriction enzyme digestion, agarose gel electrophoresis, and spectrophotometry.

The *in vivo* implementation of the proposed system revealed contradictions with model predictions. Indeed, cells showed an unstable transient behaviour with 1 fluorescence peak, instead of oscillatory behaviour. Moreover, doxycycline was found to inhibit unstable behaviour, and IPTG was found to reduce the period of fluorescence peaks. These results were contextualized with previous synthetic gene oscillator implementations, and discussed accordingly.

The next chapter provides a summary of generalized conclusions of this thesis. Additionally, on-going and future research directions, as well as strategies for improving the *in vivo* performance of the proposed system, are presented.

Chapter 6

Future work and general conclusions

The main objective of this thesis was to propose a novel mammalian synthetic gene oscillator design and provide a detailed description of its *in vitro* construction and *in vivo* performance. As described in *Chapter 1*, the fundamental motivation behind the work presented herein concerned the shortage of available mammalian synthetic gene oscillator designs and implementations, as well as the low robustness reported in previous implementations of such a system.

A comprehensive review of the most iconic synthetic gene oscillators implemented to date was presented in *Chapter 2*. Indeed, several topologies and biological organizations of these complex synthetic gene networks were characterized and described, from the early singular self-repression-based Goodwin oscillator to the complex mammalian Fussenegger oscillators. Additionally, fundamental principles underlying the design of synthetic gene oscillators were outlined and explored: (i) negative feedback is required for a cyclic process to be created, (ii) delay in negative feedback prevents a system from settling in a stable steady state, (iii) non-linearity is fundamental allows for steady state de-stabilization, and (iv) timescale matching in the synthesis and degradation rates of involved biochemical species is fundamental for oscillatory behaviour to be sustained. A comparative analysis of design and implementation features between the reviewed oscillator systems was then presented and discussed.

The design principles and biological features presented in *Chapter 2* were taken into account when designing a novel mammalian synthetic gene oscillator, presented in *Chapter 3*. Indeed, a comprehensive implementation strategy based on a 3-component delayed negative feedback topology was substantially discussed, from both a dynamical systems and molecular biology perspective. *In vivo* implementation feasibility guidelines were further established and incorporated in the biological aspects of design, which was chosen to rely on the down-regulatory activity of the Lac repressor, and the up-regulatory activities of Gal4 and tTA. An ODE-based mathematical model comprising 8 equations and representing the dynamics of the proposed synthetic gene oscillator was then presented, in conjugation with a brief recapitulation of dynamical systems theory and bifurcation analysis. The dynamic model was simulated and revealed that the chosen topology

and parameter values generated sustained oscillatory behaviour across all system components. Additionally, the model simulations revealed that IPTG and doxycycline had the potential to act as tunability factors of the system, through LacI and tTA inhibition, respectively. The bifurcation analysis performed on this dynamic model allowed for (i) a detailed mapping framework of the parametric regions corresponding to sustained oscillatory behaviour, and (ii) a theoretical exploration of the extent of system tunability using IPTG, doxycycline and alternative plasmid molarity ratios. Although providing an initial theoretical framework for system design, the dynamic model developed in *Chapter 3* did not take into account several biological features that are potentially relevant for the *in vivo* performance of the system, namely: the locations of operators and promoters relative to their corresponding transcription start sites (TSS), the mRNA and protein nuclear translocation processes, and the diffusion of biochemical species throughout the cell. These features are not considered in mathematical models of previous synthetic gene oscillator implementations, however, developing a model framework that takes these biological features into account would greatly improve the predictive power of the mathematical model presented in *Chapter 3*. Moreover, a global sensitivity analysis approach to this model would allow for a determination of a parametric ranking according to the magnitude of system dynamics disturbance. Taken in conjunction with bifurcation analysis, this analytical framework can potentially lead to a deeper understanding of the parametric space shaping oscillatory dynamics.

The insight provided by model simulations was further employed in the *in vitro* construction of synthetic gene circuits for system part *in vivo* validation, presented in *Chapter 4*. To this end, three synthetic gene circuits were assembled and implemented in HEK293T cells with the objective of validating system components. Each synthetic gene circuit consisted of a constitutively expressing module and a reporter module. The constitutively expressing modules were based on CMV promoters that were placed upstream of the three relevant ORFs: *lacI*, *tTA* and *gal4*. The reporter modules were respectively based on upstream *lacO* operator sites, a P_{TRE3G} promoter, and an UAS regulatory structure, with a downstream *gfp* ORF. All modules were encoded in six mammalian expression vectors, which were

subsequently assembled *in vitro*: (i) pCMVLacI and pOPRS_GFP comprised the LacI validation synthetic gene circuit; (ii) pCMVtTA and pZK14 comprised the tTA validation synthetic gene circuit; (iii) pCMVGal4 and pUAS-GG comprised the Gal4 validation synthetic gene circuit. The *in vitro* assembly, amplification, and purification of these six plasmids were validated using conventional molecular biology methodologies, namely restriction enzyme digestion, agarose gel electrophoresis, and spectrophotometry. Indeed, all plasmids were confirmed to have their corresponding expected sizes, a stable structure, acceptable concentration ranges in the context of mammalian cell transfection, and acceptable levels of purity according to spectrophotometric-based standards. The validation synthetic gene circuits were further implemented *in vivo* by co-transfection of HEK293T cells at equimolar ratios. Fluorescence channel images corroborated the expected results of each part validation experiment: (i) cells transfected with pOPRS_GFP, as well as cells co-transfected with pCMVLacI and pOPRS_GFP in medium supplemented with IPTG, were found to produce a progressively increasing fluorescence signal, unlike cells co-transfected with pCMVLacI and pOPRS_GFP in the absence of IPTG; (ii) cells co-transfected with pCMVtTA and pZK14 revealed an progressively increasing fluorescence signal, unlike cells transfected solely with pZK14 and cells co-transfected with pCMVtTA and pZK14 in medium supplied with doxycycline; (iii) cells co-transfected with pCMVGal4 and pUAS-GG showed a strong fluorescence signal, unlike cells transfected solely with pUAS-GG. These results implied the conclusion that the part validation aspect of biological design was successful. However, additional efforts can be carried out on the biological standardization of system parts. Further investigation and characterization of both upstream and downstream regulatory elements would provide valuable insight into the controllability of system parts, and consequently allow for a more standardized system implementation framework. Specifically, additional *in vivo* part validation experiments with alternative promoter structures, operator site organizations (in terms of number, spacing, and structure of operator sites), and poly-adenylation signals would provide a highly valuable increase in biological design modularity. Moreover, an experiment-based parameter estimation framework based on *in vivo*

part implementation would allow for a refinement of the kinetic parameters governing system dynamics. Taken in conjunction with the abovementioned additional part validation experiments with alternative regulatory elements, this approach has the potential of increasing the predictive power of the model presented in *Chapter 3*, thus permitting a more recursive and controllable synthetic gene oscillator construction/implementation process.

The successfully validated biological parts were further re-organized into finalized system modules and implemented *in vivo*, as presented in *Chapter 5*. To this end, three core system modules and one reporter module were assembled *in vitro* and encoded in four corresponding mammalian expression vectors that reflected the originally proposed delayed negative feedback topology: pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4, and pTRE3G_UbGFP. The LacI system module was encoded in pUAS_LacI, and essentially consisted on an UAS structure with a downstream *lacI* ORF. The tTA system module was encoded in pOPRS_tTA, and consisted on an upstream regulatory structure made up of *lacO* operator sites with a downstream *tTA* ORF. pTRE3G_Gal4 encoded the Gal 4 system module and was based on a P_{TRE3G} promoter with a downstream *gal4* ORF. The reporter module was based on an ubiquitin-tagged *gfp* ORF, in order to allow for GFP degradation kinetics that were in line with other system components, with an upstream P_{TRE3G} promoter. In similar fashion to the procedures carried out with the part validation plasmids presented in *Chapter 4*, the *in vitro* assembly, amplification and purification of pUAS_LacI, pOPRS_tTA, pTRE3G_Gal4, and pTRE3G_UbGFP were validated using restriction enzyme digestion, agarose gel electrophoresis, and spectrophotometry. Likewise, all plasmids were confirmed to have their corresponding expected sizes, a stable structure, acceptable concentration ranges in the context of mammalian cell transfection, and acceptable levels of purity according to spectrophotometric-based standards. The proposed mammalian synthetic gene oscillator was subsequently implemented in HEK293T cells through co-transfection at equimolar plasmid ratios. A minority of cells exhibited an unstable transient behaviour with one initial fluorescence peak. Moreover, great variability in fluorescence signal properties was recorded at the single-cell level across cellular populations: signals were

asynchronous and irregular, with different cells showing different peak amplitudes and periods. The system was also implemented in cells supplied with doxycycline and IPTG. Although doxycycline was found to abolish unstable behaviour, IPTG-supplemented cells also exhibited an unstable transient behaviour. Moreover, the period of single-cell fluorescence peaks, as well as post-peak dynamics, was found to be sensitive to IPTG. An increasing concentration of IPTG revealed an average shorter period in fluorescence peaks, also at the single-cell level, thus partially corroborating model predictions. The dynamic behaviour of the system in an experimental context did not match model predictions since no oscillatory behaviour was observed. This aspect was extensively discussed in *Chapter 5*. Indeed, additional efforts can be carried out to improve the *in vivo* performance of the system. The most fundamental aspect in this context regards the construction of a mammalian expression vector simultaneously encoding the three core components of the system: *lacI*, *tTA* and *gal4*. Implementing the proposed mammalian synthetic gene oscillator with such a plasmid would allow for two main developments: (i) it ensures that cells would internalize an equimolar ratio of system components; (ii) it reduces the combinatorial complexity of components uptake, thus increasing the overall transfection efficiency of the system. Ensuring true equimolar system component uptake would prevent cells from internalizing alternative component molarity ratios, which was shown to be a potential cause for the lack of fluorescence oscillations. Increasing system transfection efficiency would allow for more cells to exhibit oscillatory behaviour, thus permitting a more elaborate analysis, possibly at the populational level, of system dynamics. Indeed, a '*master*' plasmid simultaneously encoding the LacI, tTA and Gal4 modules is currently under construction.

The synthetic gene oscillator presented in this work represents, to date, the only mammalian synthetic gene system that is capable of dynamically reacting to a chemical inhibitor whilst presenting semi-oscillatory dynamics. Moreover, besides providing further insight into the dynamics of gene oscillators, the unstable transient behaviour generated by the proposed system is applicable to the study of multiple anti-inflammatory pathways that show similar dynamic profiles, namely p53, Hes1, and NF- κ B, thus holding a potential for therapeutic applications (Ciliberto, Novak, &

Future work and general conclusions

Tyson, 2005; Hayot & Jayaprakash, 2006; Kobayashi & Kageyama, 2010; Monk, 2003; Nelson et al., 2004).

References

- Abil, Z., Xiong, X., & Zhao, H. (2014). Synthetic Biology for Therapeutic Applications. *Molecular Pharmaceutics*. <http://doi.org/10.1021/mp500392q>
- Allgower, E. L., & Georg, K. (2003). Introduction to Numerical Continuation Methods. *Classics in Applied Mathematics*, 45, xxvi–388. <http://doi.org/10.2307/2153001>
- Alon, U. (2007). *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman Hall/CRC mathematical and computational biology series (Vol. 10). <http://doi.org/citeulike-article-id:1314150>
- Andersen, J. B., Sternberg, C., Poulsen, L. K., Bjørn, S. P., Givskov, M., & Molin, S. (1998). New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Applied and Environmental Microbiology*, 64(6), 2240–2246.
- Anderson, J. C., Clarke, E. J., Arkin, A. P., & Voigt, C. A. (2006). Environmentally controlled invasion of cancer cells by engineered bacteria. *Journal of Molecular Biology*, 355(4), 619–27. <http://doi.org/10.1016/j.jmb.2005.10.076>
- Atkinson, M. R., Savageau, M. A., Myers, J. T., & Ninfa, A. J. (2003). Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell*, 113(5), 597–607. [http://doi.org/10.1016/S0092-8674\(03\)00346-5](http://doi.org/10.1016/S0092-8674(03)00346-5)
- Aubel, D., & Fussenegger, M. (2010). Mammalian synthetic biology - From tools to therapies. *BioEssays*, 32(4), 332–345. <http://doi.org/10.1002/bies.200900149>
- Aubrecht, J., Manivasakam, P., & Schiestl, R. H. (1996). Controlled gene expression in mammalian cells via a regulatory cascade involving the tetracycline transactivator and lac repressor. *Gene*, 172(2), 227–231. [http://doi.org/10.1016/0378-1119\(96\)00216-8](http://doi.org/10.1016/0378-1119(96)00216-8)
- Ausländer, S., Ausländer, D., Müller, M., Wieland, M., & Fussenegger, M. (2012). Programmable single-cell mammalian biocomputers. *Nature*, 487(7405), 123–7. <http://doi.org/10.1038/nature11149>
- Ausländer, S., & Fussenegger, M. (2013). From gene switches to mammalian designer cells: present and future prospects. *Trends in Biotechnology*, 31(3), 155–68.

- <http://doi.org/10.1016/j.tibtech.2012.11.006>
- Balagaddé, F. K., You, L., Hansen, C. L., Arnold, F. H., & Quake, S. R. (2005). Long-term monitoring of bacteria undergoing programmed population control in a microchemostat. *Science (New York, N.Y.)*, *309*(2005), 137–140.
<http://doi.org/10.1126/science.1109173>
- Barkai, N., & Leibler, S. (2000). Circadian clocks limited by noise. *Nature*, *403*(6767), 267–268. <http://doi.org/10.1038/35002258>
- Becskei, A., & Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, *405*(6786), 590–593. <http://doi.org/10.1038/35014651>
- Bindel, D. S., Demmel, J. W., Friedman, M. J., Govaerts, W. J., & Kuznetsov, Y. a. (2005). Bifurcation analysis of large equilibrium systems in MATLAB. *Computational Science - Iccs 2005, Pt 1, Proceedings*, *3514*(1), 50–57. Retrieved from ISI:000230414400007
- Bodnar, M., & Bartomiejczyk, A. (2012). Stability of delay induced oscillations in gene expression of Hes1 protein model. *Nonlinear Analysis: Real World Applications*, *13*, 2227–2239. <http://doi.org/10.1016/j.nonrwa.2012.01.017>
- Boulos, S., Meloni, B. P., Arthur, P. G., Bojarski, C., & Knuckey, N. W. (2006). Assessment of CMV, RSV and SYN1 promoters and the woodchuck post-transcriptional regulatory element in adenovirus vectors for transgene expression in cortical neuronal cultures. *Brain Research*, *1102*(1), 27–38.
<http://doi.org/10.1016/j.brainres.2006.04.089>
- Carey, M., Lin, Y. S., Green, M. R., & Ptashne, M. (1990). A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives. *Nature*, *345*, 361–364.
<http://doi.org/10.1038/345361a0>
- Church, G. M., Elowitz, M. B., Smolke, C. D., Voigt, C. a, & Weiss, R. (2014). Realizing the potential of synthetic biology. *Nature Reviews. Molecular Cell Biology*, *15*(4), 289–94. <http://doi.org/10.1038/nrm3767>
- Ciliberto, A., Novak, B., & Tyson, J. J. (2005). Steady states and oscillations in the p53/Mdm2 network. *Cell Cycle*, *4*, 488–493. <http://doi.org/10.4161/cc.4.3.1548>
- Conrad, E. D., Tyson, J. J., Rogers, R. C., & Day, M. V. (1999). *Mathematical Models of Biochemical Oscillations*.

- Conrad, E., Mayo, A. E., Ninfa, A. J., & Forger, D. B. (2008). Rate constants rather than biochemical mechanism determine behaviour of genetic clocks. *Journal of The Royal Society Interface*, 5(Suppl 1), S9–S15.
<http://doi.org/10.1098/rsif.2008.0046.focus>
- Corish, P., & Tyler-Smith, C. (1999). Attenuation of green fluorescent protein half-life in mammalian cells. *Protein Engineering*, 12(12), 1035–1040.
<http://doi.org/10.1093/protein/12.12.1035>
- Cronin, C. a., Gluba, W., & Scrable, H. (2001). The lac operator-repressor system is functional in the mouse. *Genes and Development*, 15(12), 1506–1517.
<http://doi.org/10.1101/gad.892001>
- Danino, T., Mondragón-Palomino, O., Tsimring, L., & Hasty, J. (2010). A synchronized quorum of genetic clocks. *Nature*, 463, 326–330.
<http://doi.org/10.1038/nature08753>
- Dhooge, a., Govaerts, W., & Kuznetsov, Y. a. (2003). MATCONT: A MATLAB package for numerical bifurcation analysis of ODEs. *ACM Transactions on Mathematical Software*, 29(2), 141–164. <http://doi.org/10.1145/779359.779362>
- Dhooge, A., Govaerts, W., & Kuznetsov, Y. A. (2004). Numerical continuation of branch points of limit cycles in MATCONT. *Computational Science - Iccs 2004, Pt 2, Proceedings*, 3037, 42–49. Retrieved from ISI:000222045500006
- Dibner, C., Schibler, U., & Albrecht, U. (2010). The Mammalian Circadian Timing System: Organization and Coordination of Central and Peripheral Clocks. *ANNUAL REVIEW OF PHYSIOLOGY*, 72, 517–549.
<http://doi.org/10.1146/annurev-physiol-021909-135821>
- Distel, M., Wullimann, M. F., & Köster, R. W. (2009). Optimized Gal4 genetics for permanent gene expression mapping in zebrafish. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 13365–13370.
<http://doi.org/10.1073/pnas.0903060106>
- Elowitz, M. B., & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767), 335–338. <http://doi.org/10.1038/35002125>
- Elowitz, M. B., & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, 403, 335–338. <http://doi.org/10.1038/35002125>

- Endy, D. (2005). Foundations for engineering biology. *Nature*, *438*(7067), 449–453.
<http://doi.org/10.1038/nature04342>
- Fraser, A., & Tiwari, J. (1974). Genetical feedback-repression. *Journal of Theoretical Biology*, *47*(2), 397–412. [http://doi.org/10.1016/0022-5193\(74\)90206-9](http://doi.org/10.1016/0022-5193(74)90206-9)
- Fung, E., Wong, W. W., Suen, J. K., Bulter, T., Lee, S., & Liao, J. C. (2005). A synthetic gene-metabolic oscillator. *Nature*, *435*, 118–122.
<http://doi.org/10.1038/nature03508>
- Fussenegger, M., Morris, R. P., Fux, C., Rimann, M., von Stockar, B., Thompson, C. J., & Bailey, J. E. (2000). Streptogramin-based gene regulation systems for mammalian cells. *Nature Biotechnology*, *18*(11), 1203–1208.
<http://doi.org/10.1038/81208>
- Gardner, T. S., Cantor, C. R., & Collins, J. J. (2000). Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, *403*(6767), 339–342.
<http://doi.org/10.1038/35002131>
- Gilbert, E. S., Walker, a W., & Keasling, J. D. (2003). A constructed microbial consortium for biodegradation of the organophosphorus insecticide parathion. *Applied Microbiology and Biotechnology*, *61*, 77–81.
<http://doi.org/10.1007/s00253-002-1203-5>
- Goldbeter, a. (1995). A model for circadian oscillations in the *Drosophila* period protein (PER). *Proceedings. Biological Sciences / The Royal Society*, *261*(1362), 319–324. <http://doi.org/10.1098/rspb.1995.0153>
- Gonze, D. (2010). Coupling oscillations and switches in genetic networks. *BioSystems*, *99*(1), 60–69. <http://doi.org/10.1016/j.biosystems.2009.08.009>
- Goodwin Waddington, Conrad Hal, B. C. (1963). *Temporal organization in cells a dynamic theory of cellular control processes*. London; New York: Academic Press.
- Gossen, M., & Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A*, *89*, 5547–5551.
<http://doi.org/10.1073/pnas.89.12.5547>
- Gossen, M., Freundlieb, S., Bender, G., Müller, G., Hillen, W., & Bujard, H. (1995). Transcriptional activation by tetracyclines in mammalian cells. *Science (New*

- York, N.Y.), 268(5218), 1766–1769. <http://doi.org/10.1126/science.7792603>
- Greber, D., & Fussenegger, M. (2007). Mammalian synthetic biology: engineering of sophisticated gene networks. *Journal of Biotechnology*, 130(4), 329–45. <http://doi.org/10.1016/j.jbiotec.2007.05.014>
- Guantes, R., & Poyatos, J. F. (2006). Dynamical principles of two-component genetic oscillators. *PLoS Computational Biology*, 2(3), 188–197. <http://doi.org/10.1371/journal.pcbi.0020030>
- Hasty, J., McMillen, D., & Collins, J. J. (2002). Engineered gene circuits. *Nature*, 420(6912), 224–230. <http://doi.org/10.1038/nature01257>
- Hayot, F., & Jayaprakash, C. (2006). NF-kappaB oscillations and cell-to-cell variability. *Journal of Theoretical Biology*, 240(4), 583–91. <http://doi.org/10.1016/j.jtbi.2005.10.018>
- Heng, B. C., Aubel, D., & Fussenegger, M. (2014). G protein-coupled receptors revisited: therapeutic applications inspired by synthetic biology. *Annual Review of Pharmacology and Toxicology*, 54, 227–49. <http://doi.org/10.1146/annurev-pharmtox-011613-135921>
- Hoffmann, A., Levchenko, A., Scott, M. L., & Baltimore, D. (2002). The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. *Science (New York, N.Y.)*, 298(5596), 1241–1245. <http://doi.org/10.1126/science.1071914>
- Hu, M. C. T. C., & Davidson, N. (1987). The inducible lac operator-repressor system is functional in mammalian cells. *Cell*, 48(4), 555–566. [http://doi.org/0092-8674\(87\)90234-0](http://doi.org/0092-8674(87)90234-0) [pii]
- Kakidani, H., & Ptashne, M. (1988). GAL4 activates gene expression in mammalian cells. *Cell*, 52(lane 1), 161–167. [http://doi.org/0092-8674\(88\)90504-1](http://doi.org/0092-8674(88)90504-1) [pii]
- Karlsson, M., & Weber, W. (2012). Therapeutic synthetic gene networks. *Current Opinion in Biotechnology*, 23(5), 703–11. <http://doi.org/10.1016/j.copbio.2012.01.003>
- Keiler, K. C., & Sauer, R. T. (1996). Sequence determinants of C-terminal substrate recognition by the Tsp protease. *Journal of Biological Chemistry*, 271(5), 2589–2593. <http://doi.org/10.1074/jbc.271.5.2589>

- Kemler, I., & Fontana, A. (1999). Role of IkappaBalpha and IkappaBbeta in the biphasic nuclear translocation of NF-kappaB in TNFalpha-stimulated astrocytes and in neuroblastoma cells. *Glia*, *26*, 212–220.
- Khalil, A. S., & Collins, J. J. (2010). Synthetic biology: applications come of age. *Nature Reviews. Genetics*, *11*(5), 367–379. <http://doi.org/10.1038/nrg2775>
- Khalil, A. S., Lu, T. K., Bashor, C. J., Ramirez, C. L., Pyenson, N. C., Joung, J. K., & Collins, J. J. (2012). A synthetic biology framework for programming eukaryotic transcription functions. *Cell*, *150*(3), 647–658. <http://doi.org/10.1016/j.cell.2012.05.045>
- Khosla, C., & Keasling, J. D. (2003). Metabolic Engineering For Drug Discovery and Development. *Drug Discovery*, *2*(December), 1019–1025. <http://doi.org/10.1038/nrd1256>
- Kis, Z., Sant'Ana Pereira, H., Homma, T., Pedrigi, R. M., & Krams, R. (2015). Mammalian synthetic biology: emerging medical applications. *Journal of the Royal Society, Interface*, *12*, 1–18. <http://doi.org/10.1098/rsif.2014.1000>
- Kobayashi, T., & Kageyama, R. (2010). Hes1 regulates embryonic stem cell differentiation by suppressing Notch signaling. *Genes to Cells*, *15*, 689–698. <http://doi.org/10.1111/j.1365-2443.2010.01413.x>
- Kuznetsov, A., & Afraimovich, V. (2012). Heteroclinic cycles in the repressilator model. *Chaos, Solitons & Fractals*, *45*(5), 660–665. <http://doi.org/10.1016/j.chaos.2012.02.009>
- Kwon, I., Choe, H. K., Son, G. H., & Kim, K. (2011). Mammalian molecular clocks. *Experimental Neurobiology*, *20*(1), 18–28. <http://doi.org/10.5607/en.2011.20.1.18>
- Lienert, F., Lohmueller, J. J., Garg, A., & Silver, P. a. (2014). Synthetic biology in mammalian cells: next generation research tools and therapeutics. *Nature Reviews. Molecular Cell Biology*, *15*(2), 95–107. <http://doi.org/10.1038/nrm3738>
- Lu, T. K., Khalil, A. S., & Collins, J. J. (2009). Next-generation synthetic gene networks. *Nature Biotechnology*, *27*(12), 1139–1150. <http://doi.org/10.1038/nbt.1591>
- Luo, X., Celler, J. W., & Berndt, A. (1999). LacSwitch inducible mammalian expression

- system in mouse Swiss 3T3 fibroblasts. *Molecular & Cellular Biochemistry*, 200(1-2), 127–132.
- Lutz, R., & Bujard, H. (1997). Independent and tight regulation of transcriptional units in escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Research*, 25(6), 1203–1210.
<http://doi.org/10.1093/nar/25.6.1203>
- May, T., Eccleston, L., Herrmann, S., Hauser, H., Goncalves, J., & Wirth, D. (2008). Bimodal and hysteretic expression in mammalian cells from a synthetic gene circuit. *PloS One*, 3(6), e2372. <http://doi.org/10.1371/journal.pone.0002372>
- McAdams, H. H., & Shapiro, L. (1995). Circuit simulation of genetic networks. *Science (New York, N.Y.)*, 269(5224), 650–656. <http://doi.org/10.1126/science.7624793>
- Momiji, H., & Monk, N. A. M. (2008). Dissecting the dynamics of the Hes1 genetic oscillator. *Journal of Theoretical Biology*, 254, 784–798.
<http://doi.org/10.1016/j.jtbi.2008.07.013>
- Mondragón-Palomino, O., Danino, T., Selimkhanov, J., Tsimring, L., & Hasty, J. (2011). Entrainment of a population of synthetic genetic oscillators. *Science (New York, N.Y.)*, 333(6047), 1315–1319. <http://doi.org/10.1126/science.1205369>
- Monk, N. A. M. (2003). Oscillatory expression of Hes1, p53, and NF- κ B driven by transcriptional time delays. *Current Biology*, 13(16), 1409–1413.
[http://doi.org/10.1016/S0960-9822\(03\)00494-9](http://doi.org/10.1016/S0960-9822(03)00494-9)
- Montagne, K., Plasson, R., Sakai, Y., Fujii, T., & Rondelez, Y. (2011). Programming an in vitro DNA oscillator using a molecular networking strategy. *Molecular Systems Biology*, 7(466), 466. <http://doi.org/10.1038/msb.2010.120>
- Mukherji, S., & van Oudenaarden, A. (2009). Synthetic biology: understanding biological design from synthetic circuits. *Nature Reviews. Genetics*, 10(12), 859–871. <http://doi.org/10.1038/nrg2697>
- Müller, S., Hofbauer, J., Endler, L., Flamm, C., Widder, S., & Schuster, P. (2006). A generalized model of the repressilator. *Journal of Mathematical Biology*, 53(6), 905–937. <http://doi.org/10.1007/s00285-006-0035-9>
- Nelson, D. E., Ihekwaba, a E. C., Elliott, M., Johnson, J. R., Gibney, C. a, Foreman, B. E., ... White, M. R. H. (2004). Oscillations in NF-kappaB signaling control the

- dynamics of gene expression. *Science (New York, N.Y.)*, 306(5696), 704–708.
<http://doi.org/10.1126/science.1099962>
- Nielsen, J., Fussenegger, M., Keasling, J., Lee, S. Y., Liao, J. C., Prather, K., & Palsson, B. (2014). Engineering synergy in biotechnology. *Nature Chemical Biology*, 10(5), 319–22. <http://doi.org/10.1038/nchembio.1519>
- Novák, B., & Tyson, J. J. (2008). Design principles of biochemical oscillators. *Nature Reviews. Molecular Cell Biology*, 9(12), 981–991.
<http://doi.org/10.1038/nrm2530>
- O'Brien, E. L., Itallie, E. Van, & Bennett, M. R. (2012). Modeling synthetic gene oscillators. *Mathematical Biosciences*, 236(1), 1–15.
<http://doi.org/10.1016/j.mbs.2012.01.001>
- O'Brien, E. L., Van Itallie, E., & Bennett, M. R. (2012). Modeling synthetic gene oscillators. *Mathematical Biosciences*.
<http://doi.org/10.1016/j.mbs.2012.01.001>
- Pedraza, J. M., & van Oudenaarden, A. (2005). Noise propagation in gene networks. *Science (New York, N.Y.)*, 307(5717), 1965–1969.
<http://doi.org/10.1126/science.1109090>
- Phelps, C. B., & Brand, A. H. (1998). Ectopic gene expression in *Drosophila* using GAL4 system. *Methods (San Diego, Calif.)*, 14, 367–379.
<http://doi.org/10.1006/meth.1998.0592>
- Potter, C. J., Tasic, B., Russler, E. V., Liang, L., & Luo, L. (2010). The Q system: A repressible binary system for transgene expression, lineage tracing, and mosaic analysis. *Cell*, 141(3), 536–548. <http://doi.org/10.1016/j.cell.2010.02.025>
- Purcell, O., Peccoud, J., & Lu, T. K. (2014). Rule-based design of synthetic transcription factors in eukaryotes. *ACS Synthetic Biology*, 3(10), 737–744.
<http://doi.org/10.1021/sb400134k>
- Purcell, O., Savery, N. J., Grierson, C. S., & di Bernardo, M. (2010). A comparative analysis of synthetic genetic oscillators. *Journal of the Royal Society, Interface / the Royal Society*, 7(52), 1503–1524. <http://doi.org/10.1098/rsif.2010.0183>
- Qin, J. Y., Zhang, L., Clift, K. L., Hukur, I., Xiang, A. P., Ren, B. Z., & Lahn, B. T. (2010). Systematic comparison of constitutive promoters and the doxycycline-inducible

- promoter. *PLoS ONE*, 5(5). <http://doi.org/10.1371/journal.pone.0010611>
- Rajala, T., Häkkinen, A., Healy, S., Yli-Harja, O., & Ribeiro, A. S. (2010). Effects of transcriptional pausing on gene expression dynamics. *PLoS Computational Biology*, 6(3), 29–30. <http://doi.org/10.1371/journal.pcbi.1000704>
- Rajendran, M., & Ellington, A. D. (2008). Selection of fluorescent aptamer beacons that light up in the presence of zinc. *Analytical and Bioanalytical Chemistry*, 390, 1067–1075. <http://doi.org/10.1007/s00216-007-1735-8>
- Rao, C. V., & Arkin, A. P. (2003). Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *Journal of Chemical Physics*, 118(11), 4999–5010. <http://doi.org/10.1063/1.1545446>
- Rateitschak, K., & Wolkenhauer, O. (2007). Intracellular delay limits cyclic changes in gene expression. *Mathematical Biosciences*, 205, 163–179. <http://doi.org/10.1016/j.mbs.2006.08.010>
- Ro, D.-K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., ... Keasling, J. D. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, 440(7086), 940–943. <http://doi.org/10.1038/nature04640>
- Ruder, W. C., Lu, T., & Collins, J. J. (2011). Synthetic biology moving into the clinic. *Science (New York, N.Y.)*, 333(6047), 1248–1252. <http://doi.org/10.1126/science.1206843>
- Sadowski, I., Bell, B., Broad, P., & Hollis, M. (1992). GAL4 fusion vectors for expression in yeast or mammalian cells. *Gene*, 118, 137–141. [http://doi.org/10.1016/0378-1119\(92\)90261-M](http://doi.org/10.1016/0378-1119(92)90261-M)
- Sadowski, I., Ma, J., Triezenberg, S., & Ptashne, M. (1988). GAL4-VP16 is an unusually potent transcriptional activator. *Nature*, 335(6190), 563–564. <http://doi.org/10.1038/335563a0>
- Samad, H. El, Vecchio, D. Del, & Khammash, M. (2005). Repressilators and promotilators: loop dynamics in synthetic gene networks. *Proceedings of the 2005, American Control Conference, 2005.*, 4405–4410. <http://doi.org/10.1109/ACC.2005.1470689>
- Scheer, N., & Campos-Ortega, J. A. (1999). Use of the Gal4-UAS technique for

- targeted gene expression in the zebrafish. *Mechanisms of Development*, 80, 153–158. [http://doi.org/10.1016/S0925-4773\(98\)00209-3](http://doi.org/10.1016/S0925-4773(98)00209-3)
- Serrano, L. (2007). Synthetic biology: promises and challenges. *Molecular Systems Biology*, 3(November), 158. <http://doi.org/10.1038/msb4100202>
- Smith, H. (1987). Oscillations and multiple steady states in a cyclic gene model with repression. *Journal of Mathematical Biology*, 25, 169–190. <http://doi.org/10.1007/BF00276388>
- Smolen, P., Baxter, D. A., & Byrne, J. H. (1998). Frequency selectivity, multistability, and oscillations emerge from models of genetic regulatory systems. *The American Journal of Physiology*, 274(2 Pt 1), C531–C542.
- Steen, E. J., Chan, R., Prasad, N., Myers, S., Petzold, C. J., Redding, A., ... Keasling, J. D. (2008). Metabolic engineering of *Saccharomyces cerevisiae* for the production of n-butanol. *Microbial Cell Factories*, 7, 36. <http://doi.org/10.1186/1475-2859-7-36>
- Stricker, J., Cookson, S., Bennett, M. R., Mather, W. H., Tsimring, L. S., & Hasty, J. (2008). A fast, robust and tunable synthetic gene oscillator. *Nature*, 456(7221), 516–519. <http://doi.org/10.1038/nature07389>
- Strogatz, S. H. (1994). Nonlinear Dynamics and Chaos. In *Library* (Vol. 48, p. 498). <http://doi.org/9780738204536>
- Susaki, E. A., Stelling, J., & Ueda, H. R. (2010). Challenges in synthetically designing mammalian circadian clocks. *Current Opinion in Biotechnology*. <http://doi.org/10.1016/j.copbio.2010.07.011>
- Tavassoli, A. (2010). Synthetic biology: scope, applications and implications. *Organic & Biomolecular Chemistry*, 8(1), 24–28. <http://doi.org/10.1039/b913300n>
- Tigges, M., Dénervaud, N., Greber, D., Stelling, J., & Fussenegger, M. (2010). A synthetic low-frequency mammalian oscillator. *Nucleic Acids Research*, 38(8), 2702–2711. <http://doi.org/10.1093/nar/gkq121>
- Tigges, M., & Fussenegger, M. (2009). Recent advances in mammalian synthetic biology-design of synthetic transgene control networks. *Current Opinion in Biotechnology*, 20(4), 449–460. <http://doi.org/10.1016/j.copbio.2009.07.009>
- Tigges, M., Marquez-Lago, T. T., Stelling, J., & Fussenegger, M. (2009). A tunable

- synthetic mammalian oscillator. *Nature*, 457(7227), 309–312.
<http://doi.org/10.1038/nature07616>
- Traven, A., Jelacic, B., & Sopta, M. (2006). Yeast Gal4: a transcriptional paradigm revisited. *EMBO Reports*, 7, 496–499. <http://doi.org/10.1038/sj.embor.7400679>
- Tsai, T. Y.-C., Choi, Y. S., Ma, W., Pomerening, J. R., Tang, C., & Ferrell, J. E. (2008). Robust, tunable biological oscillations from interlinked positive and negative feedback loops. *Science (New York, N.Y.)*, 321(5885), 126–129.
<http://doi.org/10.1126/science.1156951>
- Tsimring, L., Hasty, J., Danino, T., Mondrago, O., Mondragón-Palomino, O., Tsimring, L., & Hasty, J. (2010). A synchronized quorum of genetic clocks. *Nature*, 463(7279), 326–330. <http://doi.org/10.1038/nature08753>
- Tyson, J. J., Albert, R., Goldbeter, A., Ruoff, P., & Sible, J. (2008). Biological switches and clocks. *Journal of the Royal Society, Interface / the Royal Society*, 5 Suppl 1(August), S1–S8. <http://doi.org/10.1098/rsif.2008.0179.focus>
- Ueda, H. R. (2007). Systems biology of mammalian circadian clocks. In *Cold Spring Harbor Symposia on Quantitative Biology* (Vol. 72, pp. 365–380).
<http://doi.org/10.1101/sqb.2007.72.047>
- Ueda, H. R., Hayashi, S., Chen, W., Sano, M., Machida, M., Shigeyoshi, Y., ... Hashimoto, S. (2005). System-level identification of transcriptional circuits underlying mammalian circadian clocks. *Nature Genetics*, 37(2), 187–92.
<http://doi.org/10.1038/ng1504>
- Ukai, H., & Ueda, H. R. (2010). Systems biology of mammalian circadian clocks. *Annual Review of Physiology*, 72, 579–603. <http://doi.org/10.1146/annurev-physiol-073109-130051>
- Ukai-Tadenuma, M., Kasukawa, T., & Ueda, H. R. (2008). Proof-by-synthesis of the transcriptional logic of mammalian circadian clocks. *Nature Cell Biology*, 10(10), 1154–1163. <http://doi.org/10.1038/ncb1775>
- Waks, Z., & Silver, P. a. (2009). Engineering a synthetic dual-organism system for hydrogen production. *Applied and Environmental Microbiology*, 75(7), 1867–75.
<http://doi.org/10.1128/AEM.02009-08>
- Wang, R., Jing, Z., & Chen, L. (2005). Modelling periodic oscillation in gene regulatory

- networks by cyclic feedback systems. *Bulletin of Mathematical Biology*, 67(2), 339–367. <http://doi.org/10.1016/j.bulm.2004.07.005>
- Weber, W., & Fussenegger, M. (2009). Engineering of synthetic mammalian gene networks. *Chemistry & Biology*, 16(3), 287–97. <http://doi.org/10.1016/j.chembiol.2009.02.005>
- Weber, W., & Fussenegger, M. (2010). Synthetic gene networks in mammalian cells. *Current Opinion in Biotechnology*, 21(5), 690–6. <http://doi.org/10.1016/j.copbio.2010.07.006>
- Weber, W., Kramer, B. P., Fux, C., Keller, B., & Fussenegger, M. (2002). Novel promoter/transactivator configurations for macrolide- and streptogramin-responsive transgene expression in mammalian cells. *The Journal of Gene Medicine*, 4(6), 676–86. <http://doi.org/10.1002/jgm.314>
- Webster, N., Jin, J. R., Green, S., Hollis, M., & Chambon, P. (1988). The yeast UASG is a transcriptional enhancer in human HeLa cells in the presence of the GAL4 trans-activator. *Cell*, 52, 169–178. [http://doi.org/10.1016/0092-8674\(88\)90505-3](http://doi.org/10.1016/0092-8674(88)90505-3)
- Wei, X., Potter, C. J., Luo, L., & Shen, K. (2012). Controlling gene expression with the Q repressible binary expression system in *Caenorhabditis elegans*. *Nature Methods*, 9(4), 391–5. <http://doi.org/10.1038/nmeth.1929>
- Wyborski, D. L., & Short, J. M. (1991). Analysis of inducers of the E.coli lac repressor system in mammalian cells and whole animals. *Nucleic Acids Research*, 19(17), 4647–4653. <http://doi.org/10.1093/nar/19.17.4647>
- Yao, F., Svensjö, T., Winkler, T., Lu, M., Eriksson, C., & Eriksson, E. (1998). Tetracycline repressor, tetR, rather than the tetR-mammalian cell transcription factor fusion derivatives, regulates inducible gene expression in mammalian cells. *Human Gene Therapy*, 9, 1939–1950. <http://doi.org/10.1089/hum.1998.9.13-1939>
- Ye, H., & Fussenegger, M. (2014). Synthetic therapeutic gene circuits in mammalian cells. *FEBS Letters*, 588(15), 2537–2544. <http://doi.org/10.1016/j.febslet.2014.05.003>
- Zarrin, A. A., Malkin, L., Fong, I., Luk, K. D., Ghose, A., & Berinstein, N. L. (1999).

References

Comparison of CMV, RSV, SV40 viral and V β 1 cellular promoters in B and T lymphoid and non-lymphoid cell lines. *Biochimica et Biophysica Acta - Gene Structure and Expression*. [http://doi.org/10.1016/S0167-4781\(99\)00067-6](http://doi.org/10.1016/S0167-4781(99)00067-6)

Appendix A.3

A.3.1. List of parameters for the ODE-based model described by equations 3.1 to 3.10

Table A.3.1. List of parameters for the ODE-based model described by equations 3.1 to 3.10

Parameter	Symbol	Value retrieved from literature	Reference
Lacl mRNA transcription rate (min^{-1})	k_1	30	Betrabet <i>et al.</i> (2004); Tigges <i>et al.</i> (2009)
tTA mRNA transcription rate (min^{-1})	k_2	30	Tigges <i>et al.</i> (2009)
Gal4 and GFP mRNA transcription rate (min^{-1})	k_3	30	Tigges <i>et al.</i> (2009)
Lacl mRNA degradation rate (min^{-1})	k_{dm1}	0.0173	Tigges <i>et al.</i> (2009)
tTA mRNA degradation rate (min^{-1})	k_{dm2}	0.0173	Tigges <i>et al.</i> (2009)
Gal4 mRNA degradation rate (min^{-1})	k_{dm3}	0.0173	Tigges <i>et al.</i> (2009)
GFP mRNA degradation rate	k_{dm4}	0.0173	Tigges <i>et al.</i> (2009)
General translation rate (min^{-1})	k_4	0.02	Tigges <i>et al.</i> (2009)

LacI degradation rate (min^{-1})	k_{dp1}	0.01	Pace <i>et al.</i> (1990); Elowitz and Leibler (2000)
tTA degradation rate (min^{-1})	k_{dp2}	0.0231	Tigges <i>et al.</i> (2009)
Gal4 degradation rate (min^{-1})	k_{dp3}	0.0069	Yotov <i>et al.</i> (1998); N
GFP degradation rate (min^{-1})	k_{dp4}	0.0058	Corish and Tyler-Smith (1999)
UAS:Gal4 dissociation constant (nM)	k_{d1}	24	Hong <i>et al.</i> (2008)
tetO:tTA dissociation constant (nM)	k_{d2}	5.6 nM	Tigges <i>et al.</i> (2009)
lacO:LacI dissociation constant (nM)	k_{d3}	40 nM	Elowitz and Leibler (2000)
Gal4 Hill coefficient (adimensional)	n_1	2	Hong <i>et al.</i> (2008)
tTA Hill coefficient (adimensional)	n_2	2	Tigges <i>et al.</i> (2009)
LacI Hill coefficient (adimensional)	n_3	4	Pace <i>et al.</i> (1990)
LacI:IPTG	n_4	4	Alon (2007)

interaction cooperativity coefficient			
IPTG:Lacl			
interaction affinity (nM)	K_L	1200	Alon (2007)
Tetracycline:tTA			
interaction affinity (nM)	K_T	1	Tigges <i>et al.</i> (2009)
lacl relative gene dosage (adimensional)	g_1	Variable through experimental setup	N/A
tTA relative gene dosage (adimensional)	g_2	Variable through experimental setup	N/A
gal4 relative gene dosage (adimensional)	g_3	Variable through experimental setup	N/A
gfp relative gene dosage (adimensional)	g_4	Variable through experimental setup	N/A
Tetracycline concentration (nM)	T	Variable through experimental setup	N/A
IPTG concentration (nM)	I	Variable through experimental setup	N/A

A.3.2. MATLAB code for specifying the model described by equations 3.1 to 3.10

Main .m file:

```

% This m-file describes a 3-component gene regulatory network with a
% delayed negative feedback (i.e. positive interactions delay the negative
% feedback). This topology was proposed by Tyson and Novak in their 2008
% review. This model considers the interactions of two trans-activators and
% one repressor. We will consider the activators to be tTA and Gal4
% (used in mammalian cells by Tigges and Fussenegger (2009) and Ptashne et
% al (1988)).

function dydt = System_Thesis(~, y)

% Declare variables
mRNA_LacI = y(1);
mRNA_tTA = y(2);
mRNA_Gal4 = y(3);
mRNA_GFP = y(4);
LacI = y(5);
tTA = y(6);
Gal4 = y(7);
GFP = y(8);

% Parameter values

% Transcription and translation rates
k1 = 30; % maximal LacI transcription rate, min(-1) (Tigges et al., 2009)
k2 = 30; % maximal tTA transcription rate, min(-1) (Tigges et al., 2009)
k3 = 30; % maximal Gal4 transcription rate, min(-1) (Tigges et al., 2009)
k4 = 0.02; % general translation rate, min(-1) (Tigges et al., 2009)

% Dissociation constants
kd1 = 24; % dissociation constant of UAS box, nM (Hong et al., 2008)
kd2 = 5.6; % dissociation constant of tetO box, nM (Tigges et al., 2009; Kamionka et
al., 2004)
kd3= 40; % dissociation constant of LacI from LacO box, nM (Elowitz and Leibler,
2000)

% Hill coefficients
n1 = 2; % Gal4 Hill coefficient (Pace et al., 1988)
n2 = 2; % tTA Hill coefficient (Tigges et al., 2009)

```


$n3 = 4$; % LacI Hill coefficient (Elowitz and Leibler, 2000)

% mRNA half-lives assumed to be equal to each other (Tigges et al., 2009)

mRNA_LacI_halflife = 40.0663; % LacI mRNA half-life, min

mRNA_tTA_halflife = 40.0663; % tTA mRNA half-life, min

mRNA_Gal4_halflife = 40.0663; % Gal4 mRNA half-life, min

mRNA_GFP_halflife = 40.0663; % UbGFP mRNA half-life, min

kdm1 = $\log(2)/\text{mRNA_LacI_halflife}$; % LacI mRNA degradation rate, min⁽⁻¹⁾

kdm2 = $\log(2)/\text{mRNA_tTA_halflife}$; % tTA mRNA degradation rate, min⁽⁻¹⁾

kdm3 = $\log(2)/\text{mRNA_Gal4_halflife}$; % Gal4 mRNA degradation rate, min⁽⁻¹⁾

kdm4 = $\log(2)/\text{mRNA_GFP_halflife}$; % UbGFP mRNA degradation rate, min⁽⁻¹⁾

% Protein half-lives

LacI_protein_halflife = 30; % LacI half-life, min (Pace et al., 1989 - study in E. coli)

tTA_protein_halflife = 30.0664; % tTA half-life, min (Tigges et al., 2009 - in CHO cells)

Gal4_protein_halflife = 100; % Gal4 half-life, min (Yotov et al., 1998; Nalley et al., 2006; Salghetti et al., 1999)

GFP_protein_halflife = 120; % UbGFP half-life, min (Tigges et al., 2009)

kdp1 = $\log(2)/\text{LacI_protein_halflife}$; % LacI degradation rate, min⁽⁻¹⁾

kdp2 = $\log(2)/\text{tTA_protein_halflife}$; % tTA degradation rate, min⁽⁻¹⁾

kdp3 = $\log(2)/\text{Gal4_protein_halflife}$; % Gal4 degradation rate, min⁽⁻¹⁾

kdp4 = $\log(2)/\text{GFP_protein_halflife}$; % GFP degradation rate, min⁽⁻¹⁾

% Gene dosages

g1 = 1; % LacI gene dosage, nM

g2 = 1; % tTA gene dosage, nM

g3 = 1; % Gal4 gene dosage, nM

g4 = 1; % GFP gene dosage, nM

% Chemical inhibitors

Tc = 0; % Tetracycline concentration, nM

IPTG = 0; % IPTG concentration, nM

kmt = 1; % Tetracycline-tTA interaction strength, nM (Tigges et al., 2009)

kmL = 1200; % IPTG:LacI interaction (Alon, 2007)

n4 = 4; % LacI-IPTG Hill coefficient (Alon, 2007)

Ta = $\text{tTA} * (1 - (\text{Tc} / (\text{kmt} + \text{Tc})))$; % Tetracycline transfer function for tTA activity control

% ODEs describing transcription and translation, including transcriptional

% regulation by the three components, i.e. inhibitory effect of LacI, and

% upregulating effects of tTA and Gal4

% mRNA production dynamics

dydt(1,:) = $g1 * ((k1 * (\text{Gal4}^{\text{n1}})) / (\text{kd1}^{\text{n1}} + (\text{Gal4}^{\text{n1}}))) - \text{kdm1} * \text{mRNA_LacI}$; % LacI mRNA dynamics

```
dydt(2,:) = g2*(k2/(1+(((LacI^n3)/(kd3^n3))/(1+(IPTG/kmL)^n4)))) -  
kdm2*mRNA_tTA; % tTA mRNA dynamics
```

```
dydt(3,:) = g3*((k3*(Ta^n2))/(kd2^n2+(Ta^n2))) - kdm3*mRNA_Gal4; % Gal4 mRNA  
dynamics
```

```
dydt(4,:) = g4*((k3*(Ta^n2))/(kd2^n2+(Ta^n2))) - kdm4*mRNA_GFP; % UbGFP mRNA  
dynamics
```

```
% Protein production dynamics
```

```
dydt(5,:) = k4*mRNA_LacI - kdp1*LacI; % LacI protein dynamics
```

```
dydt(6,:) = k4*mRNA_tTA - kdp2*tTA; % tTA protein dynamics
```

```
dydt(7,:) = k4*mRNA_Gal4 - kdp3*Gal4; % Gal4 protein dynamics
```

```
dydt(8,:) = k4*mRNA_GFP - kdp4*GFP; % UbGFP protein dynamics
```

Solver and plotter .m file:

```

% Solver for System_Thesis.m
% y-axis in nM, x-axis in minutes

sol = ode45(@System_Thesis, [0 7500], [0 0 0 0 0 0 0]);

% Plotter - Protein figures

% figure
% subplot(2,2,1);
% plot(sol.x, sol.y(5,:), 'r', 'LineWidth', 4)
% xlim([1500 7500])
% xlabel('Time (minutes)')
% ylabel('LacI Protein concentration (nM)')
%
% subplot(2,2,2);
% plot(sol.x, sol.y(6,:), 'r', 'LineWidth', 4)
% xlim([1500 7500])
% xlabel('Time (minutes)')
% ylabel('tTA Protein concentration (nM)')
%
% subplot(2,2,3);
% plot(sol.x, sol.y(7,:), 'r', 'LineWidth', 4)
% xlim([1500 7500])
% xlabel('Time (minutes)')
% ylabel('Gal4 Protein concentration (nM)')
%
% subplot(2,2,4);
% plot(sol.x, sol.y(8,:), 'r', 'LineWidth', 4)
% xlim([1500 7500])
% xlabel('Time (minutes)')
% ylabel('GFP Protein concentration (nM)')

% Plotter - mRNA figures

% figure
% subplot(2,2,1);
% plot(sol.x, sol.y(1,:), 'r', 'LineWidth', 4)
% xlim([1500 7500])
% xlabel('Time (minutes)')
% ylabel('LacI mRNA concentration (nM)')
%
% subplot(2,2,2);
% plot(sol.x, sol.y(2,:), 'r', 'LineWidth', 4)
% xlim([1500 7500])

```

```
% xlabel('Time (minutes)')
% ylabel('tTA mRNA concentration (nM)')
%
% subplot(2,2,3);
% plot(sol.x, sol.y(3,:), 'r', 'LineWidth', 4)
% xlim([1500 7500])
% xlabel('Time (minutes)')
% ylabel('Gal4 mRNA concentration (nM)')
%
% subplot(2,2,4);
% plot(sol.x, sol.y(8,:), 'r', 'LineWidth', 4)
% xlim([1500 7500])
% xlabel('Time (minutes)')
% ylabel('GFP mRNA concentration (nM)')
```

A.3.3. Model specification in Matcont

$$\text{Lacm}' = a1 + g1 * ((k1 * (\text{Gal}^{\wedge}n1)) / (kd1^{\wedge}n1 + (\text{Gal}^{\wedge}n1))) - kdm1 * \text{Lacm};$$

$$\text{tTAm}' = a2 + g2 * (k2 / (1 + (((\text{Lac}^{\wedge}n3) / (kd3^{\wedge}n3)) / (1 + (\text{IPTG} / kmL)^{\wedge}n4)))) - kdm2 * \text{tTAm};$$

$$\text{Galm}' = a3 + g3 * ((k3 * (\text{tTA} * (1 - (\text{Tc} / (kmt + \text{Tc})))^{\wedge}n2)) / (kd2^{\wedge}n2 + (\text{tTA} * (1 - (\text{Tc} / (kmt + \text{Tc})))^{\wedge}n2))) - kdm3 * \text{Galm};$$

$$\text{GFPm}' = g4 * ((k3 * (\text{tTA} * (1 - (\text{Tc} / (kmt + \text{Tc})))^{\wedge}n2)) / (kd2^{\wedge}n2 + (\text{tTA} * (1 - (\text{Tc} / (kmt + \text{Tc})))^{\wedge}n2))) - kdm4 * \text{GFPm};$$

$$\text{Lac}' = k4 * \text{Lacm} - kdp1 * \text{Lac};$$

$$\text{tTA}' = k4 * \text{tTAm} - kdp2 * \text{tTA};$$

$$\text{Gal}' = k4 * \text{Galm} - kdp3 * \text{Gal};$$

$$\text{GFP}' = k4 * \text{GFPm} - kdp4 * \text{GFP};$$

A.3.4. Bifurcation diagrams with respect to GFP concentration for parameters listed in tables 3.1, 3.2 and 3.3.

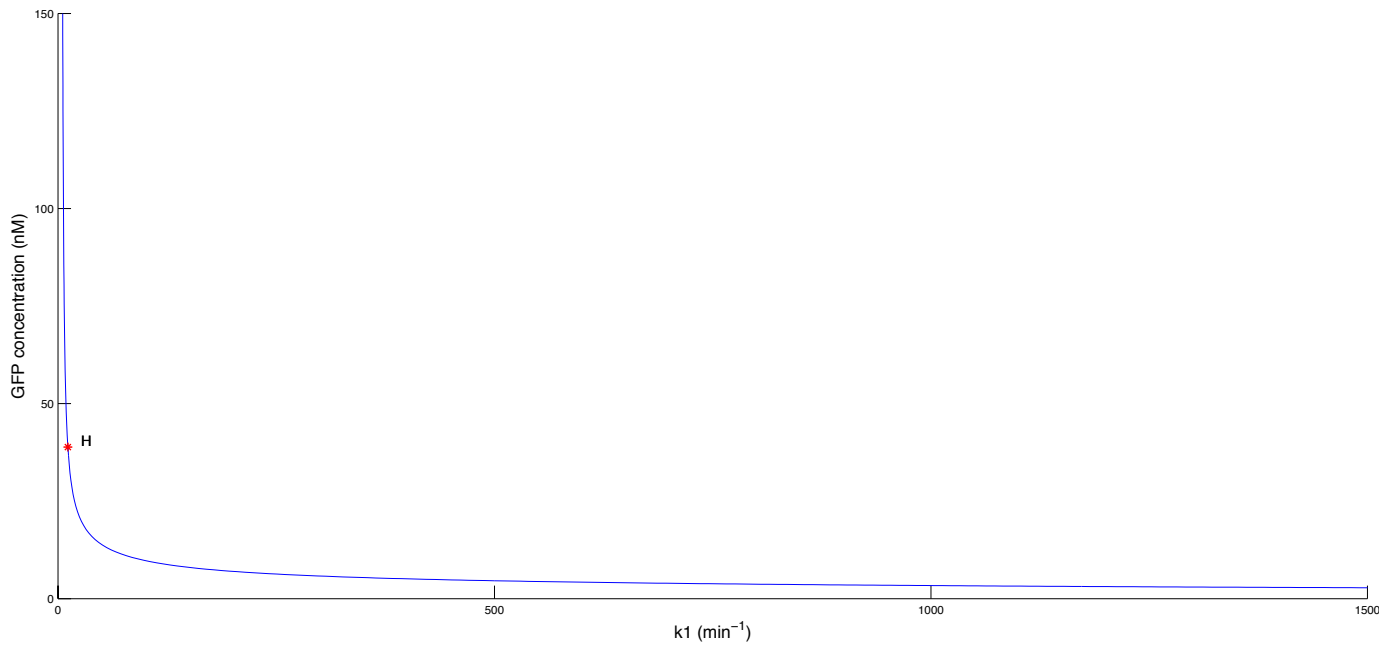


Figure A.3.1. Bifurcation diagram of parameter k_1 with respect to GFP concentration. Red dot(s) indicate Hopf points.

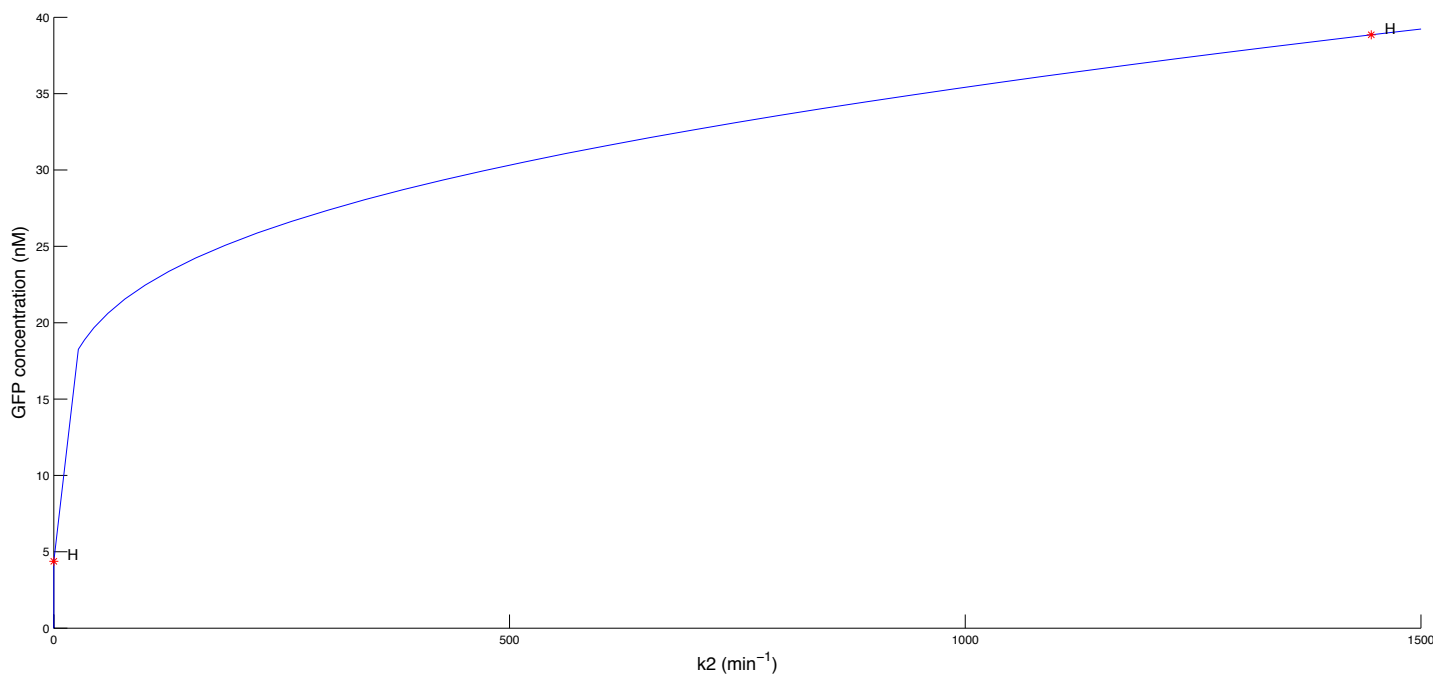


Figure A.3.2. Bifurcation diagram of parameter k_2 with respect to GFP concentration. Red dot(s) indicate Hopf points.

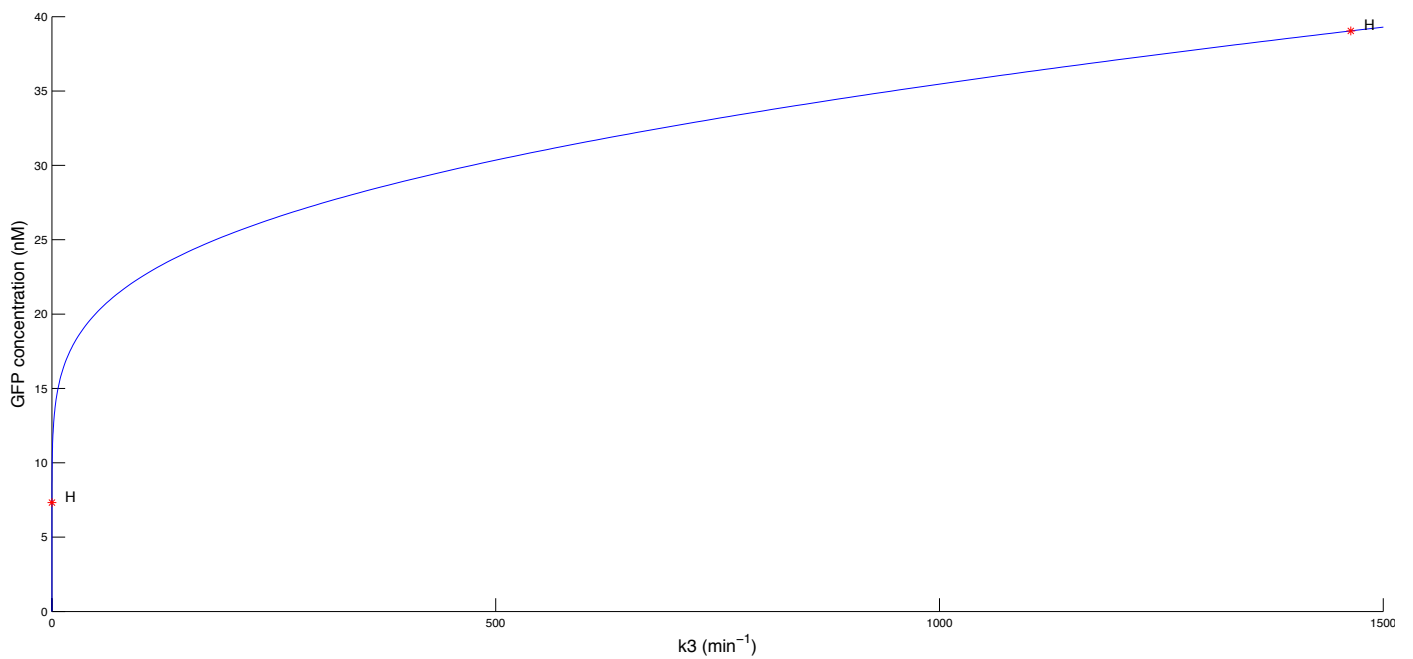


Figure A.3.3. Bifurcation diagram of parameter k_3 with respect to GFP concentration. Red dot(s) indicate Hopf points.

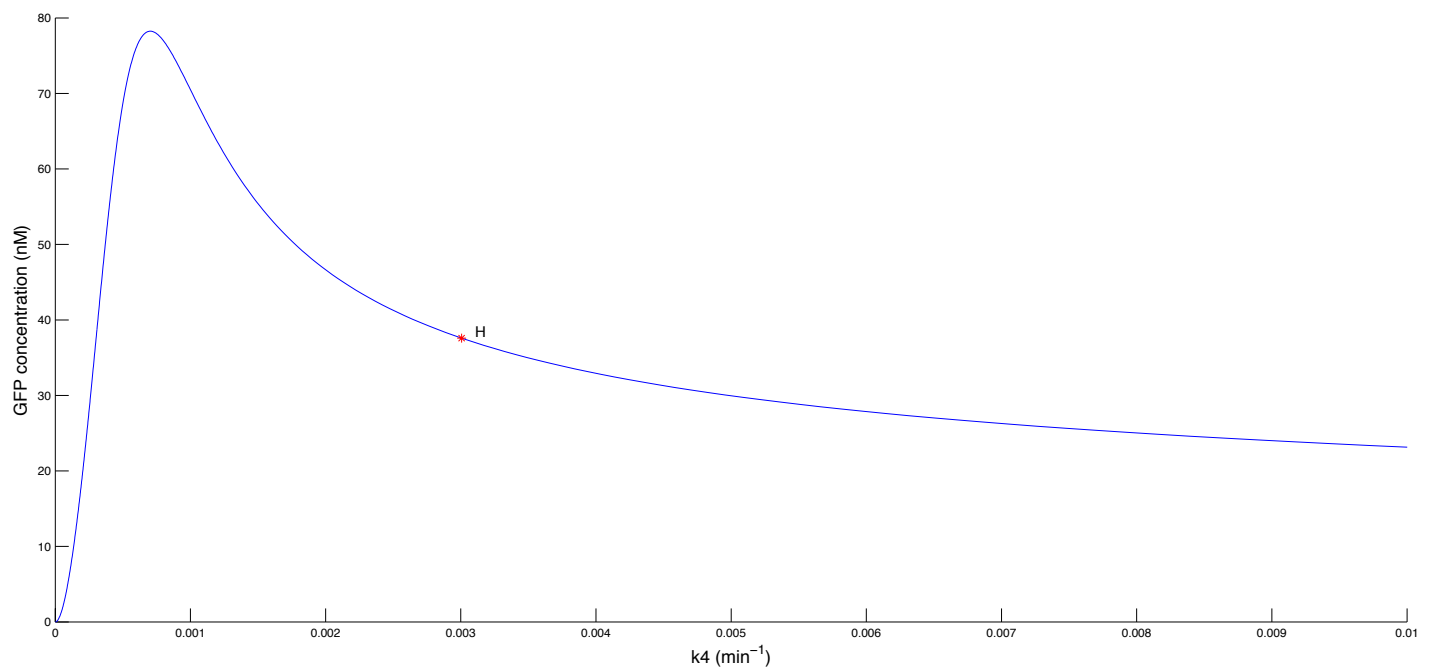


Figure A.3.4. Bifurcation diagram of parameter k_4 with respect to GFP concentration. Red dot(s) indicate Hopf points.

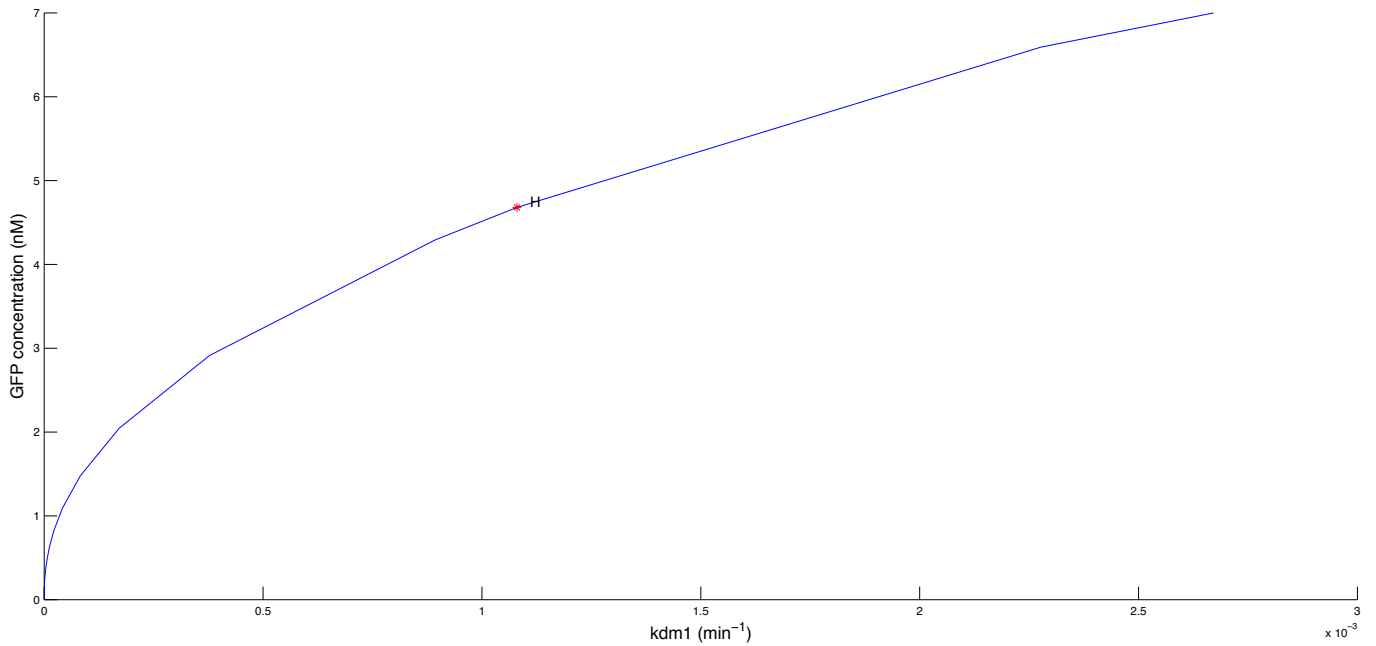


Figure A.3.5. Bifurcation diagram of parameter k_{dm1} with respect to GFP concentration. Red dot(s) indicate Hopf points.

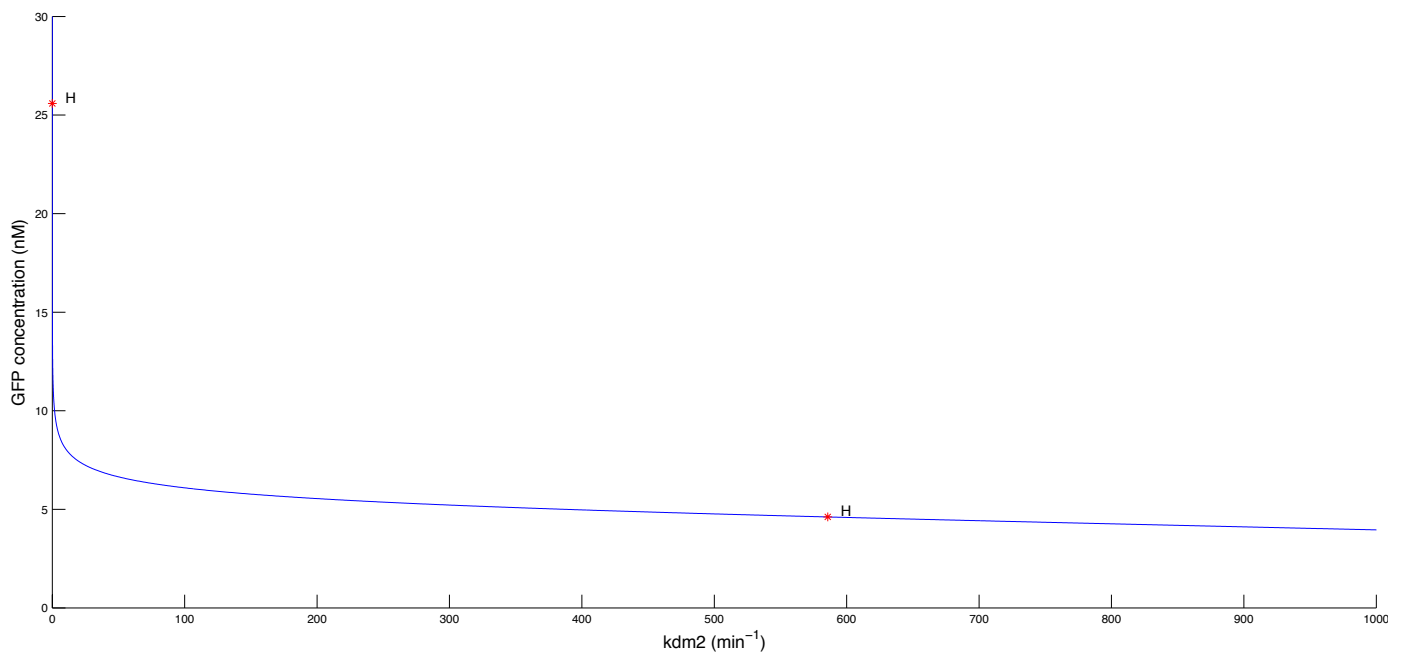


Figure A.3.6. Bifurcation diagram of parameter k_{dm2} with respect to GFP concentration. Red dot(s) indicate Hopf points.

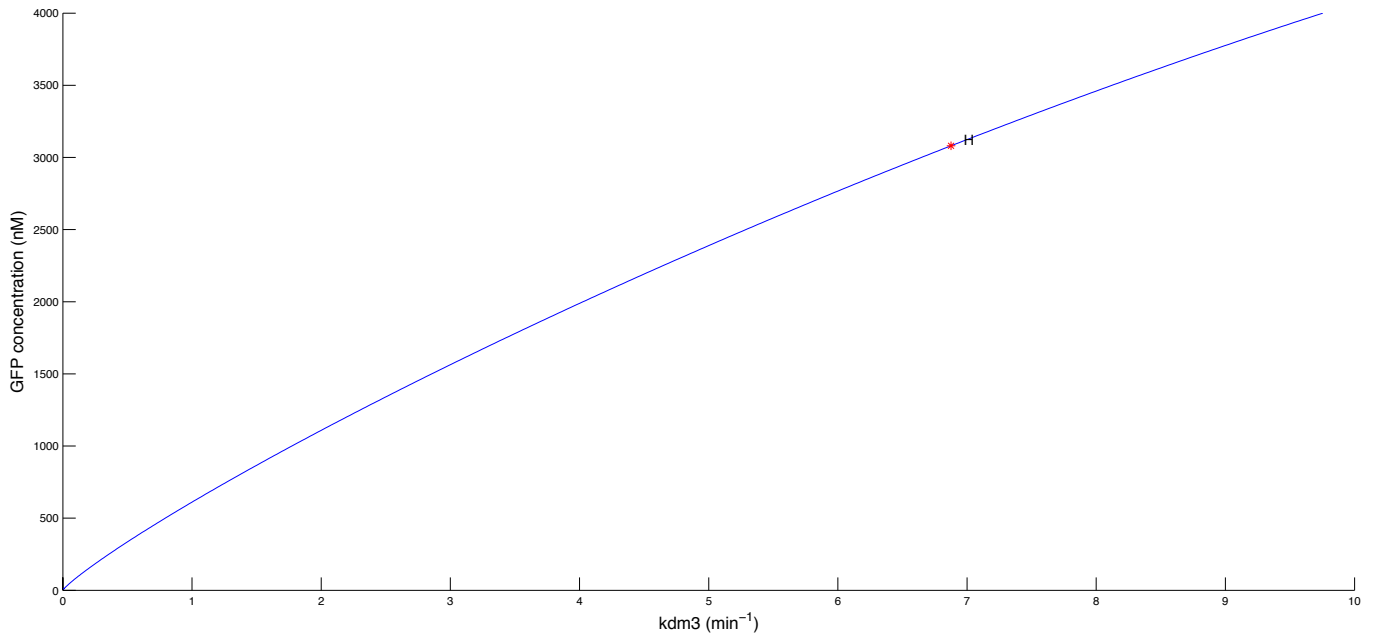


Figure A.3.7. Bifurcation diagram of parameter kdm3 with respect to GFP concentration. Red dot(s) indicate Hopf points.

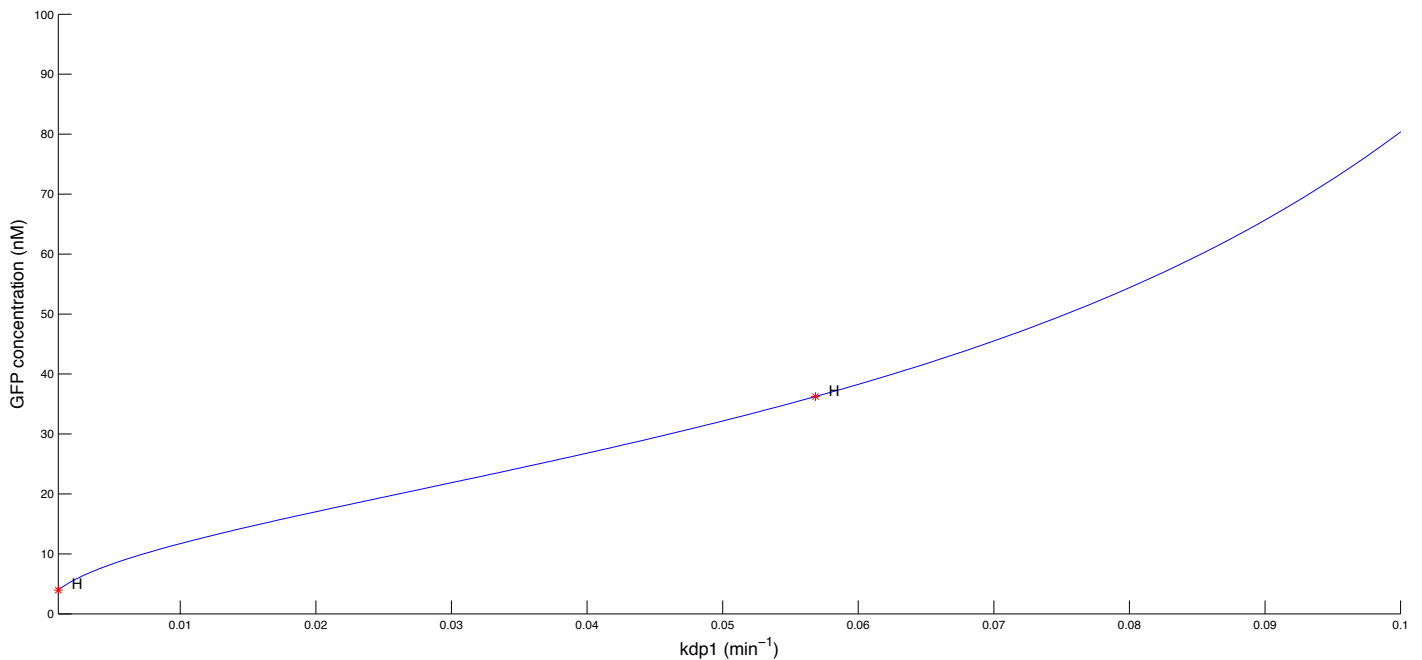


Figure A.3.8. Bifurcation diagram of parameter kdp1 with respect to GFP concentration. Red dot(s) indicate Hopf points.

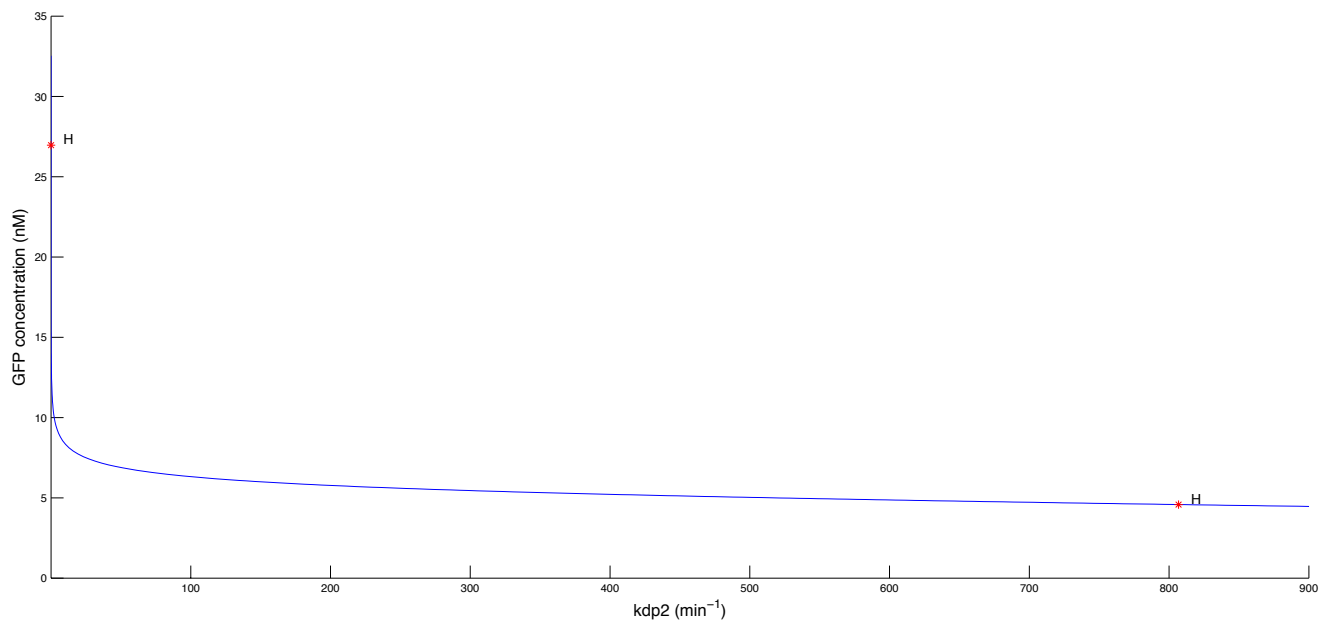


Figure A.3.9. Bifurcation diagram of parameter $kdp2$ with respect to GFP concentration. Red dot(s) indicate Hopf points.

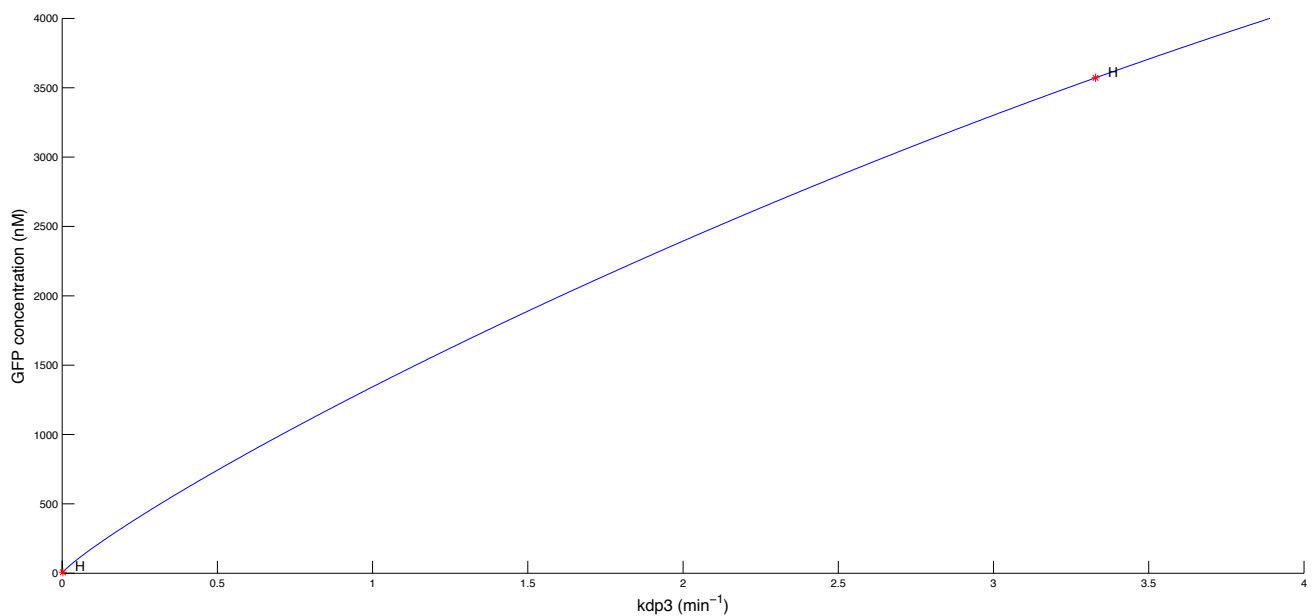


Figure A.3.10. Bifurcation diagram of parameter $kdp3$ with respect to GFP concentration. Red dot(s) indicate Hopf points.

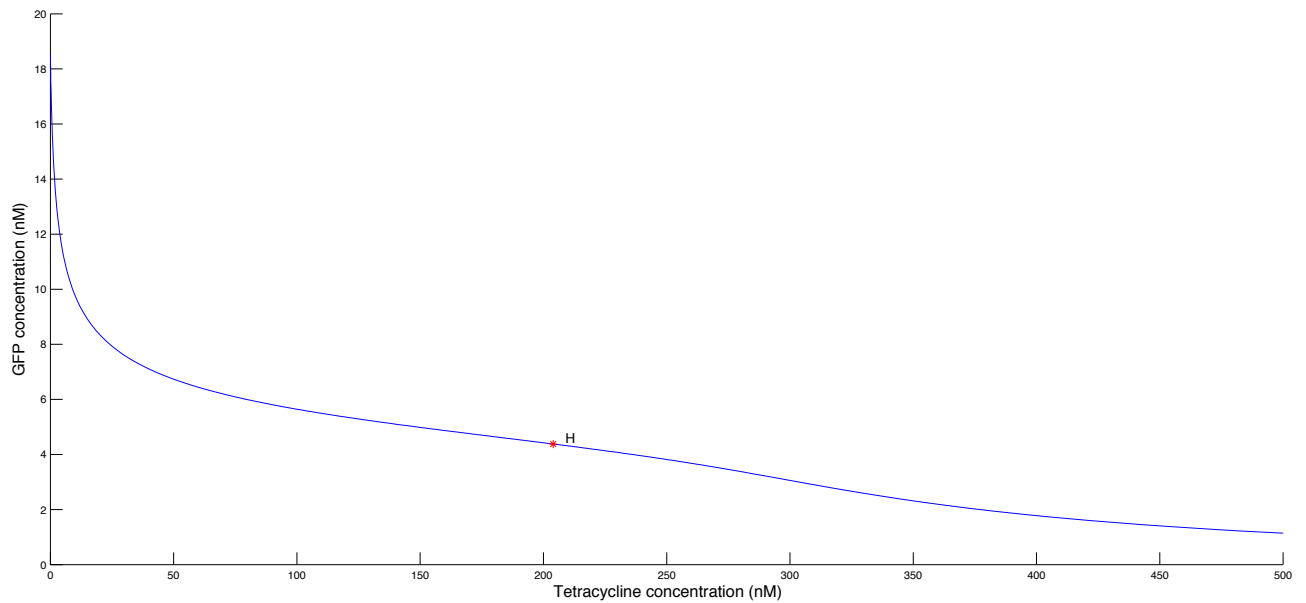


Figure A.3.11. Bifurcation diagram of tetracycline concentration with respect to GFP concentration. Red dot(s) indicate Hopf points.

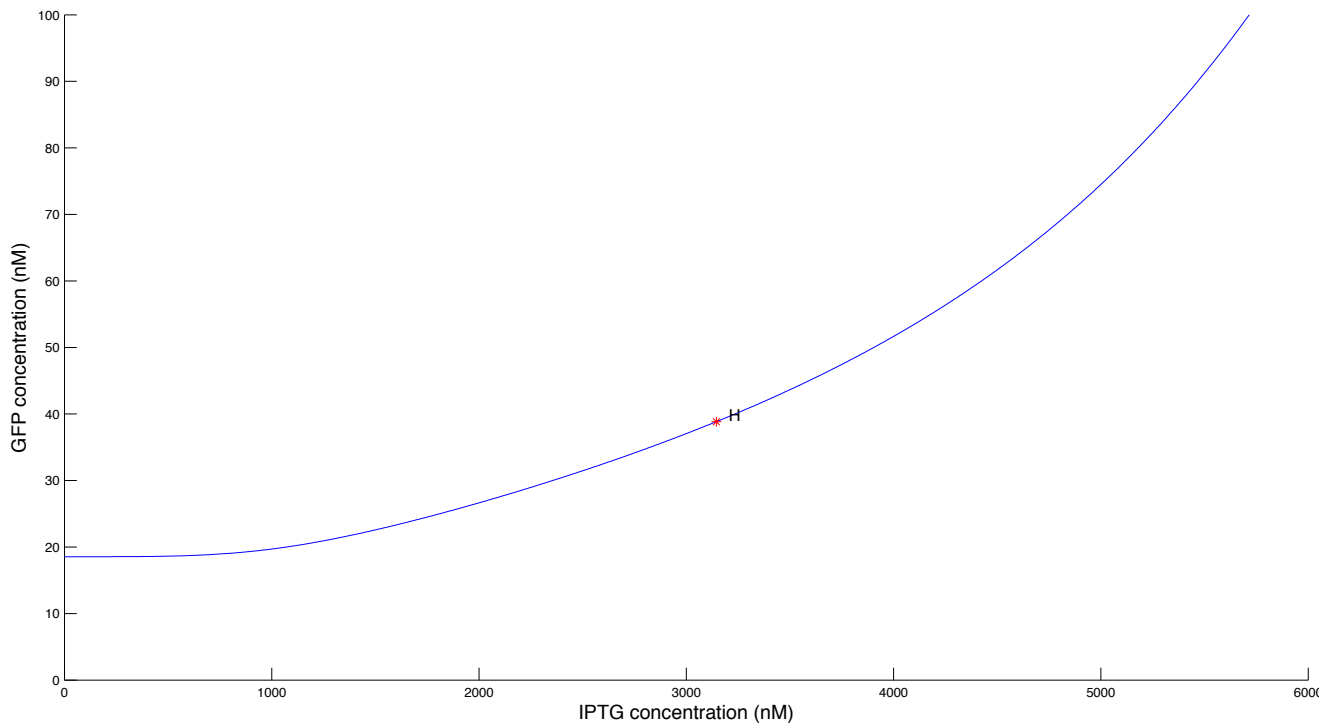


Figure A.3.12. Bifurcation diagram of IPTG concentration with respect to GFP concentration. Red dot(s) indicate Hopf points.

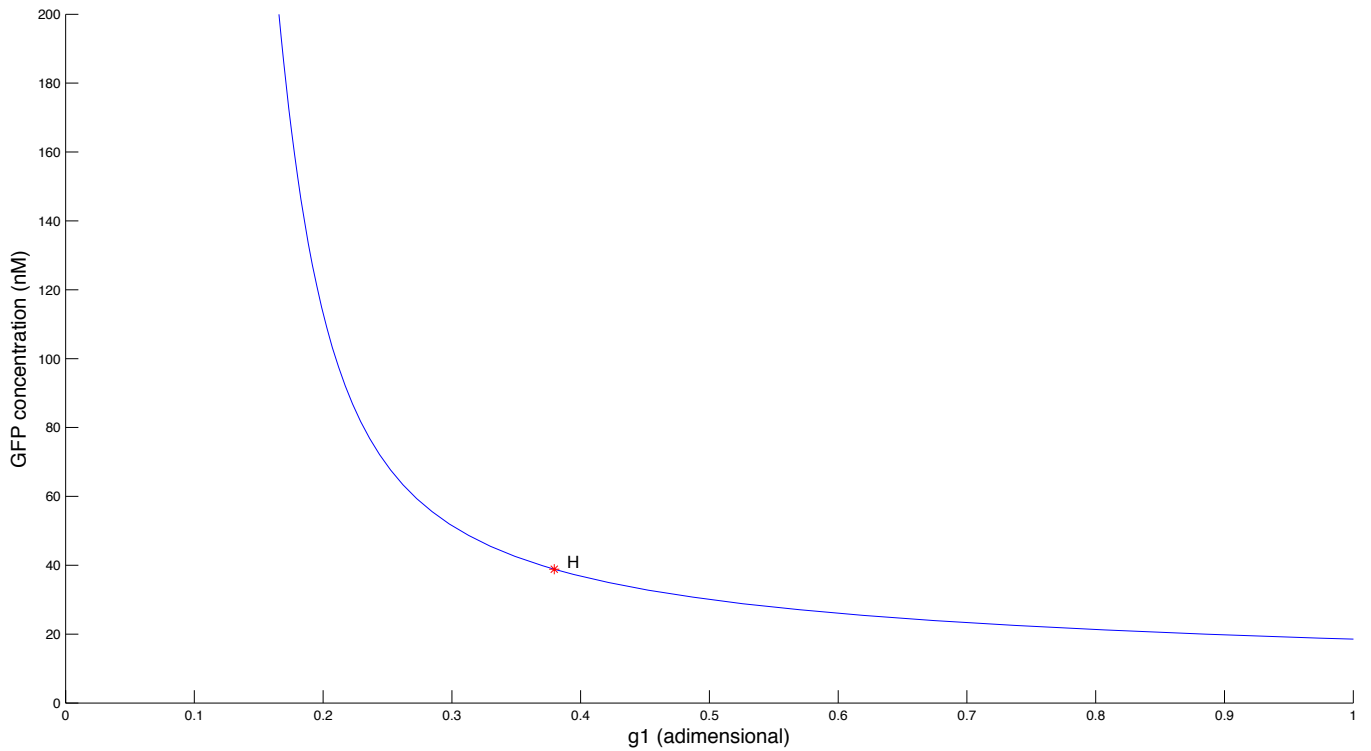


Figure A.3.13. Bifurcation diagram of parameter g_1 with respect to GFP concentration. Red dot(s) indicate Hopf points.

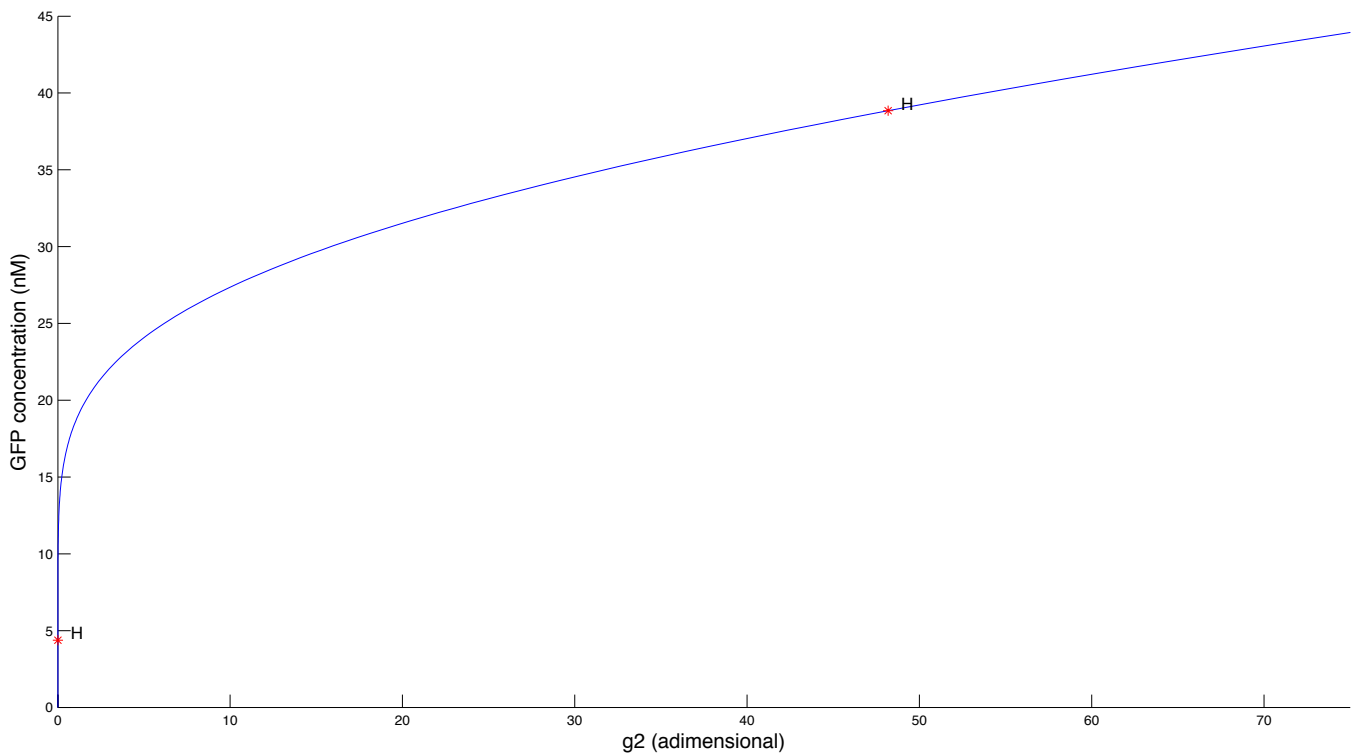


Figure A.3.14. Bifurcation diagram of parameter g_2 with respect to GFP concentration. Red dot(s) indicate Hopf points.

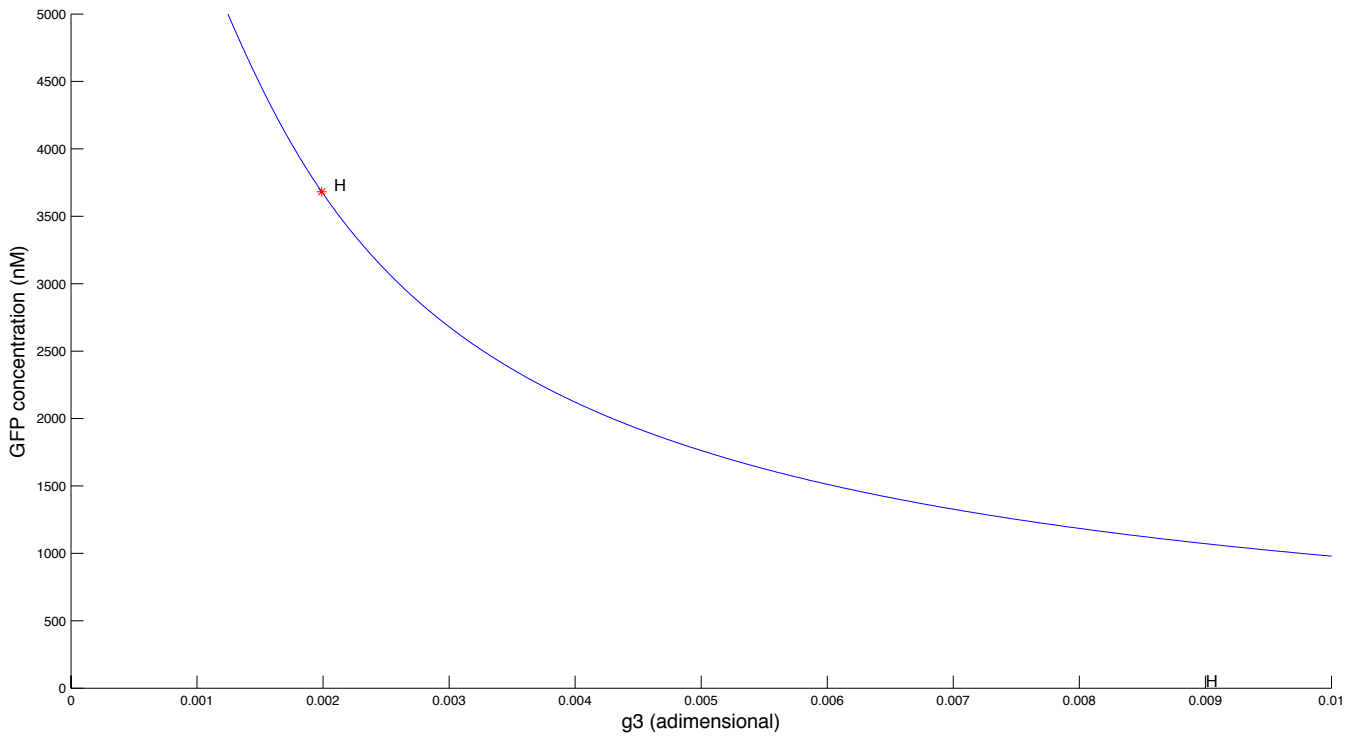


Figure A.3.15. Bifurcation diagram of parameter g_3 with respect to GFP concentration, in the range of $0 < g_3 < 0.01$. Red dot(s) indicate Hopf points.

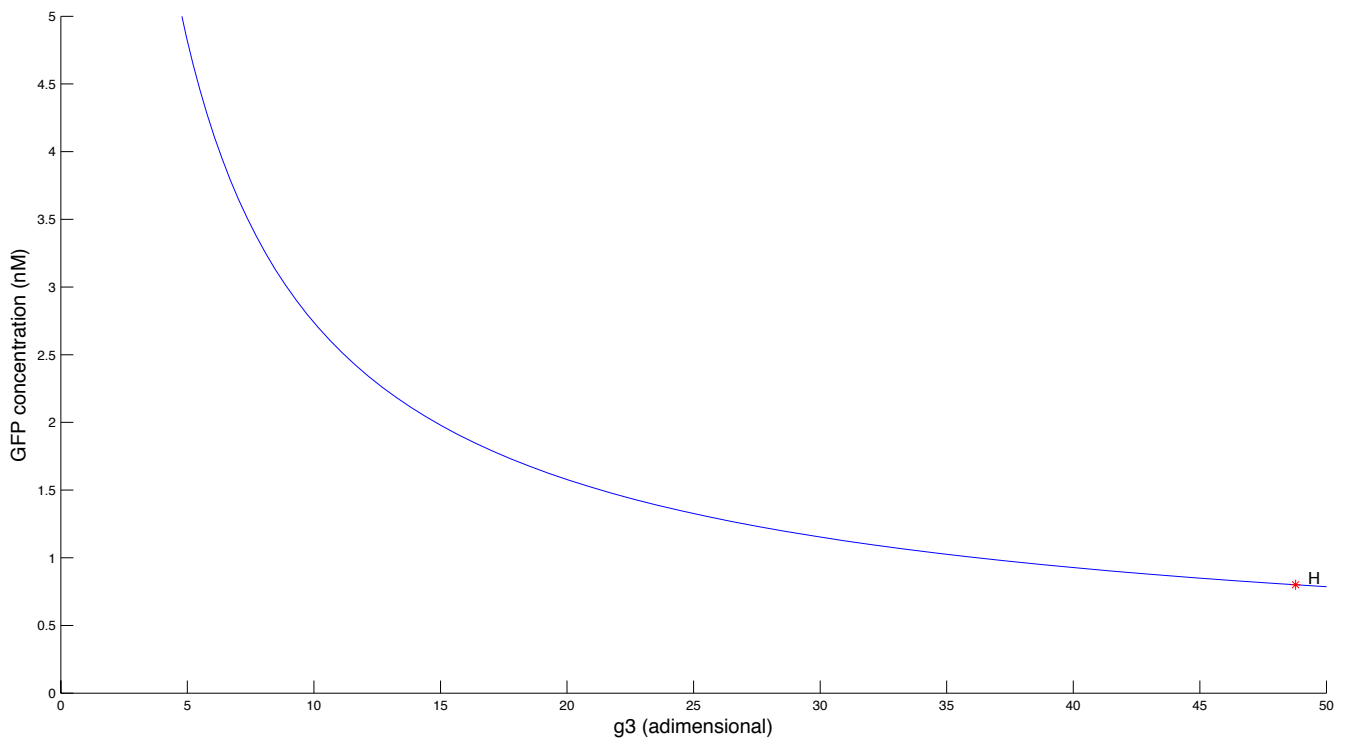


Figure A.3.16. Bifurcation diagram of parameter g_3 with respect to GFP concentration, in the range of $0 < g_3 < 50$. Red dot(s) indicate Hopf points.

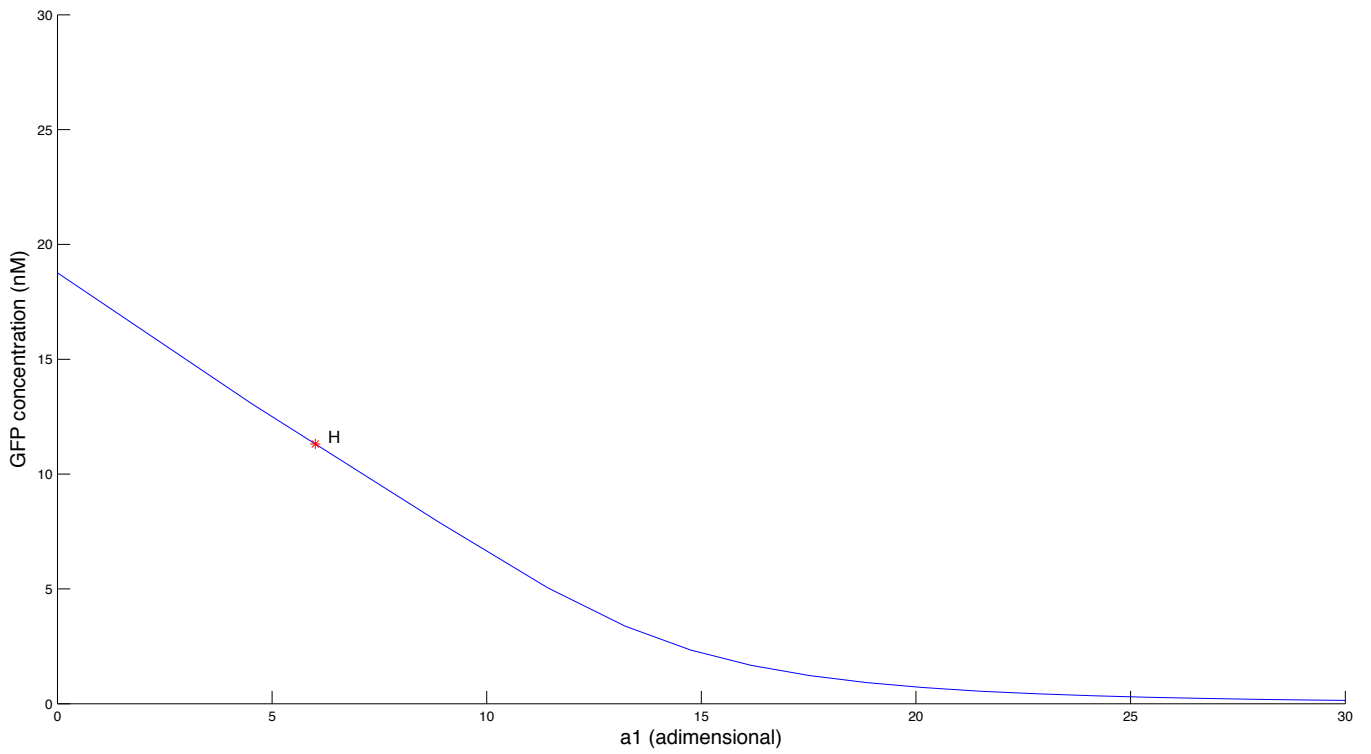


Figure A.3.17. Bifurcation diagram of parameter a_1 with respect to GFP concentration. Red dot(s) indicate Hopf points.

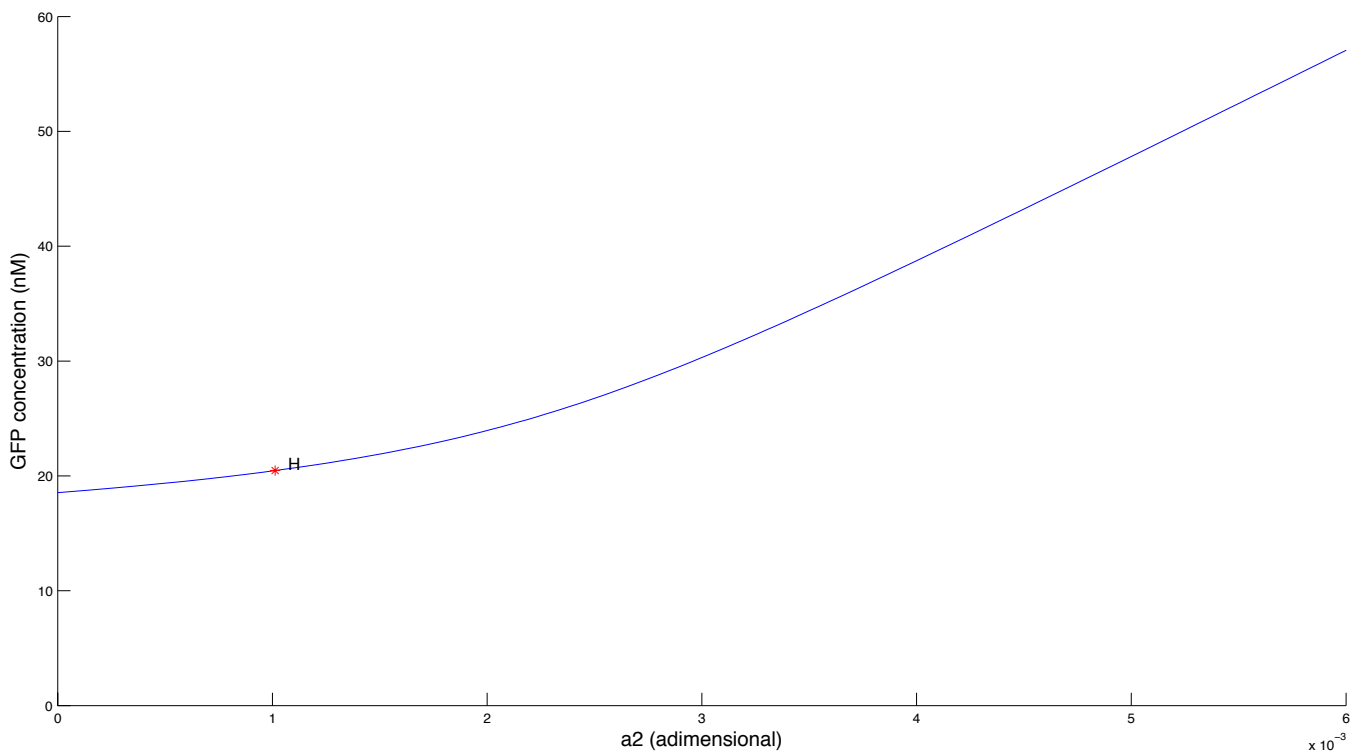


Figure A.3.18. Bifurcation diagram of parameter a_2 with respect to GFP concentration. Red dot(s) indicate Hopf points.

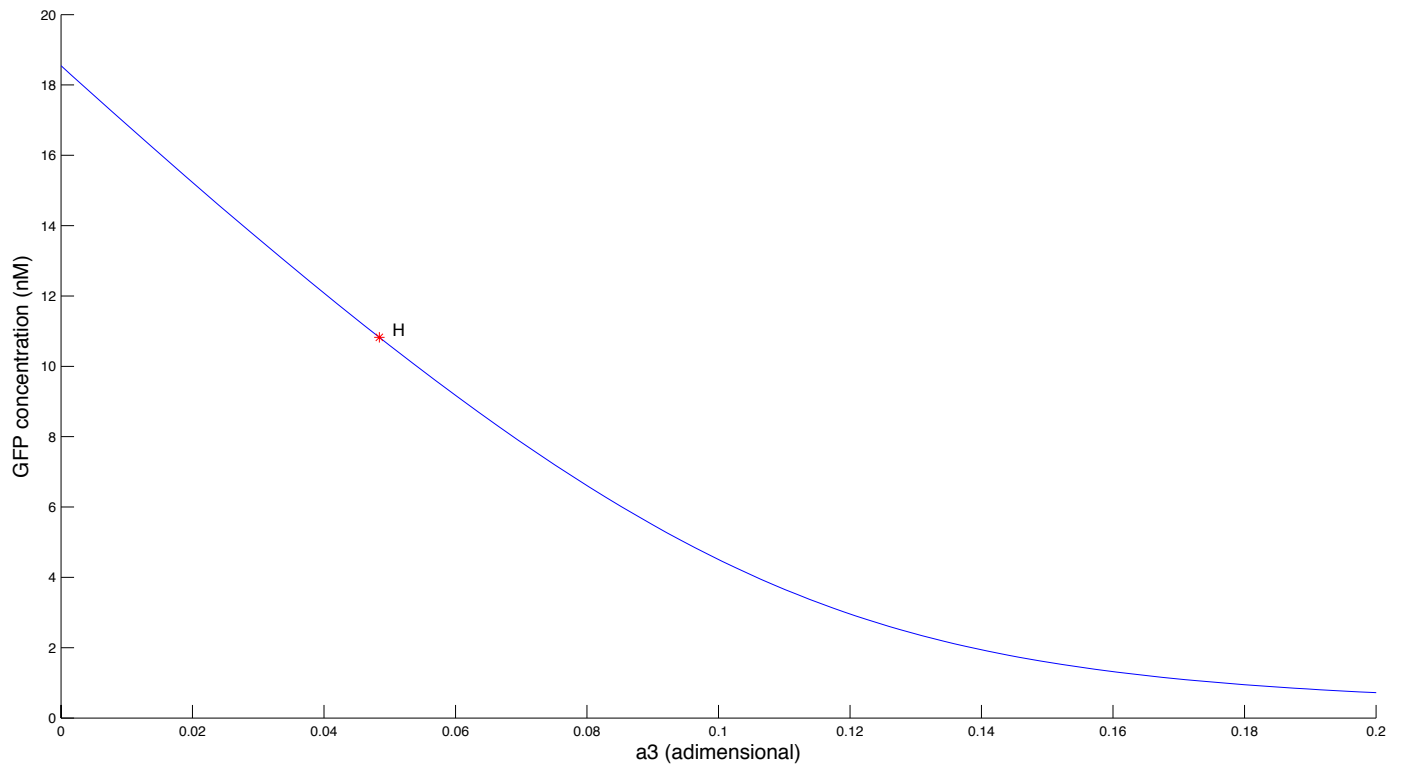


Figure A.3.19. Bifurcation diagram of parameter a_3 with respect to GFP concentration. Red dot(s) indicate Hopf points.

A.3.5. Hopf points and continuation settings

Hopf points are presented and structured as H (*LacI_mRNA*, *tTA_mRNA*, *Gal4_mRNA*, *GFP_mRNA*, *LacI*, *tTA*, *Gal4*, *GFP*, "Parameter value at bifurcation point").

Continuation of k_1

$H = (427.349369 \ 0.236838 \ 11.265179 \ 11.265179 \ 369.999453 \ 0.205055 \ 32.652692 \ 38.845444 \ 11.387196)$

Simulated from $k_1 = 0.000001$ to $k_1 = 20645$

Continuation Data

InitStepSize 0.01

MinStepSize 0.01

Max StepSize 10

Corrector Data

MaxNewton Iters 3

Max Corr Iters 10

Max Test Iters 10

Var Tolerance 1e-6

Fun Tolerance 1e-6

Test Tolerance 1e-5

Adapt 3

Stop Data

MaxNumPoints 300

ClosedCurve 50

Continuation of k_2

$H = (39.875097 \ 0.026552 \ 1.270267 \ 1.270267 \ 34.523893 \ 0.022989 \ 3.681935 \ 4.380233 \ 0.000714)$

$H = (1125.949268 \ 0.236863 \ 11.266336 \ 11.266336 \ 974.847851 \ 0.205076 \ 32.656045 \ 38.849433 \ 1445.609784)$

Simulated from $k_2 = 0.000001$ to $k_2 = 14452.479$

Continuation Data

InitStepSize 0.01

MinStepSize 0.01

Max StepSize 10

Corrector Data

MaxNewton Iters 3
 Max Corr Iters 10
 Max Test Iters 10
 Var Tolerance 1e-6
 Fun Tolerance 1e-6
 Test Tolerance 1e-5
 Adapt 3

Stop Data

MaxNumPoints 300
 ClosedCurve 50

Continuation of k_3

H = (107.093378 58.050418 2.124303 2.124303 92.721539 50.260102 6.157400
 7.325182 0.059681)

H = (1129.819544 0.004848 11.321765 11.321765 978.198740 0.004198 32.816711
 39.040570 1463.426501)

Simulated from $k_3 = 0.000001$ to $k_3 = 2639$

Continuation Data

InitStepSize 0.01
 MinStepSize 0.01
 Max StepSize 10

Corrector Data

MaxNewton Iters 3
 Max Corr Iters 10
 Max Test Iters 10
 Var Tolerance 1e-6
 Fun Tolerance 1e-6
 Test Tolerance 1e-5
 Adapt 3

Stop Data

MaxNumPoints 300
 ClosedCurve 50

Continuation of k_4

H = (1099.683467 10.517026 72.521405 72.521405 143.118263 1.368738
 31.597772 37.590453 0.003006)

Simulated from $k_4 = 0.000001$ to $k_4 = 5.2765291$

Continuation Data

InitStepSize 0.01

MinStepSize 0.01

Max StepSize 10

Corrector Data

MaxNewton Iters 3

Max Corr Iters 10

Max Test Iters 10

Var Tolerance 1e-6

Fun Tolerance 1e-6

Test Tolerance 1e-5

Adapt 3

Stop Data

MaxNumPoints 300

ClosedCurve 50

Continuation of k_{dm1}

H = (726.445078 0.028368 1.357069 1.357069 628.956777 0.024561 3.933533
4.679548 0.001080)

Simulated from $k_{dm1} = 0.000001$ to $k_{dm1} = 7249$

Continuation Data

InitStepSize 0.1

MinStepSize 0.1

Max StepSize 10

Corrector Data

MaxNewton Iters 3

Max Corr Iters 10

Max Test Iters 10

Var Tolerance 1e-6

Fun Tolerance 1e-6

Test Tolerance 1e-5

Adapt 3

Stop Data

MaxNumPoints 300

ClosedCurve 50

Continuation of k_{dm2}

H = (772.426788 0.155665 7.420691 7.420691 668.767782 0.134774 21.509248
25.588588 0.002466)

H = (44.118199 0.027965 1.337820 1.337820 38.197575 0.024212 3.877739
4.613172 585.706430)

Simulated from $k_{dm2} = 0.000001$ to $k_{dm2} = 1058$

Continuation Data

InitStepSize 0.1

MinStepSize 0.1

Max StepSize 1

Corrector Data

MaxNewton Iters 3

Max Corr Iters 10

Max Test Iters 10

Var Tolerance 1e-6

Fun Tolerance 1e-6

Test Tolerance 1e-5

Adapt 3

Stop Data

MaxNumPoints 300

ClosedCurve 50

Continuation of k_{dm3}

H = (119.027292 38.486602 2.247797 893.348857 103.053932 33.321733 6.515353
3080.513299 6.875593)

Simulated from $k_{dm3} = 0.000001$ to $k_{dm3} = 828$

Continuation Data

InitStepSize 0.01

MinStepSize 0.01

Max StepSize 10

Corrector Data

MaxNewton Iters 3

Max Corr Iters 10
 Max Test Iters 10
 Var Tolerance 1e-6
 Fun Tolerance 1e-6
 Test Tolerance 1e-5
 Adapt 3

Stop Data

MaxNumPoints 300
 ClosedCurve 50

Continuation of k_{dp1}

H = (1070.132513 0.220901 10.511735 10.511735 376.500432 0.191256 30.468796
 36.247361 0.056846)

H = (32.903830 0.024068 1.151532 1.151532 655.338174 0.020838 3.337773
 3.970799 0.001004)

Simulated from $k_{dp1} = 9.5e-9$ to $k_{dp1} = 2000$

Continuation Data

InitStepSize 0.1
 MinStepSize 0.1
 Max StepSize 10

Corrector Data

MaxNewton Iters 3
 Max Corr Iters 10
 Max Test Iters 10
 Var Tolerance 1e-6
 Fun Tolerance 1e-6
 Test Tolerance 1e-5
 Adapt 3

Stop Data

MaxNumPoints 300
 ClosedCurve 50

Continuation of k_{dp2}

H = (817.474155 0.017691 7.819344 7.819344 707.769831 0.142048 22.664767
 26.963257 0.002491)

H = (43.525661 969.967696 1.328573 1.328573 37.684555 0.024045 3.850936

4.581286 806.804175)

Simulated from $k_{dp2} = 0.000001$ to $k_{dp2} = 3104$

Continuation Data

InitStepSize 0.1
 MinStepSize 0.1
 Max StepSize 10

Corrector Data

MaxNewton Iters 3
 Max Corr Iters 10
 Max Test Iters 10
 Var Tolerance 1e-6
 Fun Tolerance 1e-6
 Test Tolerance 1e-5
 Adapt 3

Stop Data

MaxNumPoints 300
 ClosedCurve 50

Continuation of k_{dp3}

H = (109.255271 53.728494 1035.814675 1035.814675 94.593308 46.518177
 6.223375 3571.774742 3.328788)

H = (650.909504 0.044010 2.104453 2.104453 563.558012 0.038104 18.604513
 7.256736 0.002262)

Simulated from $k_{dp3} = 3.6e-7$ to $k_{dp3} = 3000$

Continuation Data

InitStepSize 0.1
 MinStepSize 0.1
 Max StepSize 10

Corrector Data

MaxNewton Iters 3
 Max Corr Iters 10
 Max Test Iters 10
 Var Tolerance 1e-6
 Fun Tolerance 1e-6
 Test Tolerance 1e-5
 Adapt 3

Stop Data

MaxNumPoints 300

ClosedCurve 50

Continuation of Tetracycline concentration

H = (39.875109 1115.228350 1.270268 1.270268 34.523904 965.565671 3.681935
4.380233 203.943297)

Simulated from $T = 0.001$ to $T = 5000$

Continuation Data

InitStepSize 0.1

MinStepSize 0.1

Max StepSize 10

Corrector Data

MaxNewton Iters 3

Max Corr Iters 10

Max Test Iters 10

Var Tolerance 1e-6

Fun Tolerance 1e-6

Test Tolerance 1e-5

Adapt 3

Stop Data

MaxNumPoints 300

ClosedCurve 50

Continuation of IPTG concentration

H = (1125.868317 0.236838 11.265181 11.265181 974.777764 0.205055 32.652698
38.845452 3144.910751)

Simulated from $I = 0.0001$ to $I = 8000$

Continuation Data

InitStepSize 0.01

MinStepSize 0.01

Max StepSize 10

Corrector Data

MaxNewton Iters 3

Max Corr Iters 10

Max Test Iters 10

Var Tolerance 1e-6
 Fun Tolerance 1e-6
 Test Tolerance 1e-5
 Adapt 3

Stop Data

MaxNumPoints 300
 ClosedCurve 50

Continuation of g_1

H = (427.349406 0.236838 11.265175 11.265175 369.999486 0.205055 32.652681
 38.845431 0.379573)

Simulated from $g_1 = 0.0001$ to $g_1 = 2564$

Continuation Data

InitStepSize 0.1
 MinStepSize 0.1
 Max StepSize 10

Corrector Data

MaxNewton Iters 3
 Max Corr Iters 10
 Max Test Iters 10
 Var Tolerance 1e-6
 Fun Tolerance 1e-6
 Test Tolerance 1e-5
 Adapt 3

Stop Data

MaxNumPoints 300
 ClosedCurve 50

Continuation of g_2

H = (39.875112 0.026552 1.270268 1.270268 34.523907 0.022989 3.681935
 4.380233 0.000024)

H = (1125.949267 0.236863 11.266336 11.266336 974.847850 0.205076 32.656045
 38.849433 48.186993)

Simulated from $g_2 = 0.000001$ to $g_2 = 3953$

Continuation Data

InitStepSize 0.1
 MinStepSize 0.1
 Max StepSize 10

Corrector Data
 MaxNewton Iters 3
 Max Corr Iters 10
 Max Test Iters 10
 Var Tolerance 1e-6
 Fun Tolerance 1e-6
 Test Tolerance 1e-5
 Adapt 3

Stop Data
 MaxNumPoints 300
 ClosedCurve 50

Continuation of g_3

H = (107.093381 58.050411 2.124303 1067.828145 92.721542 50.260096 6.157400
 3682.166017 0.001989)

H = (1129.819558 0.004848 11.321766 0.232094 978.198751 0.004198 32.816712
 0.800325 48.780886)

Simulated from $g_3 = 0.000001$ to $g_3 = 2822$

Continuation Data
 InitStepSize 0.1
 MinStepSize 0.1
 Max StepSize 10

Corrector Data
 MaxNewton Iters 3
 Max Corr Iters 10
 Max Test Iters 10
 Var Tolerance 1e-6
 Fun Tolerance 1e-6
 Test Tolerance 1e-5
 Adapt 3

Stop Data
 MaxNumPoints 300
 ClosedCurve 50

Continuation of a_1

H = (582.443299 0.068645 3.280243 3.280243 504.279912 0.059433 9.507952
11.311184 6.006596)

Simulated from $a_1 = 0.000001$ to $a_1 = 20$

Continuation Data

InitStepSize 0.1
MinStepSize 0.1
Max StepSize 10

Corrector Data

MaxNewton Iters 3
Max Corr Iters 10
Max Test Iters 10
Var Tolerance 1e-6
Fun Tolerance 1e-6
Test Tolerance 1e-5
Adapt 3

Stop Data

MaxNumPoints 300
ClosedCurve 50

Continuation of a_2

H = (588.559033 0.124392 5.934980 5.934980 509.574921 0.107698 17.202841
20.465449 0.001013)

Simulated from $a_2 = 0.000001$ to $a_2 = 20$

Continuation Data

InitStepSize 0.1
MinStepSize 0.1
Max StepSize 10

Corrector Data

MaxNewton Iters 3
Max Corr Iters 10
Max Test Iters 10
Var Tolerance 1e-6
Fun Tolerance 1e-6
Test Tolerance 1e-5
Adapt 3

Stop Data

MaxNumPoints 300

ClosedCurve 50

Continuation of a_3

H = (588.920232 0.065675 5.937737 3.138565 509.887646 0.056862 17.210833
10.822637 0.048426)

Simulated from $a_3 = 0.000001$ to $a_3 = 20$

Continuation Data

InitStepSize 0.1

MinStepSize 0.1

Max StepSize 10

Corrector Data

MaxNewton Iters 3

Max Corr Iters 10

Max Test Iters 10

Var Tolerance 1e-6

Fun Tolerance 1e-6

Test Tolerance 1e-5

Adapt 3

Stop Data

MaxNumPoints 300

ClosedCurve 50

Appendix A.4

A.4.1. Plasmid DNA sequences

pCMVLacI DNA sequence:

f1 origin of replication

CMV enhancer

CMV promoter

lacI ORF

SV40 Nuclear Localization Signal

Small t-intron

SV40 poly-adenylation signal

HSV TK promoter

HygR ORF

HSV TK poly-adenylation signal

ColE1 / pUC origin of replication

AmpR ORF / *AmpR* promoter

```

CTGACGCGCCCTGTAGCGGCGCATTAAAGCGCGCGGGTGTGGTGGTTACGCGCAGCGTGA
CCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTCGCTTCTTCCCTTCTTCTCGCCAC
GTTGCGCCGGCTTCCCGTCAAGCTCTAAATCGGGGGCTCCCTTAGGGTTCGGATTTAGTG
CTTACGGCACCTCGACCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATC
GCCCTGATAGACGGTTTTTCGCCCTTGGAGTCCACGTTCTTAAATAGTGGACTCTT
GTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTGATTTATAAGGGATTTT
GCCGATTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTA
ACAAAATATTAACGCTTACAATTTACGCGTATAGATCTAGCTTCGTGTCAAGGACGGTGACT
GCAGTGAATAATAAAATGTGTGTTTGTCCGAAATACGCGTTTTGAGATTTCTGTGCGCCGACT
AAATTCATGTCGCGCGATAGTGGTGTATCGCCGATAGAGATGGCGATATTGGAAATCGA
TATTTGAAAATATGGCATATTGAAAATGTGCGCCGATGTGAGTTTCTGTGTAAGTATATCGC
CATTTTTCCAAAAGTGATTTTTGGGCATACGCGATATCTGGCGATACGGCTTATATCGTTTTAC
GGGGGATGGCGATAGACGACTTTGGCGACTTGGGCGATTCTGTGTGTCGAAATATCGCA
GTTTCGATATAGGTGACAGACGATATGAGGCTATATCGCCGATAGAGGCGACATGAAGCTG
GCACATGGCCAATGCATATCGATCTATACATTGAATCAATATTGGCAATTAGCCATATTAGT
CATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCA
TAATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGAC

```

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGACTTCCGCGT
TACATAACTTACGGTAAATGGCCCCGCTCGTGACCGCCCAACGACCCCCGCCATTGACGTC
AATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTCAATGGGTGG
AGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCGGGCC
CCCTATTGACGTCAATGACGGTAAATGGCCCCGCTGGCATTATGCCCAGTACATGACCTTAC
GGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG
TTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTCCAAGTCTCC
ACCCATTGACGTCAATGGGAGTTTGTGGTGGCACCAAAATCAACGGGACTTCCAAAATGT
CGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTAT
ATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTG
ACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGGCCGGGAACGGTGCATTGGAA
CGCGGATTCCCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGACTCTATAGGCACACCCCT
TTGGCTCTTATGCATGCTATACTGTTTTGGCTTGGGGCCTATACACCCCGCTCCTTATGCT
ATAGGTGATGGTATAGCTTAGCCTATAGGTGTGGGTTATTGACCATTATTGACCACTCCCCT
ATTGGTGACGATACTTCCATTAATAACATGGCTCTTGGCACAACCTATCTCTATT
GGCTTATGCCAATACTCTGTCCTCAGAGACTGACACGGACTCTGTATTTTACAGGATGGG
GTCCATTTATTATTTACAAATTCACATATAACAACGCCGTCCCCGTGCCCGCAGTTTTT
ATTAACATAGCGTGGGATCTCCACGGAATCTCGGGTACGTGTTCCGGACATGGGCTCTTC
TCCGGTAGCGGCGGAGCTTCCACATCCGAGCCCTGGTCCCATGCCTCCAGCGGCTCATGGT
CGCTCGGCAGCTCCTTGCTCCTAACAGTGGAGGCCAGACTTAGGCACAGCACAATGCCAC
CACCACAGTGTGCCGACAAGGCCGTGGCGGTAGGGTATGTGTCTGAAAATGAGCTCGG
AGATTGGGCTCGACCGTGACGCAGATGGAAGACTTAAGGCAGCGGCAGAAGAAGATGCA
GGCAGCTGAGTTGTTGATTCTGATAAGAGTCAGAGGTAACCTCCGTTGCGGTGCTGTTAA
CGGTGGAGGGCAGTGTAGTCTGAGCAGTACTCGTTGCTGCCGCGCGCCACCAGACATA
ATAGCTGACAGACTAACAGACTGTTCCCTTCCATGGGTCTTTCTGCAGTACCCTCCTTGAC
ACGAAGCTTGTACCGAGCTCGGATCTCCATTGCCATTACGGCTGCGCAACTGTTGGGAA
GGGCGATCGTCTAGAGGTACCTCCCACCATGAAACCAGTAACGTTATACGATGTGCGAGA
GTATGCCGGTGTCTTATCAGACCGTTTTCCGCGTGGTGAACCAGGCCAGCCACGTTTCTG
CGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATCCCAACCGCG
TGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGC
CCTGCACGCGCCGTGCAAATTGTCGCGGCGATTAATCTCGCGCCGATCAACTGGGTGCC
AGCGTGGTGGTGTGATGGTAGAACGAAGCGGCGTCAAGCCTGTAAAGCGGCGGTGCAC
AATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATGC
CATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTATTTCTTGATGTCTGACCAGAC
ACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATCTG
GTCGATTGGGTACCAGCAAATCGCGCTGTTAGCGGGCCATTAAGTTCTGTCTCGGCGC
GTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAA
CGGGAAGGCGACTGGAGTGCCATGTCCGTTTTCAACAAACCATGCAAATGCTGAATGAGG
GCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGC
CATTACCGAGTCCGGGCTGCGCGTTGGTGGGATATCTCGGTAGTGGGATACGACGATACC
GAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAACAGGATTTTCGCTGCTGGG
GCAAACCAGCGTGGACCGCTTGTGCAACTCTCAGGGCCAGGCGGTGAAGGGCAATCA
GCTGTTGCCCGTCTCACTGGTGAAGAAAACACCCTGGCGCCCAATACGCAAACCGCC
TCTCCCCGCGGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAA
GCGGGCAGAGCAGCCTGAGGCCTCCTAAGAAGAAGAGGAAGGTTGAGCGCAACGCAATT
AATGTAAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTCCGACCTGCAGC
CCCTTGATCTTTGTGAAGGAACCTTACTTCTGTGGTGTGACATAATTGGACAAACTACCTA

CAGAGATTTAAAGCTCTAAGGTAAATATAAAATTTTAAAGTGATAATGTGTTAAACTACTG
 ATTCTAATTGTTTGTGATTTTAGATCACAGTCCCAAGGCTCATTTACAGGCCCTCAGTCCTC
 ACAGTCTGTTTCATGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAA
 ACCTCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGT
 TTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTACAAATAAAGCA
 TTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAACTCATCAATGTATCTTATCATGTCTGGG
 CCCAAGCTTGGCACTGGGATCTGCGAACGCAAGACGTAGCCCAGCGCGTCCGGCCCCG
 AGATGCGCCGCGTGCCGTGCTGGAGATGGCGGACCGGATGGATATGTTCTGCCAAGGGTT
 GTTTTGCGCATTACAGTTCTCCGCAAGAATTGATTGGCTCCAATTCTGGAGTGGTGAATC
 CGTTAGCGAGGTGCCGCCCTGCTTCATCCCCGTGGCCCGTTGCTCGCGTTTGTGGCGGTGT
 CCCCAGAAATATATTTGCATGTCTTTAGTTCTATGATGACACAAACCCCGCCAGCGTCT
 TTGTCATTGGCGAATTGCAACACGCAGATGCAGTCGGGGCGGCGCGGTCCGAGGTCCACTT
 CGCATATTAAGGTGACGCGTGTGGCCTCGAACACCGAGCGACCCTGCAGCGACCCGCTTAA
 CAGCGTCAACAGCGTGCCGAGATCCCGGGGGCAATGAGATATGAAAAAGCCTGAACTC
 ACCGCGACGTCTGTGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGC
 AGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGT
 CCTGCGGGTAAATAGCTGCGCCGATGGTTTACAAAGATCGTTATGTTTATCGGCACTTTG
 CATCGGCCGCGTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGAC
 CTATTGCATCTCCCGCCGTGCACAGGGTGTACGTTGCAAGACCTGCCTGAAACCGAACTGC
 CCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCA
 GACGAGCGGGTTCGGCCATTTCGACCCGCAAGGAATCGGTCAATACACTACATGGCGTGAT
 TTCATATGCGGATTGCTGATCCCCATGTGTATCACTGGCAAACGTGATGGACGACACCGT
 CAGTGCCTCCGTGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAA
 GTCCGGCACCTCGTGACGCGGATTCGGCTCCAACAATGCCTGACGGACAATGGCCGCA
 TAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCCGGGATTCCCAATACGAGGTCGCCA
 ACATCTTCTTGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCG
 GAGGCATCCGGAGCTTGACGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTT
 GACCAACTCTATCAGAGCTTGGTTGACGGCAATTCGATGATGCAGCTTGGGCGCAGGGTC
 GATGCGACGCAATCGTCCGATCCGGAGCCGGACTGTCGGGCGTACACAAATCGCCCGCA
 GAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGAAACCGAC
 GCCCCAGCACTCGTCCGGATCGGGAGATGGGGGAGGCTAACTGAAACACGGAAGGAGAC
 AATACCGGAAGGAACCGCGCTATGACGGCAATAAAAAGACAGAATAAAACGCACGGGTGT
 TGGGTCGTTTGTTCATAAACGCGGGGTTCCGGTCCCAGGGCTGGCACTCTGTGATACCCCAC
 CGAGACCCATTGGGCCAATACGCCCGCTTCTTCTTTTCCCACCCCAACCCCAAGTTC
 GGGTGAAGGCCAGGGCTCGCAGCCAACGTCCGGGGCGGAGGCCCTGCCATAGCCACTGG
 CCCCCTGGGTTAGGGACGGGGTCCACTAGCTAGTTCTAGTATGCATGGCGGTAATACGGTT
 ATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAGGCCAGCAAAGGC
 CAGGAACCGTAAAAAGGCCGCTTGTGCGTTTTCATAGGCTCCGCCCCCTGACGAG
 CATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACC
 AGGCGTTTTCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGTTACCGGAT
 ACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGCGCTTCTCATAGCTCACGCTGTAGGTATC
 TCAGTTCGGTGTAGGTCTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACAGCCC
 GACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC
 GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTAC
 AGAGTTCCTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGC
 GCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAA
 CCACCGCTGGTAGCGGTGGTTTTTTTGGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGG

ATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTCAC
 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTACCTAGATCCTTTTAAATTA
 AAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACCTGAGGCTGACAG**TTACCAATG**
CTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTCCATCCATAGTTGCCTGACT
CCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATG
ATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAA
GGGCCGAGCGCAGAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTGTTGC
CGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTAC
AGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTCCCAACGAT
CAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGTTAGCTCCTTCGGTCTCCG
ATCGTTGTGAGAAGTAAGTTGGCCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAA
TTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC
ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGCGTCAATACGGGATAAT
ACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAA
AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAAC
TGATCTTCAGCATCTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAA
TGCCGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATCCTCAGGACTCTT
CCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGA
ATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC

pOPRS_GFP DNA sequence:

HSV TK poly-adenylation signal

NeoR / KanR ORF

F1 origin of replication

RSV promoter

lacO operator

SV40 intron

GFP ORF

ColE1 / pUC origin of replication

AmpR ORF / AmpR promoter

TCGCGGTTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCAC
 AGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGT
 TGGCGGGTGTCCGGGCTGGCTTAACATATGCGGCATCAGAGCAGATTGTACTGAGAGTGCA

CCACGAACAAACGACCCAACACCCGTGCGTTTTATTCTGTCTTTTTATTGCCGATCCCCTCAG
 AAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACC
 GTAAAGCAGGAGGAAGCGGTACGCCATTGCGCGCAAGCTCTTCAGCAATATCACGGGTA
 GCCAACGCTATGTCCTGATAGCGGTCCGCCACACCAGCCGCCACAGTCGATGAATCCAG
 AAAAGCGGCCATTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTACGACGAG
 ATCCTCGCCGTCGGGCATGCGCGCCTGAGCCTGGCGAACAGTTCGGCTGGCGGAGCCCC
 TGATGCTCTTCGTCAGATCATCCTGATCGACAAGACCGGCTCCATCCGAGTACGTGCTCG
 CTCGATGCGATGTTTCGCTTGGTGGTGAATGGCAGGTAGCCGGATCAAGCGTATGCAGC
 CGCCGATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGA
 GATCCTGCCCCGCACTTCGCCAATAGCAGCCAGTCCCTCCCGCTTCAGTGACAACGTG
 AGCACAGCTGCGCAAGGAACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCT
 GCAGTTCATTCAGGGCACCGGACAGGTCGGTCTTGACAAAAAGAACCGGGCGCCCCTGCG
 TGACAGCCGGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCAGTCATAGCCG
 AATAGCCTCTCCACCCAAGCGGCCGAGAACCTGCGTGCAATCCATCTTGTTCAATGGCCGA
 TCCCATATTGGCTGCAGGGTCGCTCGGTGTTGAGGCCACACGCGTCACCTTAATATGCGAA
 GTGGACCTGGGACCGCGCCGCCCCGACTGCATCTGCGTGTTGCAATTCGCCAATGACAAGA
 CGCTGGGCGGGGTTTGCTCGACATTGGGTGAAACATTCCAGGCCTGGGTGGAGAGGCTT
 TTTGCTTCTCTTGCAAACCACTGCTCGACATTGGGTGAAACATTCCAGGCCTGGGTG
 GAGAGGCTTTTTGCTTCTCTTGCAAACCACTGCTCGATATGCGGTGTGAAATACCGCA
 CAGATGCGTAAGGAGAAAATACCGCATCAGGAATTGTAAACGTTAATATTTTGTAAAATT
 CGCGTTAAATTTTTGTAAATCAGCTCATTITTTAAACCAATAGGCCGAAATCGGCAAATCCC
 TTATAAATCAAAGAATAGACCGAGATAGGGTTGAGTGTGTTCCAGTTTGAACAAGAGT
 CCACTATTAAGAACGTGGACTCCAACGTCAAAGGGCGAAAACCGTCTATCAGGGCGATG
 GCCACTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTA
 AATCGGAACCCTAAAGGGAGCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTG
 GCGAGAAAGGAAGGGAAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAG
 CGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTC
 GCGCCATTGCCATTCAGGCTACGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCG
 CTATTACGCCAGCTGGCGAAGGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCA
 GGGTTTTCCAGTCACGACGTTGTA AACGACGGCCAGTGAATTGTAATACGACTGACTATA
 GCCGGAATTGGAGCTAGGCCTACGTAGCGCGGAGCTCCACCGCGGTGGCAGTACAATCT
 GCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAG
 TAGTGCGGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGA
 ATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTATC
 TGAGGGGACTAGGGTGTGTTTAGGCGAAAAGCGGGGCTTCGGTTGTACGCGGTTAGGAGT
 CCCCTCAGGATATAGTAGTTTCGCTTTTGCATAGGGAGGGGGAAATGTAGTCTTATGCAATA
 CTCTTGAGTCTTGCAACATGGTAACGATGAGTTAGCAACATGCCTTACAAGGAGAGAAAA
 AGCACCGTGCATGCCGATTGGTGAAGTAAGGTGGTACGATCGTGCCTTATTAGGAAGGC
 AACAGACGGGTCTGACATGGATTGGACGAACCACTGAATCCGCATTGCAGAGATATTGTA
 TTTAAGTGCCTACTAGGTTGTGG AATTGTGAGCGCTCACAATTCCACAGTCGACCCTAGTAT
 ACAATAACGCCATTTGACCATTACCACATTGGTGTGCACCTCCAAGCTTGGACAACTAC
 CTACAGAGATTTAAAGCTCTAA GGTAATATAAAATTTTTAAGTGTATAATGTGTTAAACTA
 CGGATCCGCTCCCATTAGGCCTACAATGGTGAGACAAGTAGCCAACAGGGAAGGGTTGCA
 AATATCATTGGGCACACCTATGATAATATTGATGAAGCAGACAGTATTCAGCAAGTAACTG
 AGAGGTGGGAAGCTCAAAGCCAAAGTCTAATGTGCAGTCAGGTGAATTTATTGAAAAATT
 TGAGGCTCCTGGTGGACTAGGGTCGACTGTGG AATTGTGAGCGCTCACAATTCCACAACCT
 AGTTCAAAGAACTGCTCCTCAGGGATCCTAATTGTTTGTGATTTTAGATTCCAACCAAGCTT

GCTGCCCCGCGTAATACGACTCACTATAGGGCGAATTGGGTACATGGTGAGCAAGGGCG
 AGGAGCTGTTACCGGGGTGGTGCCCATCTGGTCGAGCTGGACGGCGACGTAAACGGCC
 ACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGA
 AGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCTCGTGACCACCCTGACC
 TACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGGACTTCTTCAAGTC
 CGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTAC
 AAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAG
 GGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAAC
 AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGA
 TCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCG
 CATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTG
 AGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCGG
 GGATCACTCTCGGCATGGACGAGCTGTACAAGTCCGGACTCAGATCTCGAGCTCAAGCTTC
 GAATTCTAAGGATCCAATGTAAGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTCCCTTTA
 GTGAGGGTTAATTGCGCGGGCCGCAATTCCTAAGAAGTGAACACGGAAGGAGACAATAC
 CGGAAGGAACCCGCGCTATGACGGCAATAAAAAGACAGAATAAAAACGCACGGTGTGGGT
 CGTTTGTTCAATAACGCGGGGTTCCGGTCCAGGGCTGGCACTCTGTGCATACCCACCGAG
 ACCCATTGGGGCCAATACGCCCGGTTTCTCCTTTTCCCAACCCACCCCAAGTTCGGG
 TGAAGGCCAGGGCTCGCAGCCAACGTCGGGGCGCAAGCCCTGCCATAGCCACGGGCC
 CGTGGGTTAGGGACGGGGTCCCCATGGGGAATGTTTTATGGTTCGTGGGGGCTAGCCGG
 CAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGGCGTTTTGCGTATTGGGCGCTCT
 CCGTTCCTCGCTCACTGACTCGCTGCGCTCGGTGTTCCGGTGCGGCGAGCGGTATCAGCT
 CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAATG
 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAGGCCGCGTTGCTGGCGTTTTC
 CATAGGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGGTGGCGAA
 ACCGACAGGACTATAAAGATAACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCGCTCTCT
 GTTCCGACCCTGCCGTTACCGGATACCTGTCCGCCTTTCTCCCTCGGGAAGCGTGGCGCT
 TTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCT
 GTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG
 TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA
 GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCCTGAAGTGGTGGCCTAACTACGGCTACAC
 TAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAATAAGAGTT
 GGTAGCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTTGTGGCAACGA
 GCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTACGGGGTCT
 GACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGA
 TCTTACCTAGATCCTTTTAAATTAATAAATGAAGTTTTAAATCAATCTAAAGTATATAGAT
 AACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTA
 TTTTCGTTCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTT
 ACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTACCCGGCTCCAGATTTAT
 CAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTGCAACTTTATCCGC
 CTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTT
 TGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTTTGGTATGGCT
 TCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAA
 AGCGGTTAGTCTCTCGGTCCTCCGATCGTTGTGAGAAGTAAGTTGGCCGAGTGTATCAC
 TCATGGTTATGGCAGCACTGCATAATTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTG
 TGAAGTGGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTC
 TTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA

TTGAAAACGTTCTTCGGGGCGAAAACCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCG
 ATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTACCAGCGTTTCTGGG
 TGAGCAAAAACAGGAAGGCAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAAATG
 TTGAATACTCATACTCTTCCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATG
 AGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCC
 CCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATA
 GCGTATCACGAGGCCCTTTCGTC

pCMVtTA DNA sequence:

CMV enhancer / CMV promoter

tTA ORF

SV40 poly-adenylation signal

EM7 promoter / *BleoR* ORF

SV40 origin of replication

ColE1 / pUC origin of replication

GCGCGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTT
 CATAGCCCATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCGCTGGCTGACC
 GCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
 GGACTTTCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCACTTGGCAGTACAT
 CAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCTG
 GCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGT
 CATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTT
 GACTCACGGGGATTTCCAAGTCTCCACCCATTGACGTCAATGGGAGTTTGTGTTTGGCACCA
 AAATCAACGGGACTTTCAAAATGTGCTAACAACCTCCGCCCATTTGACGCAAATGGGCGGT
 AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTG
 CTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTTGCATTCCTGCA
 GGTGCACATCGATCTTAAGCAGTACTTCTAGAGAATTGTGAGCGCTCACAATTCCATGAGAT
 TAGATAAAAGTAAAGTGATTAACAGCGCATTAGAGCTGCTTAATGAGGTCGGAATCGAAGG
 TTTAAACAACCCGTAAACTCGCCCAGAAGCTAGGTGTAGAGCAGCCTACATTGTATTGGCATG
 TAAAAATAAGCGGGCTTTGCTCGACGCCTTAGCCATTGAGATGTTAGATAGGCACCATACT
 CACTTTTGCCTTTAGAAGGGGAAAGCTGGCAAGATTTTTTACGTAATAACGCTAAAAGTTT
 TAGATGTGCTTTACTAAGTCATCGCGATGGAGCAAAAGTACATTTAGGTACACGGCCTACA
 GAAAAACAGTATGAACTCTCGAAAATCAATTAGCCTTTTTATGCCAACAAGGTTTTTCACTA
 GAGAATGCATTATATGCACTCAGCGCTGTGGGGCATTACTTTAGGTTGCGTATTGGAAGA
 TCAAGAGCATCAAGTCGCTAAAGAAGAAAGGGAAACACCTACTACTGATAGTATGCCGCCA
 TTATTACGACAAGCTATCGAATTATTTGATCACCAAGGTGCAGAGCCAGCCTTCTTATTCGG
 CCTTGAATTGATCATATGCGGATTAGAAAAACAATTAATGTGAAAGTGGGTCCGCGTAC
 AGCCGCGCGCTACGAAAAACAATTACGGGTCTACCATCGAGGGCCTGCTCGATCTCCCGG
 ACGACGACGCCCCGAAGAGGCGGGGCTGGCGGCTCCGCGCTGTCTTTCTCCCCGCGG

ACACACGCGCAGACTGTCGACGGCCCCCGACCGATGTCAGCCTGGGGGACGAGCTCCAC
 TTAGACGGCGAGGACGTGGCGATGGCGCATGCCGACGCGCTAGACGATTTTCGATCTGGAC
 ATGTTGGGGGACGGGGATTCCCCGGTCCGGGATTTACCCCCACGACTCCGCCCCCTACG
 GCGCTCTGGATATGGCCGACTTCGAGTTTGAGCAGATGTTTACCGATGCCCTTGAATTGAC
 GAGTACGGTGGGTAGGGATCCGAACAAAACTCATCTCAGAAGAGGATCTGAATATGCAT
 ACCGGTCATCATCACCATCACCATTGAGTTTGATCCCCGGGAATTCAGACATGATAAGATAC
 ATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAAT
 TTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTGGGGTGGGCG
 AAGAACTCCAGCATGAGATCCCCGCGCTGGAGGATCATCCAGCCGGCGTCCCGGAAAACG
 ATTCCGAAGCCCAACCTTTATAGAAGGCGGGCGGTGGAATCGAAATCTCGTAGCACGTGTC
 AGTCTGCTCCTCGGCCACGAAGTGCACGCAGTTGCCGGCCGGGTGCGCGAGGGCGAACTC
 CCGCCCCACGGCTGCTCGCCGATCTCGGTTCATGGCCGGCCCGGAGGCGTCCCGGAAGTTC
 GTGGACACGACCTCCGACCACTCGGCGTACAGCTCGTCCAGGCCGCGCACCCACACCCAGG
 CCAGGGTGTGTCCGGCACCACCTGGTCTGGACCGCGCTGATGAACAGGGTTCACGTCGTC
 CCGGACCACACCGGCGAAGTCGTCCTCCACGAAGTCCCGGGAGAACCCGAGCCGGTCCGT
 CCAGAACTCGACCGCTCCGGCGACGTGCGCGCGGTGAGCACCGGAACGGCACTGGTCAA
 CTTGGCCATGGTTTAGTTTCTCACCTTGTCTATTATACTATGCCGATATACTATGCCGATGA
 TTAATTGTCAACACGTGCTGATCAGATCCGAAAATGGATATACAAGCTCCCGGGAGCTTTTT
 GCAAAGCCTAGGCCTCCAAAAAGCCTCCTCACTACTTCTGGAATAGCTCAGAGGCAGAG
 GCGGCCTCGGCCTCTGCATAAATAAAAAAATTAGTCAGCCATGGGGCGGAGAATGGGCG
 GAACTGGGCGGAGTTAGGGGCGGGATGGGCGGAGTTAGGGGCGGGACTATGGTTGCTGA
 CTAATTGAGATGCATGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACACC
 TGGTTGCTGACTAATTGAGATGCATGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGGGAC
 TTTCCACACCCTCGTCGATTCTGAAGGTACCAGCACAGTGGACTCGAGAGATCTGGCCGGCT
 GGGCCCCGTTTCGAAGGTAAGCCTATCCCTAACCTCTCCTCGGTCTCGATTCTACGCGTACC
 GGTATCATCACCATCACCATTGAGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGT
 TGCCAGCCATCTGTTGTTTGCCCCTCCCCGTGCCTTCTTGACCCTGGAAGGTGCCACTCCC
 ACTGTCTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATT
 CTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGC
 ATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGTGGCGGTAATA
 CGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAA
 AAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTCATAGGCTCCGCCCCCTG
 ACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA
 GATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGCTTA
 CCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTA
 GGTATCTCAGTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTT
 CAGCCCCACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGA
 CTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGG
 TGCTACAGAGTTCTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGT
 ATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAA
 ACAAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAA
 AAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAA
 ACTCACGTTAAGGGATTTTGGTCATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCC
 TTTCTGCTCGCGGTTTCGGTGTGACGGTGAACCTCTGACACATGCAGCTCCCGGAGAC
 GGTACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGC
 GGGTGTGGCGGGTGTGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTAAGT
 AGTGCACCATATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCA

GGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTC
GCTATTACGCCA

pZK14 DNA sequence:

PTRE3G promoter

TetO operator sites

eGFP ORF

ColE1 / pUC origin of replication

AmpR ORF / AmpR promoter

CTC GAGTTTACTCCCTATCAGTGATAGAGA ACGTATGAAGAGTTTACTCCCTATCAGTGATA
GAGAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGA ACGTATAAGGAGTTTACTCCCT
ATCAGTGATAGAGA ACGTATGACCAGTTTACTCCCTATCAGTGATAGAGA ACGTATCTACAG
TTTACTCCCTATCAGTGATAGAGA ACGTATATCCAGTTTACTCCCTATCAGTGATAGAGA ACG
GTATAAGCTTTAGGCGGTGTACGGTGGGCGCCTATAAAAGCAGAGCTCGTTTAGTGAACCGT
CAGATCGCCTGGAGCAATTCCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCCTA
CCCTCGTAAAGTCGACAACGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCCACCGGGGT
GGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG
CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGC
AAGCTGCCCGTGCCCTGGCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAG
CCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACG
TCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAA
GTTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGA
CGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATCATG
GCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGAC
GGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGC
TGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAA
GCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGAC
GAGCTGTACAAGTCCGGAICTCAGATCTCGAGCTCAAGCTTCGAATTCTAAGGATCCAATGTA
ACTGTATTCAGCGATGACGAAATTCTTAGCTATTGTAATACTCTAGAGGATCTTTGTGAAGG
AACCTTACTTCTGTGGTGTGACATAATTGGACAACTACCTACAGAGATTTAAAGCTCTAAG
GTAATATAAAATTTTAAAGTGTATAATGTGTTAAACTACTGATTCTAATTGTTTGTGATTTT
AGATTCCAACCTATGGAAGTATGAATGGGAGCAGTGGTGAATGCCTTTAATGAGGAAAA
CCTGTTTTGCTCAGAAGAAATGCCATCTAGTGATGATGAGGCTACTGCTGACTCTCAACATT
CTACTCCTCCAAAAAAGAAGAGAAAGGTAGAAGACCCCAAGGACTTTCCTTCAAGATTGCT
AAGTTTTTTGAGTCATGCTGTGTTTAGTAATAGAACTCTTGCTTGCTTTGCTATTTACACCAC
AAAGGAAAAAGCTGCACTGCTATAACAAGAAAATTATGGAAAAATATTCTGTAACCTTTATAA
GTAGGCATAACAGTTATAATCATAACATACTGTTTTTTCTTACTCCACACAGGCATAGAGTGT
CTGCTATTAATAACTATGCTCAAAAATTGTGTACCTTTAGCTTTTTAATTTGTAAGGGGTTA
ATAAGGAATATTTGATGTATAGTGCCTTGACTAGAGATCATAATCAGCCATACCACATTTGT
AGAGGTTTTACTTGCTTTAAAAAACCTCCACACCTCCCCCTGAACCTGAAACATAAAATGA

ATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCA
 TCACAAATTTACAAATAAAGCATTTCCTACTGCATTCTAGTTGTGGTTTGTCCAAACTCAT
 CAATGTATCTTATCATGTCTGCGGCTCTAGAGCTGCATTAATGAATCGGCCAACGCGCGGG
 GAGAGGCGGTTTGCATTGGGCGCTCTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGG
 TCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGA
 ATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAGGCCAGGAACC
 GAAAAAGGCCGCTTGCTGGCGTTTTCATAGGCTCCGCCCCCTGACGAGCATCACAAA
 AATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAACCAGGCGTTTC
 CCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCG
 CCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG
 TGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGC
 GCCTTATCCGGTAACACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGC
 AGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT
 GAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTG
 AAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG
 GTAGCGGTGGTTTTTTTGGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGA
 AGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACACTCACGTTAAGGGA
 TTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAATGAAGTT
 TTAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGT
 GAGGCACCTATCTCAGCGATCTGTCTATTTGTTTCATCCATAGTTGCCTGACTCCCCGTCTGT
 TAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAG
 ACCCAGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCG
 CAGAAGTGGTCTGCAACTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA
 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTG
 GTGTCACGCTCGTCTTTGGTATGGCTTCATTAGCTCCGGTCCCAACGATCAAGGCGAGT
 TACATGATCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTCA
 GAAGTAAGTTGGCCGAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACT
 GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGA
 ATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCA
 CATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAAG
 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAG
 CATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAA
 AAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTTCTCTTTTCAATATTATT
 GAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAAT
 AAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCA
 TTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAAGAAT
 TC

pCMVGal4 DNA sequence:

CMV enhancer / CMV promoter

Gal4 ORF

bGH poly-adenylation signal

f1 origin of replication

SV40 promoter / NeoR / KanR ORF

SV40 poly-adenylation signal

ColE1 / pUC origin of replication

AmpR ORF / AmpR promoter

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACCTCTCAGTACAATCTGCTCTGATGCC
 GCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGA
 GCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAG
 GGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTAT
 TGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC
 CGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCATT
 GACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTCAAT
 GGGTGGAGTATTTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAG
 TACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGA
 CCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGA
 TCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGTTTTGACTCACGGGGATTTCCAAG
 TCTCCACCCCATTGACGTCAATGGGAGTTTTGTTTTGGCACCAAATCAACGGGACTTTCCAA
 AATGTCGTAACAACCTCCGCCCATTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGG
 TCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCTGCTTACTGGCTTATCGAAATT
 AATACGACTCACTATAGGGAGACCCAAGCTGGCTAGTTAAGCTTGGTACCGAGCTCGGATC
 CCAACATGAAGCTACTGTCTTCTATCGAACAAGCATGCGATATTTGCCGACTTAAAAAGCTC
 AAGTGCTCCAAAGAAAAACCGAAGTGCGCCAAGTGTCTGAAGAACAACCTGGGAGTGTGCG
 TACTCTCCAAAACCAAAGGTCTCCGCTGACTAGGGCACATCTGACAGAAGTGGAAATCAA
 GGCTAGAAAGACTGGAACAGCTATTTCTACTGATTTTTCTCGAGAAGACCTTGACATGATT
 TTGAAAATGGATTCTTTACAGGATATAAAAGCATTGTTAACAGGATTATTTGTACAAGATAA
 TGTGAATAAAGATGCCGTCACAGATAGATTGGCTTCAGTGGAGACTGATATGCCTCTAACAT
 TGAGACAGCATAGAATAAGTGCGACATCATCATCGGAAGAGAGTAGTAACAAAGGTCAA
 GACAGTTGACTGTATCGATTGACTCGGCAGCTCATCATGATAACTCCACAATTCCGTTGGAT
 TTTATGCCCAGGGATGCTCTTCATGGATTTGATTGGTCTGAAGAGGATGACATGTCGGATG
 GCTTGCCCTTCTGAAAACGGACCCCAACAATAATGGGTTCTTTGGCGACGGTTCTCTCTTAT
 GTATTCTTCGATCTATTGGCTTTAAACCGGAAAATTACACGAACTCTAACGTTAACAGGCTCC
 CGACCATGATTACGGATAGATACACGTTGGCTTCTAGATCCACAACATCCCGTTTACTTCAA
 AGTTATCTCAATAATTTTACCCTACTGCCCTATCGTGCACCTACCGACGCTAATGATGTTG
 TATAATAACCAGATTGAAATCGCGTCAAGGATCAATGGCAAATCCTTTTAACTGCATATT
 AGCCATTGGAGCCTGGTGTATAGAGGGGAATCTACTGATATAGATGTTTTTACTATCAA
 ATGCTAAATCTATTTGACGAGCAAGGTCTTCGAGTCAGGTTCCATAATTTGGTGACAGCC
 CTACATCTTGTGCGGATATACACAGTGGAGGCAGAAAACAATACTAGCTATAATTTTCA
 CAGCTTTTCCATAAGAATGGCCATATCATTGGGCTTGAATAGGGACCTCCCCTCGTCCTTCA
 GTGATAGCAGCATTCTGGAACAAGACGCCGAATTTGGTGGTCTGTCTACTCTGGGAGAT
 CCAATTGTCCTGCTTTATGGTCGATCCATCCAGCTTCTCAGAATAACAATCTCCTTCCCTTCT

TCTGTCGACGATGTGCAGCGTACCACAACAGGTCCCACCATATATCATGGCATCATTGAAAC
AGCAAGGCTCTTACAAGTTTTACAAAAATCTATGAACTAGACAAAACAGTAACTGCAGAA
AAAAGTCCTATATGTGCAAAAAAATGCTTGATGATTTGTAATGAGATTGAGGAGGTTTCGA
GACAGGCACCAAAGTTTTTACAATGGATATTTCCACCACCGCTCTAACCAATTTGTTGAAG
GAACACCCTTGGCTATCCTTTACAAGATTGGAAGTGAAGTGGAAACAGTTGTCTCTTATCATT
TATGTATTAAGAGATTTTTTCACTAATTTTACCCAGAAAAAGTCACAACTAGAACAGGATCA
AAATGATCATCAAAGTTATGAAGTTAAACGATGCTCCATCATGTTAAGCGATGCAGCACAAA
GAACTGTTATGTCTGTAAGTAGCTATATGGACAATCATAATGTCACCCCATATTTTGCCTGG
AATTGTTCTTATTACTTGTTCATGCAGTCCTAGTACCCATAAAGACTCTACTCTCAAACCTCA
AAATCGAATGCTGAGAATAACGAGACCGCACAAATTATTACAACAAATTAACACTGTTCTGAT
GCTATTAAAAAAAGTGGCCACTTTTAAAATCCAGACTTGTGAAAAATACATTCAAGTACTGG
AAGAGGTATGTGCGCCGTTTCTGTTATCACAGTGTGCAATCCATTACCGCATATCAGTTAT
AACAAATAGTAATGGTAGCGCCATTAAAAATATTGTCGGTTCTGCAACTATCGCCCAATACCC
TACTCTTCCGGAGGAAAATGTCAACAATATCAGTGTTAAATATGTTTCTCCTGGCTCAGTAG
GGCCTTCACCTGTGCCATTGAAATCAGGAGCAAGTTTCAAGTATCTAGTCAAGCTGTTATCT
AACGTCCACCCCTCTGTAACCTCCAGTGACAATACCAAGAAGCACACCTTCGCATCGCTC
AGTCACGCCTTTTCTAGGGCAACAGCAACAGCTGCAATCATTAGTGCCACTGACCCCGTCTG
CTTTGTTGGTGGCGCCAATTTAATCAAAGTGGGAATATTGCTGATAGCTCATTGCTCTTCA
CTTTCACTAACAGTAGCAACGGTCCGAACCTCATAACAACCTCAAACAAATTCTCAAGCGCTT
CACAAACCAATTGCCTCCTTAACGTTTATGATAACTTATGAATAATGAAATCACGGCTAGTA
AAATTGATGATGGTAATAATTCAAACCACTGTACCTGGTTGGACGGACCAAACCTGCGTAT
AACGCGTTTGAATCACTACAGGGATGTTAATACCACTACAATGGATGATGTATATAACTA
TCTATTCGATGATGAAGATACCCACCAAACCAAAAAAAGAGTAAAGCGGCCGCTCGAGTC
TAGAGGGCCCTTCGAACAAAAACTCATCTCAGAAGAGGATCTGAATATGCATACCGGTCAT
CATCACCATCACCATTGAGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAG
CCATCTGTTGTTTGGCCCTCCCCGTCCTTCTTGACCCTGGAAGGTGCCACTCCACTGTC
CTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGG
GGGTGGGGTGGGGCAGGACAGCAAGGGGGGAGGATTGGGAAGACAATAGCAGGCATGCT
GGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGG
GTATCCCCACGCGCCCTGTAGCGGCGCATTAAAGCGCGGGGGTGTGGTGGTTACGCGCAG
CGTGACCGCTACACTTGCCAGCGCCCTAGCGCCGCTCCTTTGCTTTCTTCCCTTCTTTCTC
GCCACGTTCCGCCGCTTTCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATT
TAGTGCTTACGGCACCTCGACCCAAAAAATGATTAGGGTGTGGTTACGTTAGTGGG
CCATCGCCCTGATAGACGGTTTTTCCGCCCTTGACGTTGGAGTCCACGTTCTTAATAGTGG
CTCTTGTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTGATTATAAGGG
ATTTTGGCGATTTCCGGCCTATTGGTTAAAAAATGAGCTGATTAACAAAAATTTAACGCGAA
TTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCA
GAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTC
CCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCC
CTAACTCCGCCCATCCCGCCCCTAACTCCGCCAGTCCGCCCATTTCTCCGCCCATGGCTGA
CTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCTCTGCCTCTGAGCTATTCCAGAAGTAG
TGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGTCCCGGGAGCTTGTATATCCAT
TTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTTTCGCATGATTGAACAAGATGGATTG
CACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCGGCTATGACTGGGCACAACAGA
CAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGGCGCCCGTTCTTTTT
GTCAAGACCGACCTGTCCGGTGCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCG
TGGCTGGCCACGACGGGCGTTCCTTGCAGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAA

GGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCC
TGCCGAGAAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCT
ACCTGCCATTTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTA CTGGATGGAAG
CCGGTCTTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACT
GTTCCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGAT
GCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCG
GCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAG
CTTGCGGGCAATGGGCTGACCGCTTCTCGTGCTTACGGTATCGCCGCTCCCGATTGCA
GCGCATCGCTTCTATCGCCTTCTGACGAGTTCTTCTGA GCGGGACTCTGGGGTTCGCGAA
ATGACCGACCAAGCGACGCCAACCTGCCATCACGAGATTTGATTCCACCGCCGCCTTCTA
TGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGG
GATCTCATGCTGGAGTTCTTCGCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAA
TAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGT
TTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTG
GCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAAC
ATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACAT
TAATTGCGTTGCGCTCACTGCCGCTTTCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAA
TGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTTCCGCTTCTCGC
TCACTGACTCGCTGCGCTCGGTCGTTCCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGC
GGTAATACGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAAGG
CCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTGGCGTTTTCATAGGCTCCGC
CCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA
CTATAAAGATAACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCTG
CCGCTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCA
CGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAAC
CCCCGTTACGCCGACCGCTGCGCCTTATCCGTA ACTATCGTCTTGAGTCCAACCCGGTA
AGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTAT
GTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAG
TATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGA
TCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGCAAGCAGCAGATTACGC
GCAGAAAAAAGGATCTCAA GAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTG
GAACGAAA ACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTACCTAG
ATCCTTTTAAATTA AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCT
GACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTGTTTCATCC
ATAGTTGCCTGACTCCCCGTGCTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCC
CAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACC
AGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCTCCATCCAGTC
TATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTG
TTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTAGCTCC
GGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGGTTAGCT
CCTTCGGTCTCCGATCGTTGTCAGAAGTAAGTTGGCCGAGTGTTATCACTCATGGTTATG
GCAGCACTGCATAATTCTTACTGTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAG
TACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCGGCGTC
AATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGT
TCTTCGGGGCGAAA ACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCAC
TCGTGCACCCAACTGATCTTACGATCTTTTACTTTACCAGCGTTTCTGGGTGAGCAAAAAC
AGGAAGGCAAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATACTCAT

ACTCTTCCTTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACAT
 ATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC
 CACCTGACGTC

pUAS-GG DNA sequence:

f1 origin of replication

UAS promoter

GFP ORF

β -globin intron

SV40 poly-adenylation signal

ColE1 / pUC origin of replication

AmpR ORF / AmpR promoter

CACCTAAATTGTAAGCGTTAATATTTTGTAAAATTCGCGTTAAATTTTTGTAAATCAGCTC
 ATTTTTAAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAGAATAGACCGAGA
 TAGGGTTGAGTGTTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTCCAA
 CGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCACTACGTGAACCATCACCTAA
 TCAAGTTTTTTGGGGTTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCC
 GATTTAGAGCTTGACGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCG
 AAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCACCACA
 CCCGCCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTTCAGGCTGCGCAACT
 GTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGAT
 GTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTCACGACGTTGTAAAAC
 GACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGGATCCAAGCTTGCATGC
 CTGCAGGTCGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTC
 CGAGCGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAGCGGAGACTCTAGCGAGCG
 CCGGAGTATAAATAGAGGCGCTTCGTCTACGGAGCGACAATTCAATTCAAACAAGCAAAGT
 GAACACGTCGCTAAGCGAAAGCTAAGCAAATAAACAAGCGCAGCTGAACAAGCTAAACAA
 TCTGCAGTAAAGTGCAAGTTAAAGTGAATCAATTAAGTAACCAGCAACCAAGTAAATCA
 ACTGCAACTACTGAAATCTGCCAAGAAGTAATTATTGAATACAAGAAGAGAACTCTGAATA
 GGGAATTGGGAATTCGTTAACAGATCCGATATCCGCCACCATGGTGAGCAAGGGCGAGGA
 GCTGTTACCGGGGTGGTGCCCATCTGGTTCGAGCTGGACGGCGACGTAAACGGCCACAA
 GTTCAGCGTGCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTT
 CATCTGCACCACCGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCTGACCTACG
 GCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGACTTCTTCAAGTCCGCC
 ATGCCCGAAGGTTGGTATCAAAATGATCTATAACTTCGTATAGCATAATTATACGAAGTTATG
 GTTACAAGACAGGTTTAAGGAGACCAATAGAAACTGGGCATGTGGAGACAGAGAAGACTC

TTGGGTTTCTGATAGGCACTGACTCTCTGCCTATTGGTCTATTTTCCCACCCCTTAGGGTAC
 GTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGA
 AGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG
 ACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATCAT
 GGCCGACAAGCAGAAGAACGGCATCAAGGCGAAGTTCAAGATCCGCCACAACATCGAGGA
 CGGCGGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTG
 CTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGA
 AGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGA
 CGAGCTTTACAAGGATATCCCTGAATTCTAGATAACTGATCATAATCAGCCATACCACATTTG
 TAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGA
 ATGCAATTGTTGTTGTTAACTTGTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCA
 TCACAAATTTACAAATAAAGCATTITTTTCACTGCATTCTAGTTGTGGTTGTCCAAACTCAT
 CAATGTATCTTAATTTAAATTGGGCGCGCCCAACTAGTTCTAGAGCGGCCGCCACCAGCGGTG
 GAGCTCCAGCTTTTGTCCCTTATGAGGGTTAATTTGAGCTTGGCGTAATCATGGTCAT
 AGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGC
 AATAAGTGAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCT
 CACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGC
 GCGGGGAGAGGGCGGTTTGCATTGGGCGCTTCCGCTTCCGCTCCTCGCTCACTGACTCGTGC
 GCTCGGTGTTTCGGTGC GGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATC
 CACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCA
 GGAACCGTAAAAAGGCCGCTTGCTGGCGTTTCCATAGGCTCCGCCCCCTGACGAGCA
 TCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCA
 GGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGCTTACCGGATA
 CCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCATAGCTCACGCTGTAGGTATCT
 CAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTACGCC
 GACCGCTGCGCCTTATCCGGTAACATATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC
 GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTAC
 AGAGTTCCTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGC
 GCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAA
 CCACCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAAAGG
 ATCTCAA GAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACCTCAC
 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTACCTAGATCCTTTTAAATTA
 AATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATG
 CTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTGTTTCATCCATAGTTGCCTGACT
 CCCCCTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATG
 ATACCGCGAGACCCACGCTCACC GGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAA
 GGGCCGAGCGCAGAAGTGGTCTGCAACTTATCCGCTCCATCCAGTCTATTAATTGTTGC
 CGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCACAACGTTGTTGCCATTGCTAC
 AGGCATCGTGGTGTACGCTCGTCTTGGTATGGCTTCAATCAGCTCCGGTCCCAACGAT
 CAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGTTAGCTCCTTCGGTCTCCG
 ATCGTTGTGAGAAGTAAGTTGGCCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAA
 TTCTTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC
 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGGCCGGCGTCAATACGGGATAAT
 ACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAA
 AACTCTCAAGGATCTTACCCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAAC
 TGATCTCAGCATCTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAA
 TGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTC

AATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTT
 AGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC

pOPRS DNA sequence:

HSV TK poly-adenylation signal

NeoR / KanR ORF / HSV TK promoter

f1 origin of replication

Multiple Cloning Site (MCS)

ColE1 / pUC origin of replication

AmpR ORF / AmpR promoter

TCGCGCGTTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCAC
 AGTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGT
 TGGCGGGTGTGCGGGCTGGCTTAACATGCGGCATCAGAGCAGATTGTACTGAGAGTGCA
 CCACGAACAAACGACCCAACACCCGTGCGTTTTATTCTGTCTTTTTATTGCCGATCCCTCAG
 AAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACC
 GTAAAGCAGGGAAGCGGTGAGCCATTGCGCCGAAGCTCTTCAAGCAATATCACGGGTA
 GCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGCCACAGTCGATGAATCCAG
 AAAAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTACGACGAG
 ATCCTCGCCGTCGGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCC
 TGATGCTCTTCGTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCG
 CTCGATGCGATGTTTTCGTTGGTGGTGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGC
 CGCCGATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGA
 GATCCTGCCCGGCACTTCGCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTG
 AGCACAGCTGCGCAAGGAACGCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCT
 GCAGTTCATTGAGGACCGGACAGGTCGGTCTTGACAAAAAGAACCGGGCGCCCCTGCGC
 TGACAGCCGGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCAGTCATAGCCG
 AATAGCCTCTCCACCAAGCGGCCGAGAACCTGCGTGCAATCCATCTTGTTCAATGGCCGA
 TCCCATATTGGCTGCAGGGTCGCTCGGTGTTGAGGCCACACGCGTCACCTTAATATGCGAA
 GTGGACCTGGGACCGCGCCGCCCCGACTGCATCTGCGTGTTGCAATTCGCCAATGACAAGA
 CGCTGGGCGGGGTTTGCTCGACATTGGGTGAAACATTCCAGGCCTGGGTGGAGAGGCTT
 TTTGCTTCTCTTGCAAAACCACACTGCTCGACATTGGGTGAAACATTCCAGGCCTGGGTG
 GAGAGGCTTTTTGCTTCTCTTGCAAAACCACACTGCTCGATATGCGGTGTGAAATACCGCA
 CAGATGCGTAAGGAGAAAATACCGCATCAGGAAATTGTAAACGTTAATATTTTGTAAAT
 CGCGTTAAATTTTTGTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAATCCC
 TTATAAATCAAAGAATAGACCGAGATAGGGTTGAGTGTGTTCCAGTTTGAACAAGAGT
 CCACTATTAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATG
 GCCACTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTA
 AATCGGAACCCTAAAGGGAGCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTG

GCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAG
 CGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTC
 GCGCCATTCGCCATTCAGGCTACGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCG
 CTATTACGCCAGCTGGCGAAGGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCA
 GGGTTTTCCAGTCACGACGTTGTAAACGACGGCCAGTGAATTGTAATACGACTGACTATA
 GCCGAATTGGAGCTAGGCTACGTAGCGCGGAGCTCCACCGCGGTGGCAGTACAATCT
 GCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAG
 TAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGA
 ATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTATC
 TGAGGGGACTAGGGTGTGTTTAGGCGAAAAGCGGGGCTTCGGTTGTACGCGGTTAGGAGT
 CCCCTCAGGATATAGTAGTTTTGCTTTTTGCATAGGGAGGGGAAATGTAGTCTTATGCAATA
 CTCTTGTAGTCTTGAACATGGTAACGATGAGTTAGCAACATGCCTTACAAGGAGAGAAAA
 AGCACCGTGCATGCCGATTGGTGAAGTAAGGTGGTACGATCGTGCCTTATTAGGAAGGC
 AACAGACGGGTCTGACATGGATTGGACGAACCACTGAATTCCGCATTGCAGAGATATTGTA
 TTTAAGTGCCTACTAGGTTGTGGAATTGTGAGCGCTACAATCCACAGTCGACCCTAGTAT
 ACAATAACGCCATTTGACCATTACCACATTGGTGTGCACCTCCAAGCTTGGACAACTAC
 CTACAGAGATTTAAAGCTCTAAGGTAAATATAAAATTTTTAAGTGTATAATGTGTTAAACTA
 CGGATCCGCTCCCATTAGGCCTACAATGGTGAGACAAGTAGCCAACAGGGAAGGGTTGCA
 AATATCATTTGGGCACACCTATGATAATATTGATGAAGCAGACAGTATTCAGCAAGTAACTG
 AGAGGTGGGAAGCTCAAAGCCAAAGTCTAATGTGCAGTCAGGTGAATTTATTGAAAAATT
 TGAGGCTCCTGGTGGACTAGGGTCGACTGTGGAATTGTGAGCGCTACAATCCACAACCT
 AGTTCAAAGAAGTCTCCTCAGGGATCCTAATTGTTTGTGATTTTAGATTCCAACCAAGCTT
 GCTGCCCGCGCTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCTCGAGG
 TCGACGGTATCGATAAGCTTGATATCGAATTCCTGCAGCCCGGGGATCCACTAGTTCTAGA
 CGGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTCCCTTTAGTGAGGGTTAATTGCGCGG
 GCCGAATTTTAAAGAAGTAAACACGGAAGGAGACAATACCGGAAGGAACCCGCGCTAT
 GACGGCAATAAAAAGACAGAATAAAACGCACGGTGTGGGTGTTGTTGTTGATAAACGCGG
 GGTTCCGGTCCCAGGGCTGGCACTGTGCGATACCCACCGAGACCCATTGGGGCCAATAC
 GCCGCGTTTTCTTCTTTTCCCCACCCCAAGTTCGGGTGAAGGCCAGGGCTCGC
 AGCCAACGTCGGGGCGGCAAGCCCTGCCATAGCCACGGGCCCGTGGGTTAGGGACGGG
 GTCCCCATGGGGAATGGTTTATGGTTCGTGGGGGCTAGCCGGCAGCTGCATTAATGAATC
 GGCCAACGCGCGGGGAGAGGCGGTTTGGTATTGGGCGCTCTCCGCTTCTCGCTACTG
 ACTCGCTGCGCTCGGTCGTTCCGGTGTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAAT
 ACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAGGCCAGCA
 AAAGGCCAGGAACCGTAAAAAGGCCGCGTGTGCTGGCGTTTTCATAGGCTCCGCCCCCT
 GACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAA
 AGATACCAGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGCTT
 ACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGT
 AGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGT
 TCAGCCCGACCGCTGCGCCTTATCCGGTAACATCGTCTTGAGTCCAACCCGGTAAGACACG
 ACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCG
 GTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGG
 TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCA
 AACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGAACGAGCAGATTACGCGCAGAAA
 AAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAA
 AACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTACCTAGATCCTTTT
 AAATTAATAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTT

ACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTG
 CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCT
 GCAATGATACCGCGAGACCCACGCTCACGGGCTCCAGATTTATCAGCAATAAACCAGCCAG
 CCGGAAGGGCCGAGCGCAGAAGTGGTCTGCAACTTTATCCGCCTCCATCCAGTCTATTAAT
 TGGTGCCGGGAAGCTAGAGTAAGTAGTTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCAT
 TGCTACAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCG
 AACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGT
 CCTCCGATCGTTGTCAGAAGTAAGTTGGCCGAGTGTTATCACTCATGGTTATGGCAGCACT
 GCATAATTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAAC
 CAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGG
 GATAATACCGCGCCACATAGCAGAAGTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGG
 GCGAAAACCTCAAGGATCTTACCCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGCAC
 CCAACTGATCTTCAGCATCTTTACTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGG
 CAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCC
 TTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAAT
 GTATTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGA
 CGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCT
 TTCGTC

pZK11 DNA sequence:

CMV enhancer / CMV promoter

Chimeric intron

hmGFP ORF

SV40 poly-adenylation signal

AmpR promoter / AmpR ORF

ColE1 / pUC origin of replication

TCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTG
 GCCATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATG
 ACCGCCATGTTGGCATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAG
 TTCATAGCCCATATATGGAGTTCGCGTTACATAAATTACGGTAAATGGCCCCGCTGGCTGA
 CCGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAAT
 AGGGACTTTCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA
 CATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCCG
 CTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCTACTTGGCAGTACATCTACGTATT
 AGTCATCGCTATTACCATGGTGTATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGG
 TTTGACTCACGGGATTTCCAAGTCTCACCCATTGACGTCAATGGGAGTTTGTTTTGGCA
 CAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGC
 GGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTGTAGTGAACCGTCAGATCA

CTAGAAGCTTTATTGCGGTAGTTTATCACAGTTAAATTGCTAACGCAGTCAGTGCTTCTGAC
 ACAACAGTCTCGAACTTAAGCTGCAGAAAGTTGGTCGTGAGGCACTGGGCAGGTAAGTATCA
 AGTTACAAGACAGGTTTAAGGAGACCAATAGAACTGGGCTTGTGCGAGACAGAGAAGAC
 TCTTGCCTTCTGATAGGCACCTATTGGTCTTACTGACATCCACTTGCCTTCTCTCCACAGG
 TGCCACTCCCAGTTCAATTACAGCTCTAAGGCTAGAGTACTTAATACGACTCACTATAGGC
 TAGCCCCGGGGATATCGCCACCATGGGCGTGATCAAGCCCGACATGAAGATCAAGCTGCG
 GATGGAGGGCGCCGTGAACGGCCACAAATTCGTGATCGAGGGCGACGGGAAAGGCAAGC
 CCTTGGAGGGTAAGCAGACTATGGACCTGACCGTGATCGAGGGCGCCCCCTGCCCTTCGC
 TTATGACATTCTACCACCGTGTTCGACTACGGTAACCGTGTCTTCGCCAAGTACCCCAAGG
 ACATCCCTGACTACTTCAAGCAGACCTTCCCCGAGGGCTACTCGTGGGAGCGAAGCATGAC
 ATACGAGGACCAGGGAATCTGTATCGCTACAAACGACATCACCATGATGAAGGGTGTGGAC
 GACTGCTTCGTGTACAAAATCCGCTTCGACGGGGTCAACTTCCCTGCTAATGGCCCGGTGAT
 GCAGCGCAAGACCCTAAAGTGGGAGCCAGTACCGAGAAGATGTACGTGCGGGACGGCGT
 ACTGAAGGGCGATGTTAATATGGCACTGCTCTTGGAGGGAGGCGGCCACTACCGCTGCGA
 CTTCAAGACCACCTACAAAGCCAAGAAGGTGGTGCAGCTTCCGACTACCACTTCGTGGACC
 ACCGCATCGAGATCGTGAGCCACGACAAGGACTACAACAAGTCAAGCTGTACGAGCACG
 CCGAAGCCACAGCGGACTACCCCGCCAGGCCGGCTAATAGTTCTAGAGCGGCGCTTCCC
 TTTAGTGAGGGTAAATGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAAC
 CACAAGTGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATT
 TGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCA
 GGTTACAGGGGGAGATGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTAAA
 ATCCGATAAGGATCGATCCGGGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTC
 CCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAAGCGCG
 GCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTC
 CTTTCGCTTCTTCCCTTCTTCTCGCCACGTTTCGCGGGCTTCCCCGTCAAGCTCTAAATCG
 GGGGCTCCCTTAGGGTTCCGATTTAGTGCTTACGGCACCTCGACCCCAAAAACTTGATT
 AGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTGGACGTTG
 GAGTCCACGTTCTTAATAGTGGACTCTGTTCCAAACTGGAACAACACTCAACCCTATCTCG
 GTCTATTCTTTGATTTATAAGGGATTTTGCCGATTTCCGCCTATTGGTTAAAAAATGAGCTG
 ATTTAACAAAAATTAACGCGAATTTAACAAAATATTAACGTTACAATTTCTGATGCGGT
 ATTTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATATGGTGCCTCTCAGTACAATCT
 GCTCTGATGCCGCATAGTTAAGCCAGCCCCGACCCCGCCAACACCCGCTGACGCGCCCTG
 ACGGGCTTGTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGC
 ATGTGTCAGAGTTTTACCGTCATACCGAAACGCGCGAGACGAAAGGGCCTCGTGATAC
 GCCTATTTTATAGGTTAATGTCATGATAATAATGTTTTCTTAGACGTGAGGTGGCACTTTTC
 GGGGAAATGTGCGCGGAACCCCTATTTGTTATTTTCTAAATACATTCAAATATGTATCCGC
 TCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATT
 CAACATTTCCGTGTCGCCCTTATTCCCTTTTTGCGGCATTTGCCTTCTGTTTTGCTCACCC
 AGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACAT
 CGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAA
 TGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAA
 GAGCAACTCGGTGCGCCGATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCAC
 AGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATG
 AGTGATAAACAAGTGCAGGCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCG
 CTTTTTGCACAACATGGGGGATCATGTAACCTGCCTTGATCGTTGGGAACCGGAGCTGAAT
 GAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGC
 GCAAACATTAACCTGGCGAACTACTTACTCTAGCTTCCCGCAACAATTAATAGACTGGATG

GAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTCCGGCTGGCTGGTTTATTG
CTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAG
ATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGA
ACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGAC
CAAGTTTACTCATATATACTTTAGATTGATTTAAAACCTTCATTTTTAATTTAAAAGGATCTAG
GTGAAGATCCTTTTTGATAATCTCATGACCAAATCCCTTAACGTGAGTTTTCGTTCCACTGA
GCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAAT
CTGCTGCTTGCAAACAAAAAACCACCGCTACCAGCGGTGGTTTGTGGCCGGATCAAGAG
CTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCG
CTCTGCTAATCCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTG
GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCCGGCTGAACGGGGGGTTTCGTGC
ACACAGCCCAGCTTGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTAT
GAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGG
GTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGT
CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCG
GAGCCTATGAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTT
TTGCTCACATGGCTCGACAGATCT

A.4.2. Brightfield images – *in vivo* implementation of the LacI, tTA and Gal4 validation synthetic gene circuits

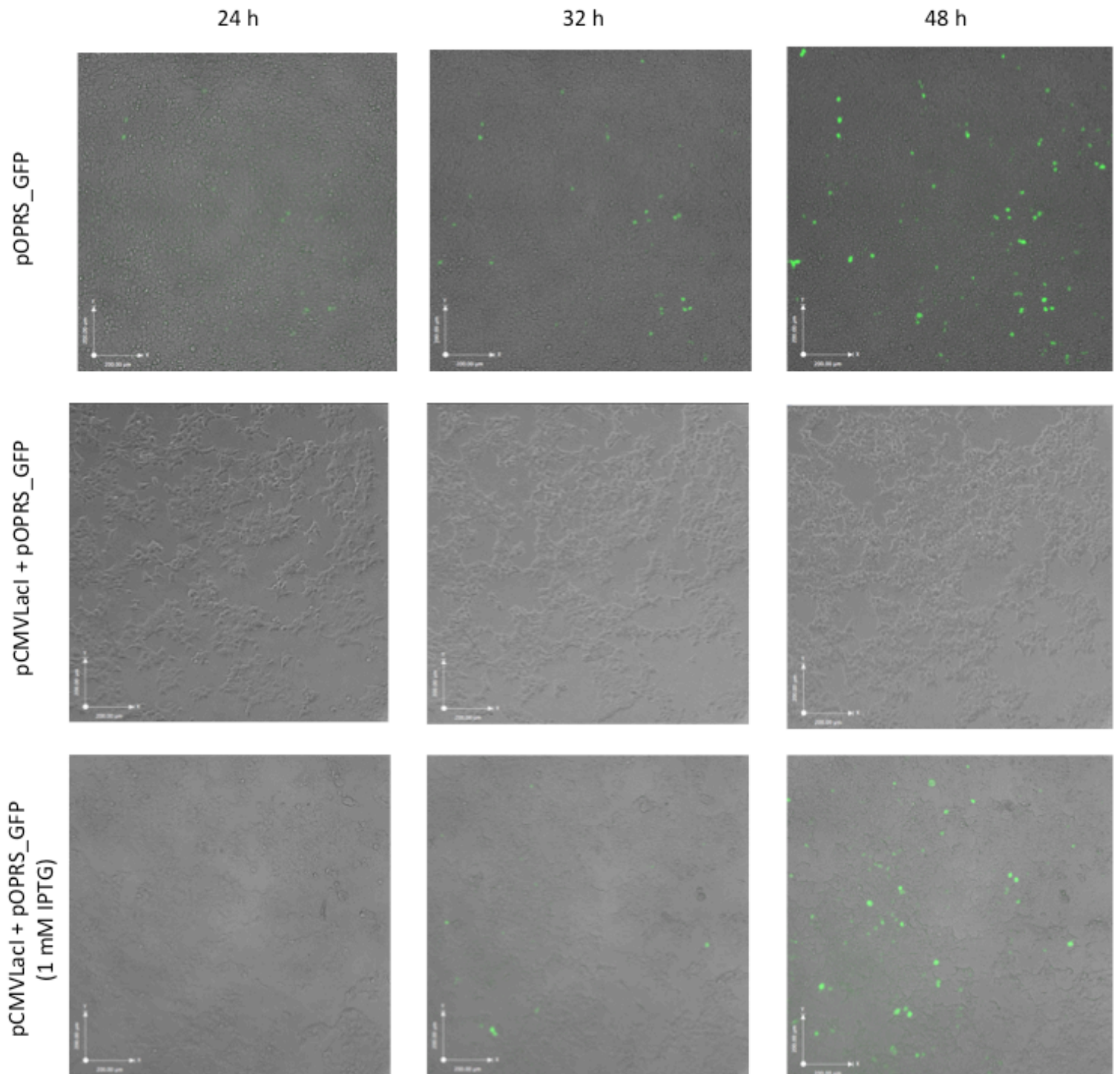


Figure A.4.2.1. Fluorescence channel images superimposed to brightfield channel images of HEK cells transfected with LacI validation synthetic gene circuit components. White arrows on the bottom left of each image represent a 200 μm distance (Top) Cells transfected with pOPRS_GFP 24, 32 and 46 hours after transfection. (Middle) Cells co-transfected with pCMVLacI and pOPRS_GFP 24, 32 and 46 hours after transfection. (Bottom) Cells co-transfected with pCMVLacI and pOPRS_GFP in DMEM medium supplemented with a final concentration of 1 mM IPTG.

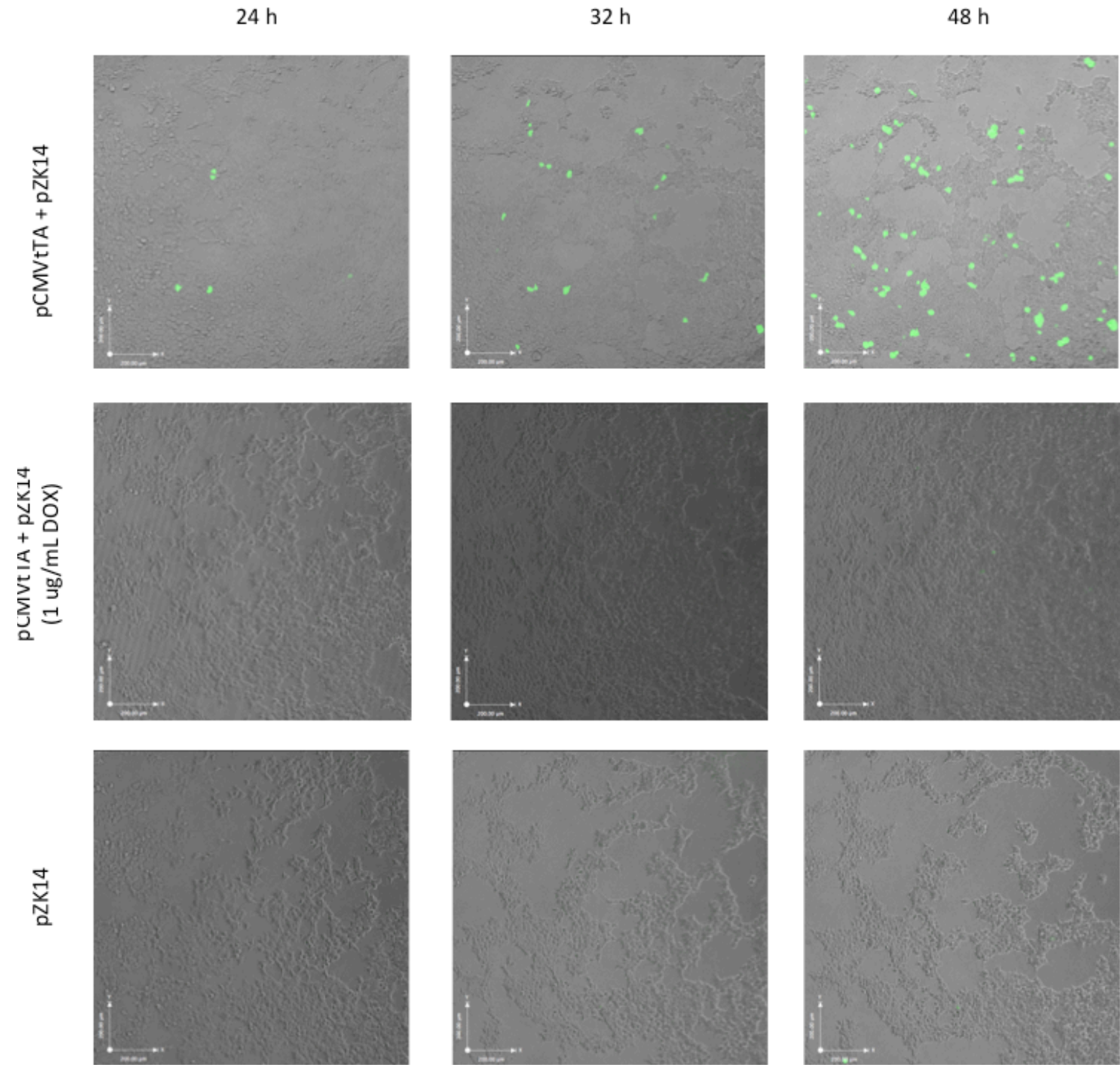


Figure A.4.2.2. Fluorescence channel images superimposed to brightfield channel images of HEK cells transfected with tTA validation synthetic gene circuit components. White arrows on the bottom left of each image represent a 200 μm distance (Top) Cells co-transfected with pCMVtTA and pZK14 24, 32 and 46 hours after transfection. (Middle) Cells co-transfected with pCMVtTA and pZK14 24, 32 and 46 hours after transfection in DMEM medium supplemented with a final concentration of 1 $\mu\text{g}/\text{mL}$ doxycycline. (Bottom) Cells transfected with pZK14 24, 32 and 46 hours after transfection.

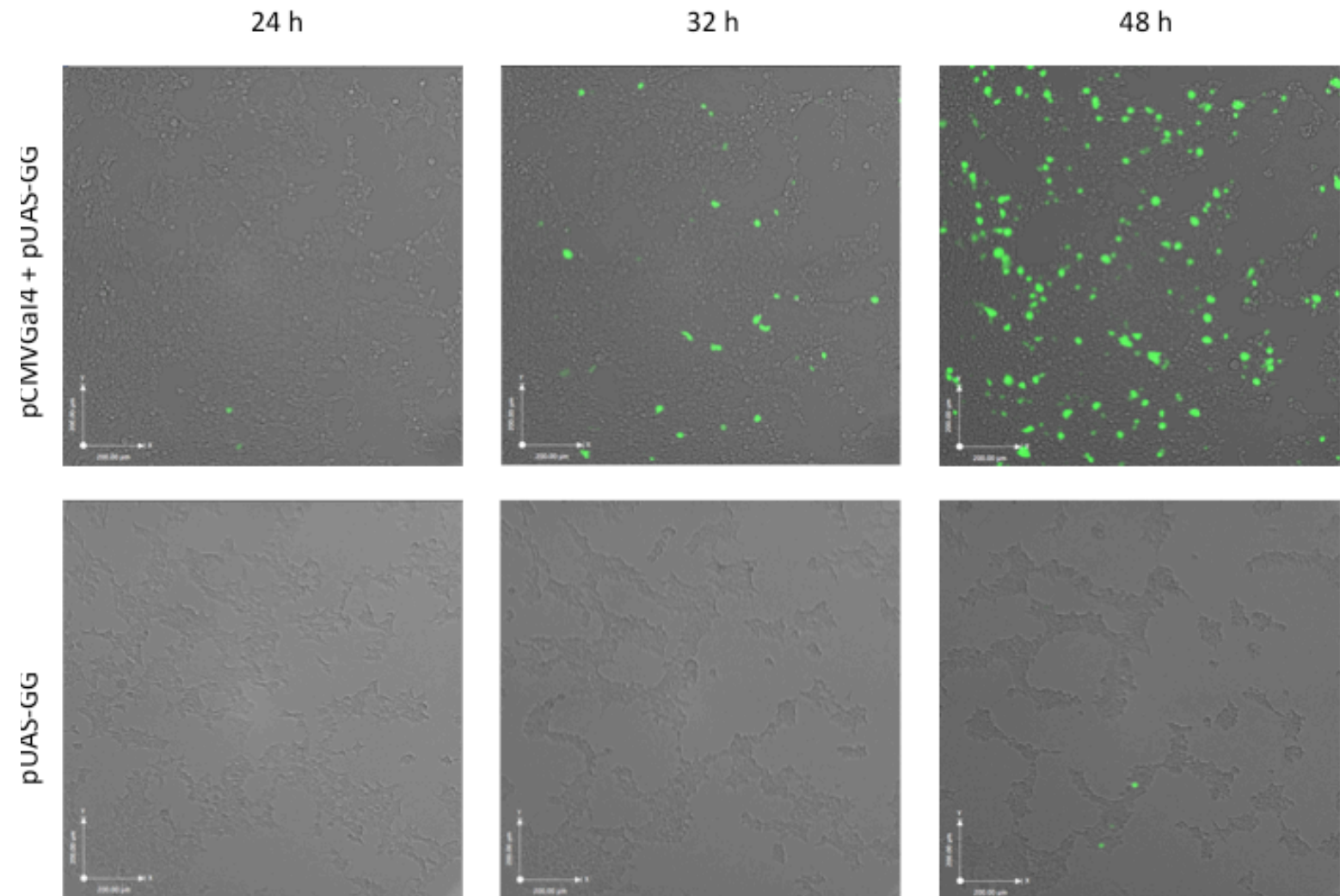


Figure A.4.2.3. Fluorescence channel images superimposed to brightfield channel images of HEK cells transfected with Gal4 validation synthetic gene circuit components. White arrows on the bottom left of each image represent a 200 μm distance (Top) Cells co-transfected with pCMVGal4 and pUAS-GG 24, 32 and 46 hours after transfection. (Bottom) Cells transfected with pUAS-GG 24, 32 and 46 hours after transfection.

Appendix A.5

A.5.1. Plasmid DNA sequences

pUAS_LacI DNA sequence:

6 adjacent UAS sites

IVS8 synthetic intron

lacI ORF / SV40 Nuclear Localization Signal

bGH poly-adenylation signal

f1 origin of replication

SV40 promoter

SV40 origin of replication

NeoR / *KanR* ORF

SV40 poly-adenylation signal

ColE1 / pUC origin of replication

AmpR ORF / *AmpR* promoter

```
GACGGATCGGGAGATCTCCCGATCCCCTATGGTGC ACTCTCAGTACAATCTGCTCTGATGCC
GCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGGAGGTCGCTGAGTAGTGCGCGA
GCAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAG
GGTTAGGCGTTTTGCGCTGCTTCGCGAAAGCTTCCGAGCTCTTACGCGGGTCAAGCGGAG
TACTGTCCTCCGAGTGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAGTCGAGGGTC
GAAGCGGAGTACTGTCCTCCGAGTGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAG
TCGACTCTAGAGGGTATATAATGGATCTCGAGATATCGGAGCTCGTTTAGTGAACCGTCAG
ATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCA
GCCTCCGCGGCCGGGAACGGTGCATTGGAACGCGCATTCCCCGTGTTAATTAACAGGTAAG
TGTCTTCTCCTGTTTCTTCCCCTGCTATTCTGCTCAACCTTCTATCAGAACTGCAGTATC
TGTATTTTTGCTAGCAGTAATACTAACGGTTCTTTTTTCTTTTACAGATGAAGCCCGTGAC
CCTGTACGACGTGGCCGAGTACGCTGGCGTGTCTACCAGACCGTGTCCCGGGTGGTCAAC
CAGGCCAGCCACGTGTCCGCAAGACCCGCGAGAAGGTGGAAGCCGCCATGGCCGAGCTG
AACTACATCCCCAACCGGGTGGCCAGCAGCTGGCCGGCAAGCAGTCTCTGCTGATCGGCG
TGGCCACCAGCAGCCTGGCCCTGCACGCCCTTCTCAGATCGTGGCCGCCATCAAGAGCAG
AGCCGACCAGCTGGGCGCCAGCGTGGTGGTGTCCATGGTGGAAAGATCCGGCGTGGAAGC
CTGCAAGGCCCGCGTGCACAACCTGCTGGCCAGAGAGTGTCCGGCCTGATCATCAACTAC
CCCCTGGACGACCAGGACGCCATTGCCGTGGAAGCTGCCTGCACCAACGTGCCCGCCCTGT
TCTGGACGTGTCCGATCAGACCCCATCAACAGCATCATCTTACGCCACGAGGACGGCACC
CGGCTGGGCGTGGAACATCTGGTGGCTCTGGGACACCAGCAGATCGCCCTGCTGGCTGGC
```

CCCCTGTCCTCCGTGTCTGCTAGACTGAGACTGGCCGGCTGGCACAAGTACCTGACCCGGA
 ACCAGATCCAGCCAATCGCCGAGAGAGAGGGCGATTGGAGCGCCATGTCCGGCTTCCAGC
 AGACCATGCAGATGCTGAACGAGGGGCATCGTGCCACCGCCATGCTGGTGGCTAACGACCA
 GATGGCCCTGGGCGCCATGCGGGCCATCACCGAGTCTGGACTGAGAGTGGGCGCCGACAT
 CAGCGTCGTGGGCTACGACGACACCGAGGACTCCAGCTGCTACATCCCCCCCCTGACCACC
 ATCAAGCAGGACTTCAGACTGCTGGGCCAGACCAGCGTGGACCGGCTGCTGCAGCTGTCTC
 AGGGCCAGGCCGTGAAGGGCAACCAGCTGCTGCCCGTGTCCCTGGTCAAGAGAAAGACCA
 CCCTGGCCCCAACACCCAGACCGCCTCTCAAGAGCCCTGGCCGACAGCCTGATGCAGCT
 GGCCAGACAGGTGTCCCGGCTGGAAAGCGGCCAGAGCAGCCTGAGGCCCCCCAAGAAGAA
 GCGGAAAGTCTGAGGTACCGAGCTCGGATCCACTAGTCCAGTGTGGTGGAAATTCTGCAGAT
 ATCCAGCACAGTGGCGGCCCGCTCGAGTCTAGAGGGCCCCGTTTAAACCCGCTGATCAGCCTC
 GACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCGTGCCTTCTTGACCCT
 GGAAGGTGCCACTCCACTGTCCTTCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGA
 GTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGG
 GAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGA
 ACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAAGCGCGGGC
 GGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTT
 CGCTTCTTCCCTTCTTCTCGCCACGTTGCGCGGCTTCCCCGTCAAGCTCTAAATCGGGG
 GCTCCCTTAGGGTCCGATTTAGTGCTTACGGCACCTCGACCCCAAAAACTTGATTAGG
 GTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTGGACGTTGGAG
 TCCACGTTCTTAATAGTGGACTCTTGTCCAAACTGGAACAACACTCAACCCTATCTCGGTC
 TATTCTTTGATTATAAGGGATTTTGGCGATTCGGCCTATTGGTTAAAAAATGAGCTGATT
 TAACAAAAATTAACGCGAATTAAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTC
 CCCAGGCTCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGG
 TGTGGAAAGTCCCAGGCTCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGT
 CAGCAACCATAGTCCCGCCCCTAACTCCGCCATCCCGCCCTAACTCCGCCAGTTCGCCCC
 ATTCTCCGCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCT
 CTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCT
 CCCGGGAGCTTGTATATCCATTTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTTTCGC
 ATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCG
 GCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGC
 GCAGGGGCGCCGGTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAG
 GACGAGGCAGCGGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCAGCTGTGCTC
 GACGTTGTCAGTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGAT
 CTCCTGTATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCG
 GCTGCATACGCTTGATCCGGCTACCTGCCATTGACCAACGAAACATCGCATCGAGC
 GAGCACGTAAGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGAGCATC
 AGGGGCTCGCGCCAGCCGAACTGTTCCGAGGCTCAAGGCGCGCATGCCGACGGCGAGG
 ATCTCGTCGTGACCATGGCGATGCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTT
 TCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCCTATCAGGACATAGCGTTGG
 CTACCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGTTCCTCGTGCTTTAC
 GGTATCGCCGCTCCCGATTGCGAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA
 GCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCAACCTGCCATCACGAGATTT
 CGATTCCACCGCCGCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCT
 GGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCACCCCACCTGTTTATT
 GCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTCACAAATAAAGCATTITTT
 TTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCG

TCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTAT
 CCGCTCACAATTCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCT
 AATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTCCAGTCGGGAAAC
 CTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGTATTG
 GGCGCTCTCCGCTTCCCTCGCTCACTGACTCGCTGCGCTCGGTGTTTCGGCTGCGGCGAGCG
 GTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAA
 AGAACATGTGAGCAAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGG
 CGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACA AAAAATCGACGCTCAAGTCAGAGG
 TGGCGAAACCCGACAGGACTATAAAGATAACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC
 GCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCG
 TGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAG
 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCCACCGCTGCGCCTTATCCGGTAACTATCG
 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGG
 ATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACG
 GCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAA
 AGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTTTTTTTTGTTTGCAA
 GCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGG
 TCTGACGCTCAGTGGAACGAAAACACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAA
 GGATCTTACCTAGATCCTTTTAAATTA AAAATGAAGTTTTAAATCAATCTAAAGTATATATG
 AGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGT
 CTATTTTCGTTCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGG
 CTTACCATCTGGCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATT
 TATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTGCAACTTTATC
 CGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATA
 GTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTGTTTGGTATG
 GCTTCATTAGCTCCGGTTCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAA
 AAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGACGTGTTAT
 CACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTT
 CTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTG
 CTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGA ACTTTAAAAGTGCTCA
 TCATTGGAAAACGTTCTTCGGGGCGAAAACACTCTCAAGGATCTTACCGCTGTTGAGATCCAGT
 TCGATGTAACCCACTCGTGCACCCA ACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCT
 GGGTGAGCAAAAACAGGAAGGCCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAA
 ATGTTGAATACTCATACTCTTCTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTC
 ATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACA AATAGGGGTTCCGCGCACAT
 TTCCCCGAAAAGTGCCACCTGACGTC

pOPRS_tTA DNA sequence:

HSV TK poly-adenylation signal

NeoR / KanR ORF / HSV TK promoter

f1 origin of replication

RSV promoter

lacO operator sites

SV40 chimeric intron

tTA ORF

ColE1 / pUC origin of replication

AmpR ORF / AmpR promoter

TCGCGCGTTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCAC
AGTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGT
TGGCGGGTGTGGGGCTGGCTTAACCTATGCGGCATCAGAGCAGATTGTAAGTACTGAGAGTGCA
CCACGAACAAACGACCCAACACCCGTGCGTTTTATTCTGTCTTTTTATTGCCGATCCCCTCAG
AAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACC
GTAAGCAGCAGGAAGCGGTGAGCCATTGCGCGCAAGCTCTTCAGCAATATCACGGGTA
GCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGCCACAGTCGATGAATCCAG
AAAAGCGGCCATTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTGACGACGAG
ATCCTCGCCGTCGGGCATGCGCGCCTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCC
TGATGCTCTTCGTCAGATCATCCTGATCGACAAGACCGGCTCCATCCGAGTACGTGCTCG
CTCGATGCGATGTTTCGCTTGGTGGTGAATGGGAGGTAGCCGGATCAAGCGTATGCAGC
CGCCGATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGA
GATCCTGCCCCGGCACTTCGCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTG
AGCACAGCTGCGCAAGGAACGCCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCT
GCAGTTCATTAGGGCACCGGACAGGTCGGTCTTGACAAAAAGAACCGGGCGCCCCTGCGC
TGACAGCCGGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCAGTCATAGCCG
AATAGCCTCTCCACCAAGCGGCCGAGAACCTGCGTGCAATCCATCTTGTTCAATGGCCGA
TCCCATATTGGCTGCAGGGTCGCTCGGTGTTGAGGCCACACGCGTCACCTTAATATGCGAA
GTGGACCTGGGACCGCGCCGCCCCGACTGCATCTGCGTGTTGCAATTCGCCAATGACAAGA
CGCTGGGCGGGGTTTGCTCGACATTGGGTGAAACATTCCAGGCCTGGGTGGAGAGGCTT
TTTGCTTCTCTTGCAAAACCACTGCTCGACATTGGGTGAAACATTCCAGGCCTGGGTG
GAGAGGCTTTTTGCTTCTCTTGCAAAACCACTGCTCGATATGCGGTGTGAAATACCGCA
CAGATGCGTAAGGAGAAAATACCGCATCAGGAAATTGTAAACGTTAATATTTTGTAAAATT
CGCGTTAAATTTTTGTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCC
TTATAAATCAAAGAATAGACCGAGATAGGGTTGAGTGTGTTCCAGTTTGAACAAGAGT
CCACTATTAAGAACGTGGACTCAAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATG
GCCACTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTA

AATCGGAACCCTAAAGGGAGCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTG
GCGAGAAAGGAAGGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAG
CGTACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTC
GCGCCATTCGCCATTCAGGCTACGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCG
CTATTACGCCAGCTGGCGAAGGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCA
GGTTTTCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTGACTATA
GCCGAATTGGAGCTAGGCCTACGTAGCGCGGAGCTCCACCGCGGTGGCAGTACAATCT
GCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAG
TAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGA
ATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTATC
TGAGGGGACTAGGGTGTGTTAGGCGAAAAGCGGGGCTTCGGTTGTACGCGTTAGGAGT
CCCCTCAGGATATAGTAGTTTTGCTTTTGCATAGGGAGGGGAAATGTAGTCTTATGCAATA
CTCTTGTAGTCTTGCAACATGGTAACGATGAGTTAGCAACATGCCTTACAAGGAGAGAAAA
AGCACCGTGCATGCCGATTGGTGGAAAGTAAGGTGGTACGATCGTGCCTTATTAGGAAGGC
AACAGACGGGTCTGACATGGATTGGACGAACCACTGAATTCCGCATTGCAGAGATATTGTA
TTAAGTGCCTACTAGGTTGTGGAATTGTGAGCGCTCACAATCCACAGTCGACCCTAGTAT
ACAATAACGCCATTTGACCATTCACCACATTGGTGTGCACCTCCAAGCTTGGACAACTAC
CTACAGAGATTTAAAGCTCTAAGGTAATATAAAATTTTTAAGTGTATAATGTGTTAACTA
CGGATCCGCTCCATTAGGCCTACAATGGTGAAGCAAGTAGCCAACAGGGAAGGGTTGCA
AATATCATTGGGCACACCTATGATAATATTGATGAAGCAGACAGTATTCAGCAAGTAACTG
AGAGGTGGAAGCTCAAAGCCAAAGTCTAATGTGCAGTCAGGTGAATTTATTGAAAAATT
TGAGGCTCCTGGTGGACTAGGGTCGACTGTGGAATTGTGAGCGCTCACAATCCACAACCT
AGTTCAAAGAAGTCTCCTCAGGGATCCTAATTGTTTGTGTATTTTAGATTCCAACCAAGCTT
GCTGCCCCGCGGTAATACGACTCACTATAGGGCGAATTGGGTACATGAGATTAGATAAAAG
TAAAGTGATTAACAGCGCATTAGAGCTGCTTAATGAGGTGCGAATCGAAGGTTTAAACAACC
CGTAAACTCGCCAGAAGCTAGGTGTAGAGCAGCCTACATTGTATTGGCATGTAAAAAATA
AGCGGGCTTTGCTCGACGCCTTAGCCATTGAGATGTTAGATAGGCACCATACTCACTTTTGC
CCTTTAGAAGGGGAAAGCTGGCAAGATTTTTACGTAATAACGCTAAAAGTTTTAGATGTGC
TTTACTAAGTCATCGCGATGGAGCAAAAGTACATTTAGGTACACGGCCTACAGAAAAACAG
TATGAACTCTCGAAAATCAATTAGCCTTTTATGCCAACAAGGTTTTTACTAGAGAATGCA
TTATATGCACTCAGCGCTGTGGGGCATTTTACTTTAGGTTGCGTATTGGAAGATCAAGAGCA
TCAAGTCGCTAAAGAAGAAAGGGAAACACCTACTACTGATAGTATGCCGCCATTATTACGA
CAAGCTATCGAATTATTTGATCACCAAGGTGCAGAGCCAGCCTTCTTATTGGCCTTGAATT
GATCATATGCGGATTAGAAAAACAATTAATGTGAAAGTGGGTCCGCGTACAGCCGCGCG
CGTACGAAAAACAATTACGGGTCTACCATCGAGGGCCTGCTCGATCTCCCGGACGACGACG
CCCCGAAGAGGCGGGGCTGGCGGCTCCGCGCCTGTCTTTCTCCCGCGGGACACACGCG
CAGACTGTCGACGGCCCCCGACCGATGTCAGCCTGGGGGACGAGCTCCAATTAGACGGC
GAGGACGTGGCGATGGCGCATGCCGACGCGCTAGACGATTCGATCTGGACATGTTGGGG
GACGGGGATTCCCCGGTCCGGGATTTACCCCCACGACTCCGCCCTACGGCGCTCTGG
ATATGGCCGACTTCGAGTTTGAAGCAGATGTTACCGATGCCCTTGAATTGACGAGTACGG
TGGGTAGGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTCCCTTTAGTGAGGGTTAATTG
CGGGGCCGCAATTTCTAAGAAGTGAACACGGAAGGAGACAATACCGGAAGGAACCCGC
GCTATGACGGCAATAAAAAGACAGAATAAAAACGCACGGTGTGGGTCGTTTGTTCATAAAC
GCGGGGTTCCGTTCCAGGGCTGGCACTCTGTGATACCCACCGAGACCCATTGGGGCCA
ATACGCCGCGTTTCTTCTTTTCCCCACCCACCCCAAGTTCGGGTGAAGGCCAGGGC
TCGAGCCAACGTCGGGGCGGCAAGCCCTGCCATAGCCACGGGCCCCGTGGGTTAGGGAC
GGGTCCCCATGGGGAATGGTTTATGGTTCGTGGGGGCTAGCCGGCAGCTGCATTAATG

AATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGCTATTGGGCGCTCTTCCGCTTCCTCGCTC
ACTGACTCGCTGCGCTCGGTCGTTCCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGG
TAATACGGTTATCCACAGAATCAGGGGATAACGCGAGGAAAGAACATGTGAGCAAAGGCC
AGCAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTCCATAGGCTCCGCC
CCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACT
ATAAAGATAACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGC
CGTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCAC
GCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACC
CCCCGTTACGCCGACCCTGCGCCTTATCCGGTAACATCGTCTTGAGTCCAACCCGTA
GACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGT
AGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTA
TTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATC
CGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGAACGAGCAGATTACGCGC
AGAAAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGA
ACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATC
CTTTAAATTAATAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGAC
AGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATA
GTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCA
GTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCA
GCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCGCAACTTTATCCGCCTCCATCCAGTCT
ATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGT
TGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTAGCTCCG
GTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCC
TTCGGTCTCCGATCGTTGTCAGAAGTAAGTTGGCCGAGTGTATCACTCATGGTTATGGC
AGCACTGCATAATTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTA
CTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCGGCGTCAA
TACGGGATAATACCGCGCCACATAGCAGAAGTTAAAAGTGCTCATCATTGGAAAACGTTCT
TCGGGGCGAAAACCTCAAGGATCTTACCCTGTTGAGATCCAGTTTCGATGTAACCCACTCG
TGCACCCAACTGATCTTCAGCATCTTTACTTTACCAGCGTTTCTGGGTGAGCAAAAACAG
GAAGGCAAATGCCGCAAAAAAGGAATAAGGGCGACACGAAATGTTGAATACTCATACT
TCTTCCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT
TGAATGTATTTAGAAAAATAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAGTGCCA
CCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAG
GCCCTTCGTC

pTRE3G_Gal4 DNA sequence:

TRE3G promoter

TetO operator sites*Gal4* ORF

Small t intron

SV40 Nuclear Localization Signal

SV40 poly-adenylation signal

ColE1 / pUC origin of replication

AmpR ORF / *AmpR* promoter

CTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGAAGAGTTTACTCCCTATCAGTGATA
 GAGAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGAACGTATAAGGAGTTTACTCCCT
 ATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGAACGTATCTACAG
 TTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAAC
 GTATAAGCTTTAGGCGTGTACGGTGGGCGCCTATAAAAGCAGAGCTCGTTTAGTGAACCGT
 CAGATCGCCTGGAGCAATTCCACAACACTTTTGTCTTATACCAACTTCCGTACCACTTCCCTA
 CCCTCGTAAAGATGAAGCTACTGTCTTCTATCGAACAAGCATGCGATATTTGCCGACTTAAA
 AAGCTCAAGTGCTCCAAAGAAAAACCGAAGTGCGCCAAGTGTCTGAAGAACAACCTGGGAG
 TGTCGCTACTCTCCAAAACCAAAGGTCTCCGCTGACTAGGGCACATCTGACAGAAGTGG
 AATCAAGGCTAGAAAGACTGGAACAGCTATTTCTACTGATTTTCTCGAGAAGACCTTGAC
 ATGATTTTAAAATGGATTCTTTACAGGATATAAAAGCATTGTTAACAGGATTATTTGTACA
 AGATAATGTGAATAAAGATGCCGTCACAGATAGATTGGCTTCAGTGGAGACTGATATGCCT
 CTAACATTGAGACAGCATAGAATAAGTGCGACATCATCATCGGAAGAGAGTAGTAACAAAG
 GTCAAAGACAGTTGACTGTATCGATTGACTCGGCAGCTCATCATGATAACTCCACAATTCCG
 TTGGATTTTATGCCCAGGGATGCTCTTCATGGATTTGATTGGTCTGAAGAGGATGACATGTC
 GGATGGCTTGCCCTTCTGAAAACGGACCCCAACAATAATGGGTTCTTTGGCGACGGTTCTC
 TCTTATGTATTCTTCGATCTATTGGCTTTAAACCGGAAAATTACACGAACTCTAACGTTAACA
 GGCTCCCGACCATGATTACGGATAGATACACGTTGGCTTCTAGATCCACAACATCCCGTTTA
 CTTCAAAGTTATCTCAATAATTTTACCCCTACTGCCCTATCGTGCACCTACCGACGCTAATG
 ATGTTGTATAATAACCAGATTGAAATCGCGTCAAGGATCAATGGCAAATCCTTTTTAACTG
 CATATTAGCCATTGGAGCCTGGTGTATAGAGGGGAATCTACTGATATAGATGTTTTTACT
 ATCAAATGCTAAATCTCATTGACGAGCAAGGTCTTCGAGTCAGGTTCCATAATTTTGGTG
 ACAGCCCTACATCTTCTGTCGCGATATACACAGTGGAGGCAGAAAACAAATACTAGCTATAA
 TTTTACAGCTTTTCCATAAGAATGGCCATATCATTGGGCTTGAATAGGGACCTCCCCTCGTC
 CTTCAGTGATAGCAGCATTCTGGAACAAAGACGCCGAATTTGGTGGTCTGTCTACTCTTGGG
 AGATCCAATTGTCCCTGCTTTATGGTCGATCCATCCAGCTTCTCAGAATACAATCTCCTCCC
 TTCTTCTGTGACGATGTGCAGCGTACCACAACAGGTCCACCATATATCATGGCATCATTG

AACAGCAAGGCTCTTACAAGTTTTACAAAAATCTATGAACTAGACAAAACAGTAACTGCA
 GAAAAAAGTCCTATATGTGCAAAAAAATGCTTGATGATTTGTAATGAGATTGAGGAGGTTT
 CGAGACAGGCACCAAAGTTTTTACAAATGGATATTTCCACCACCGCTCTAACCAATTTGTTG
 AAGGAACACCCTTGGCTATCCTTACAAGATTGCAACTGAAGTGGAACAGTTGTCTCTTAT
 CATTTATGTATTAAGAGATTTTTTACTAATTTTACCCAGAAAAAGTCACA ACTAGAACAGGA
 TCAAAATGATCATCAAAGTTATGAAGTTAAACGATGCTCCATCATGTTAAGCGATGCAGCAC
 AAAGA ACTGTTATGTCTGTAAGTAGCTATATGGACAATCATAATGTCACCCCATATTTTGCCT
 GGAATTGTTCTTATTACTTGTTCATGCAGTCTAGTACCATAAAGACTCTACTCTCAA ACT
 CAAAATCGAATGCTGAGAATAACGAGACCGCACAATTATTACAACAAATTAACACTGTTCTG
 ATGCTATTAAAAAA ACTGGCCACTTTTAAAATCCAGACTTGTGAAAAATACATTCAAGTACT
 GGAAGAGGTATGTGCGCGTTTCTGTTATCACAGTGTGCAATCCATTACCGCATATCAGTT
 ATAACAATAGTAATGGTAGCGCCATTAAAATATTGTGCGTTCTGCAACTATCGCCCAATAC
 CCTACTCTCCGGAGGAAAATGTCAACAATATCAGTGTTAAATATGTTTCTCCTGGCTCAGTA
 GGGCCTTACCTGTGCCATTGAAATCAGGAGCAAGTTTCAGTGATCTAGTCAAGCTGTTATC
 TAACCGTCCACCCTCTCGTAACTCTCCAGTGACAATACCAAGAAGCACACCTTCGCATCGCTC
 AGTCACGCCTTTTCTAGGGCAACAGCAACAGCTGCAATCATTAGTGCCACTGACCCCGTCTG
 CTTTGTGTTGGTGCCCAATTTTAAATCAAAGTGGAATATTGCTGATAGCTCATTGTCCTTCA
 CTTTACTAACAGTAGCAACGGTCCGAACCTCATAACA ACTCAACAAATTTCTCAAGCGCTTT
 CACAACCAATTGCCTCCTCTAACGTTTCATGATAACTTCATGAATAATGAAATCACGGCTAGTA
 AAATTGATGATGGTAATAATTCAAACCACTGTCACCTGGTTGGACGGACCAA ACTGCGTAT
 AACGCGTTTGAATCACTACAGGGATGTTTAAATACCACTACAATGGATGATGTATATAACTA
 TCTATTGATGATGAAGATACCCACCAAACCCAAAAAAGAGTAAGATCCAATGTA ACTGT
 ATTCAGCGATGACGAAATCTTAGCTATTGTAATACTCTAGAGGATCTTTGTGAAGGAACCT
 TACTTCTGTGGTGTGACATAATTGGACAAACTACCTACAGAGATTTAAAGCTCTAAGGTA AA
 TATAAAATTTTTAAGTGTATAATGTGTTAAACTACTGATTCTAATTGTTTGTGATTTTTAGATT
 CCAACCTATGGA ACTGATGAATGGGAGCAGTGGTGGAAATGCCTTTAATGAGGAAAACCTGT
 TTTGCTCAGAAGAAATGCCATCTAGTGATGATGAGGCTACTGCTGACTCTCAACATTCTACT
 CCTCCAAAAAAGAAGAGAAAGGTA GAAGACCCCAAGGACTTTTCTTCAGAATTGCTAAGTT
 TTTTGAGTCATGCTGTGTTTAGTAATAGA ACTCTTGCTTGCTTTGCTATTTACACCACAAAGG
 AAAAAGCTGCACTGCTATACAAGAAAATTATGGAAAAATATTCTGTAACCTTTATAAGTAGG
 CATAACAGTTATAATCATAACATACTGTTTTTTCTTACTCCACACAGGCATAGAGTGTCTGCT
 ATTAATAACTATGCTCAAAAATTGTGTACCTTTAGCTTTTTAATTTGTAAGGGGTTAATAAG
 GAATATTTGATGTATAGTGCCTTACTAGAGATCATAATCAGCCATACCACATTTGTAGAGG
 TTTTACTTGCTTTAAAAAACCTCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAA
 TTGTTGTTGTTAACTTGTATTATGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAA
 ATTTACAAATAAAGCATTTTTTTCTACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGT
 ATCTTATCATGTCTGCGGCTCTAGAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGG
 CGGTTTGCATATTGGGCGCTCTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGTTTCG
 GCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGG
 GATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAA
 GGCCGCGTTGCTGGCGTTTCCATAGGCTCCGCCCCCTGACGAGCATCACA AAAATCGAC
 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGG
 AAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCCTTTCT
 CCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGG
 TCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTAGCCCGACCGCTGCGCCTTA
 TCCGTA ACTATCGTCTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGC
 CACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTG

GTGGCCTAACTACGGCTACACTAGAAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA
 GTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCG
 GTGGTTTTTTTGTGGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAA GAAGATCC
 TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACACGTTAAGGGATTTTGG
 TCATGAGATTATCAAAAAGGATCTTACCTAGATCCTTTTAAATTA AAAATGAAGTTTTAAAT
 CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA
 CCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATA
 ACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCAC
 GCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAA
 GTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTA
 AGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTG
 ACGCTCGTCGTTTGGTATGGCTTCATTAGCTCCGGTTCCCAACGATCAAGGCGAGTTACAT
 GATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTCAGAAGT
 AAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCAT
 GCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGT
 GTATGCGGCGACCGAGTTGCTCTTGGCCGGCGTCAATACGGGATAATACCGCGCCACATAG
 CAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACCTCAAGGATCT
 TACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTT
 TTACTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAATGCCGCAAAAAAGGG
 AATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCAT
 TTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAA
 TAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATC
 ATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTCGTCTTCAAGAATTC

pTRE3G_UbGFP DNA sequence:

PTRE3G promoter

TetO operator sites

Ubiquitin tag

eGFP ORF

SV40 poly-adenylation signal

f1 origin of replication

SV40 promoter / SV40 origin of replication

NeoR / KanR ORF

HSV TK poly-adenylation signal

ColE1 / pUC origin of replication

TAGTTATGAGTTTACTCCCTATCAGTGATAGAGAACGTATGAAGAGTTTACTCCCTATCAGT
 GATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGAACGTATAAGGAGTTTACT
 CCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGAACGTATCT
 ACAGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGA
 GAACGTATAAGCTTTAGGCGTGTACGGTGGGCGCCTATAAAAGCAGAGCTCGTTTAGTGAA
 CCGTCAGATCGCCTGGAGCAATTCCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTT
 CCTACCCTCGTAAACTAGCGCTACCGGACTCAGATCTCGAGCTCAAGCTTCGAATTCACCAT
 GCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCTCGAGGTTGAGCCAGTGAC
 ACCATTGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGA
 GGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCTGTCTGACTACAACATCCA
 GAAAGAGTCCACCCTGCACCTGGTACTCCGTCTCAGAGGTGTGGTGGGGAAGCTTGGTCTGA
 CAGGATCCACCGGTGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTG
 CCCATCCTGGTGCAGCTGGACGGCGACGTAACGGCCACAAGTTCAGCGTGTCCGGCGAG
 GGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGC
 TGCCCGTGCCCTGGCCACCTCGTGACCACCTGACCTACGGCGTGCAGTGCTTCAGCCGC
 TACCCCGACCACATGAAGCAGCAGACTTCTCAAGTCCGCCATGCCGAAGGCTACGTCCA
 GGAGCGCACCATCTTCTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTC
 GAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGC
 AACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATCATGGCCG
 ACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCA
 GCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCT
 GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGC
 GATCACATGGTCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGC
 TGTACAAGTAAAGCGGCCGCGACTCTAGATCATAATCAGCCATACCACATTTGTAGAGGTTT
 TACTTGCTTTAAAAAACCTCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTG
 TTGTTGTTAACTTGTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATT
 TCACAAATAAAGCATTITTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATC
 TTAAGGCGTAAATTGTAAGCGTTAATATTTTGTAAAATTTCGCGTAAATTTTTGTAAATCA
 GCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAGAATAGACC
 GAGATAGGGTTGAGTGTTGTTCCAGTTTGAACAAGAGTCCACTATTAAGAACGTGGACT
 CCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCAC
 CTAATCAAGTTTTTTGGGGTGCAGGTGCCGTAAGCACTAAATCGGAACCCTAAAGGGAGC
 CCCCATTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGAAGAA
 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC
 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCAGGTGGCACTTTTCGGGGAAT
 GTGCGCGGAACCCCTATTTGTTTATTTTTCTAATAACATTCAAATATGTATCCGCTCATGAGA
 CAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTCCTGAGGCGGAAAGAA
 CCAGCTGTGGAATGTGTGTGAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAG
 AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCC
 CCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCC GCCCT
 AACTCCGCCATCCCGCCCCTAACTCCGCCAGTTCGCCCCATTCTCGGCCCATGGCTGACT
 AATTTTTTTTATTATGCAGAGGCCGAGGCCGCTCGGCCTCTGAGCTATTCCAGAAGTAGT
 GAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAGATCGATCAAGAGACAGGATGAGGA
 TCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGA

GGCTATTGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCG
GCTGTCAGCGCAGGGGCGCCCGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAAT
GAACTGCAAGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCA
GCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCG
GGGCAGGATCTCTGTCATCTCACCTTGCTCTGCCGAGAAAGTATCCATCATGGCTGATGC
AATGCGGCGGCTGCATACGTTGATCCGGCTACCTGCCATTGACCACCAAGCGAAACATC
GCATCGAGCGAGCACGTA CTGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACG
AAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCCGACGGCTCAAGGCGAGCATGCCCG
ACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGA
TGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGAC
ATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTC
TCGTGCTTTACGGTATCGCCGCTCCCATTGCGAGCGCATCGCCTTCTATCGCCTTCTTGACG
AGTTCCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCAACCTGCCA
TCAGGATTTGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCG
GGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCACCCT
AGGGGGAGGCTAACTGAAACACGGAAGGAGACAATACCGGAAGGAACCCGCGCTATGAC
GGCAATAAAAAGACAGAATAAAAACGCACGGTGTGGGTCGTTGTTGATAAACGCGGGGTT
CGGTCCAGGGCTGGCACTCTGTGATACCCACCGAGACCCATTGGGGCCAATACGCC
GCGTTTCTCCTTTTCCCCACCCACCCCAAGTTCGGGTGAAGGCCAGGGCTCGCAGCC
AACGTCGGGGCGGCAGGCCCTGCCATAGCCTCAGGTTACTCATATATACTTTAGATTGATTT
AAAACCTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAA
AATCCCTAACGTGAGTTTTCGTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT
CTTCCTTGAGATCCTTTTTTTCTGCGGTAATCTGCTGCTTGCAAACAAAAAACCACCGCTAC
CAGCGGTGGTTTGTGGCCGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTC
AGCAGAGCGCAGATACCAATACTGTCCTTAGTGTAGCCGTAGTTAGGCCACCACTTCAA
GAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCA
GTGGCGATAAGTCGTGTCTTACCGGGTGGACTCAAGACGATAGTTACCGGATAAGGCGCA
GCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACAC
CGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAA
GGCGGACAGGTATCCGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTC
CAGGGGGAAACGCCTGGTATCTTTATAGTCTGTCGGGTTTCGCCACCTCTGACTTGAGCGT
CGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAACGCCAGCAACGCGGCC
TTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTCTGCGTTATCCCCTG
ATTCTGTGGATAACCGTATTACCGCCATGCAT

A.5.2. Fluorescence channel images corresponding to the single-cell signals in figures 5.14 to 5.16

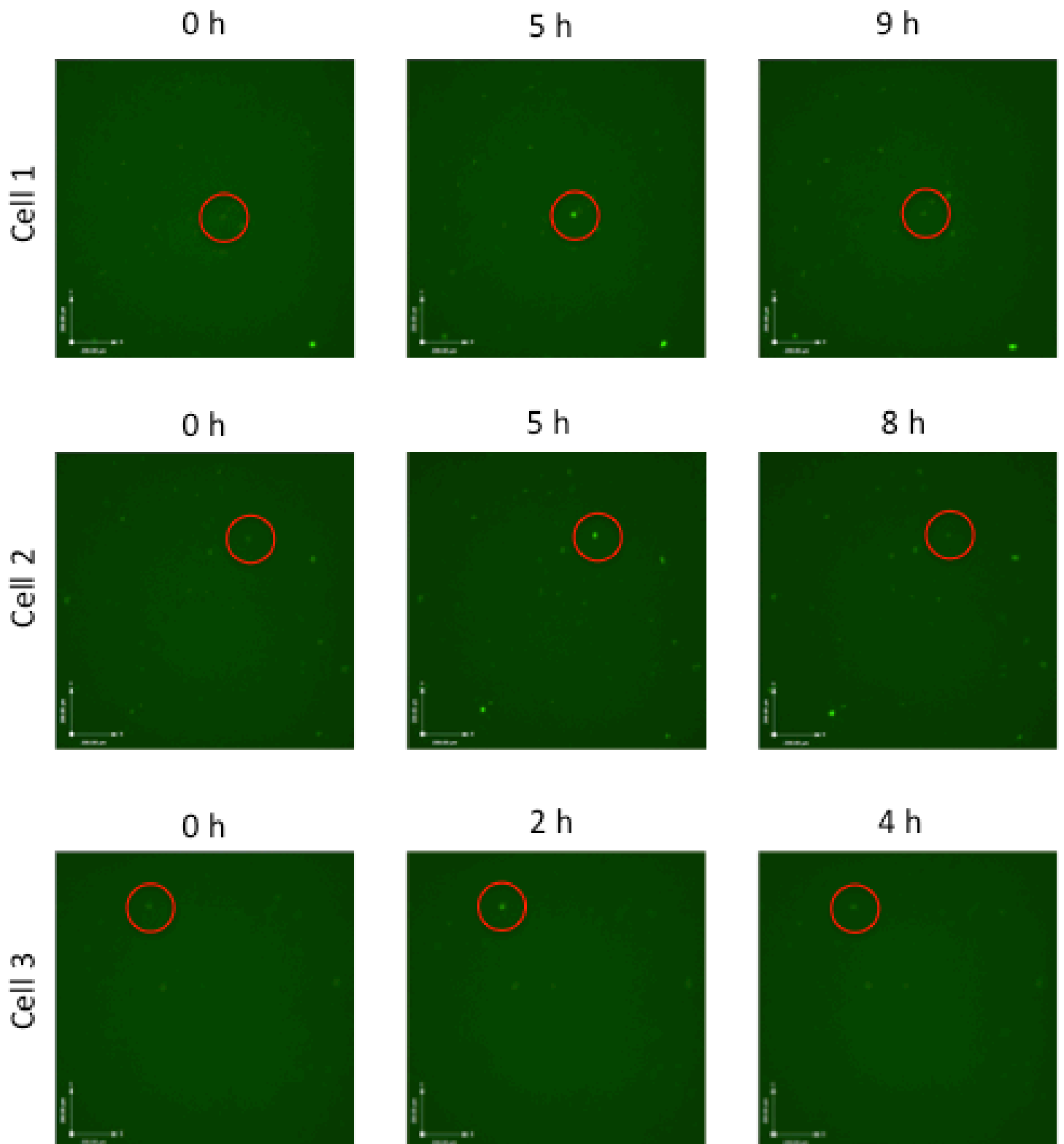


Figure A.5.1. Fluorescence channel images corresponding to the single-cell signals illustrated in figure 5.14. Red circles highlight the respective cells in each image. The time frames indicated above each

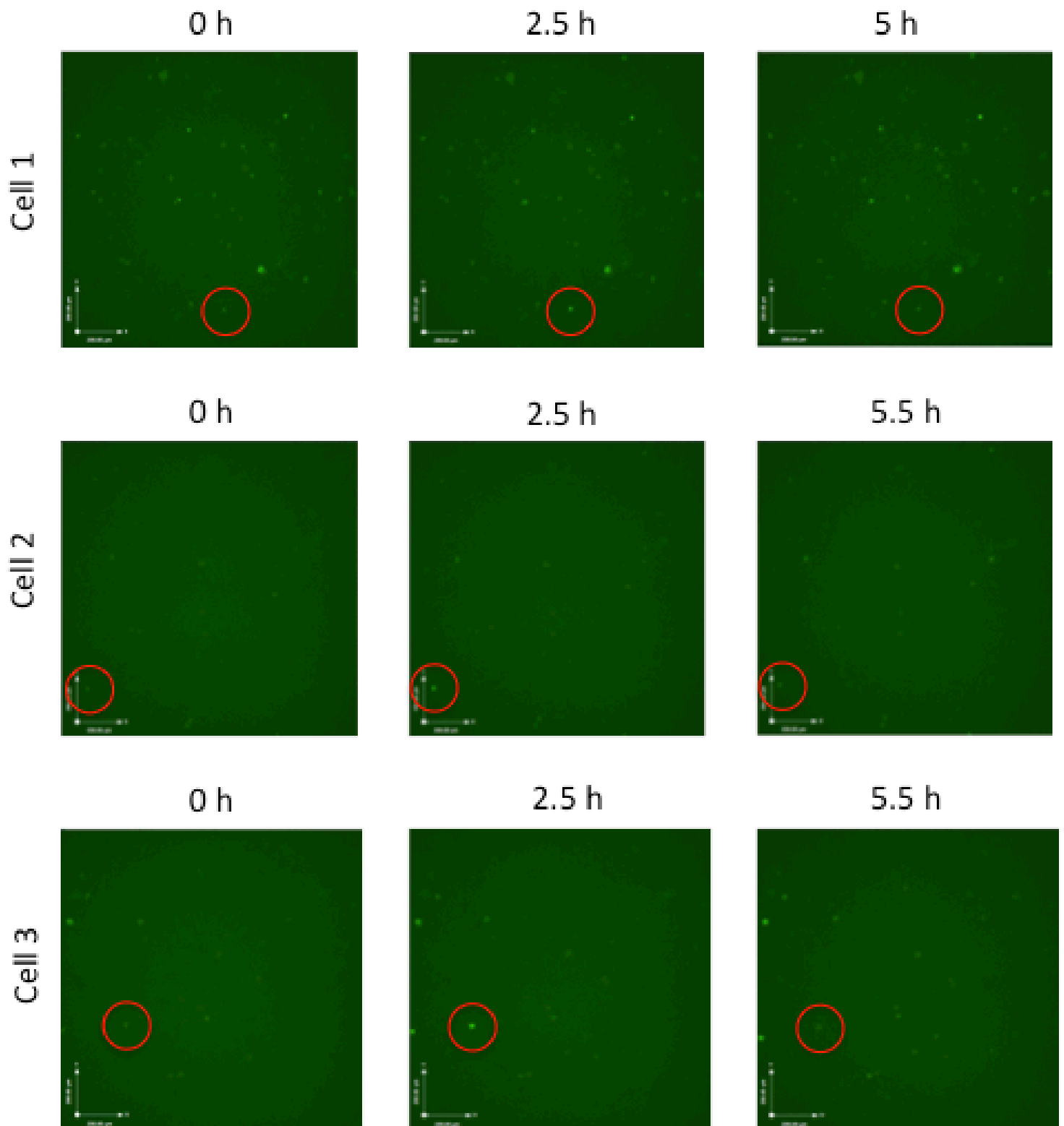


Figure A.5.2. Fluorescence channel images corresponding to the single-cell signals illustrated in figure 5.15. Red circles highlight the respective cells in each image. The time frames indicated above each picture are in accordance with the x-axis in figure 5.15.

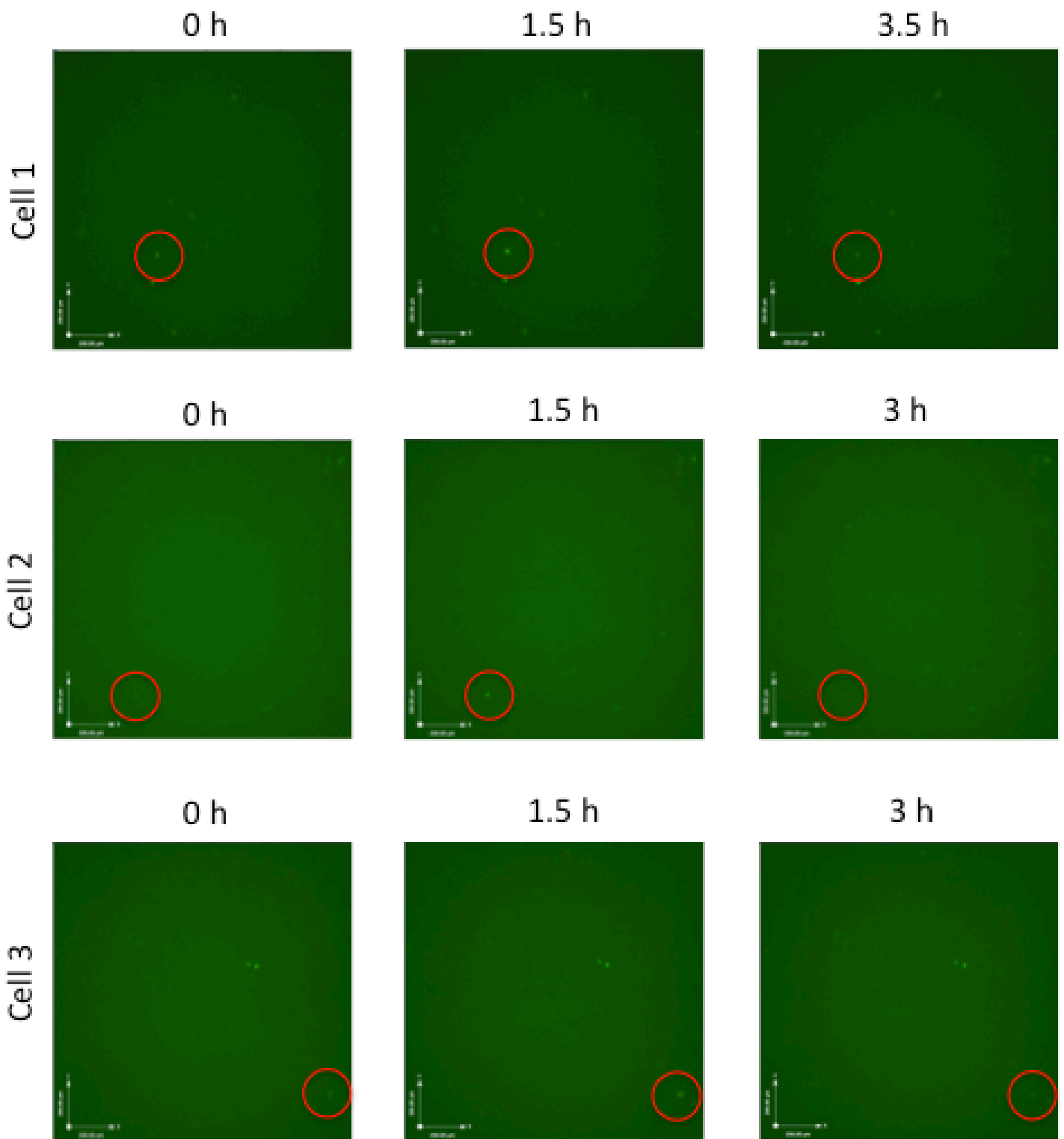


Figure A.5.3. Fluorescence channel images corresponding to the single-cell signals illustrated in figure 5.16. Red circles highlight the respective cells in each image. The time frames indicated above each picture are in accordance with the x-axis in figure 5.16.