



**Design, Mathematical Modelling, Construction and
Testing of Synthetic Gene Network Oscillators to
Establish *Roseobacter* Clade Bacteria and the
Protozoan *Trypanosoma brucei* as Synthetic
Biology Chassis.**

Yanika Borg

June 2015

University College London

A thesis submitted in partial fulfillment of the requirements of the degree of
Doctor of Philosophy

Declaration

I, Yanika Borg confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:.....

Abstract

The aim of this project is to establish *Roseobacter* marine bacteria and *Trypanosoma brucei* (*T. brucei*) protozoa as synthetic biology chassis. This work addresses the gap within synthetic biology resulting from the limited choice of host cells available for use in practice. This was done by developing synthetic bacterial and trypanosomal genetic regulatory networks (GRNs) which function as an oscillator as well as by developing the necessary protocols and set-ups to allow for the analysis of GRN dynamics within the host.

Roseobacter clade bacteria are naturally found in diverse oceanic habitats and have an important ecological role in balancing global carbon levels. This makes *Roseobacter* an ideal chassis for future efforts to apply synthetic biology to bioremediation and geo-engineering challenges. The aim of this investigation was to establish straight-forward molecular biology procedures in *Roseobacter* bacteria followed by characterisation and modelling of an *E. coli* oscillator in *Roseobacter*. Results showed that *Roseobacter* synthetic biology is non-trivial.

Protozoa are exploited as host cells for industrial production of biotherapeutics due to fast doubling times and host proteins' mammalian-like post-translational glycosylation. As an established model organism for studying protozoa, *T. brucei* provided a test case for establishing synthetic biology in this phylum for the first time. *T. brucei* is highly divergent from eukaryotes commonly used in synthetic biology and possesses a sophisticated genomic machinery to evade host immune systems. The establishment of standard synthetic biology approaches in mathematical modelling and gene network design in *T. brucei* will underpin application of synthetic biology to enhance the industrial capability of the protozoa as a chassis and to probe its pathobiology. This investigation involved design and assembly of a Goodwin oscillator, followed by characterisation and modelling of the network and the development of a novel experimental set-up for live-cell imaging of single motile trypanosomes. Results showed that *T. brucei* is a promising novel synthetic biology chassis.

Signature:.....

Acknowledgements

I would like to express my deep gratitude towards my supervisors Dr Darren Nesbeth and Prof. Alexey Zaikin, for giving me, a Mathematics and Statistics graduate, the opportunity to pursue a PhD within the field of synthetic biology. Throughout the PhD, they have provided me with guidance and given me plenty of opportunities to develop and grow as a scientist. I would also like to thank Dr Sam Alsford at the London School of Hygiene and Tropical Medicine, for allowing me to carry out some of the research in his lab and for making suggestions which helped to strengthen my work.

I would also like to thank Ms Aurelija Grigonyte, who assisted with some of the experiments in Chapter 3, for her companionship and help.

Last but not least, I would like to thank my family for supporting me unconditionally and for always making me smile.

Contents

1	Introduction	25
1.1	Synthetic biology and its application	25
1.1.1	Synthetic biology adopts a multidisciplinary approach: The role of wet lab and mathematical analysis in synthetic biology	26
1.1.2	The democratisation of synthetic biology and the DIYbio movement	27
1.2	The choice of chassis for synthetic GRNs	28
1.2.1	First generation synthetic biology chassis	28
1.2.1.1	Gene structure, expression and regulatory mechanisms in prokaryote and eukaryote chassis	30
1.2.1.1.1	Initiation of transcription and transcription.	30
1.2.1.1.2	Post-transcriptional modification.	31
1.2.1.1.3	Initiation of translation, translation and post-translational modifications.	31
1.2.1.1.4	Transcriptional regulatory mechanisms in prokaryote and eukaryote chassis.	32
1.2.1.1.5	Post-transcriptional regulatory mechanisms in prokaryote and eukaryote chassis.	33
1.2.1.2	Transcriptional and post-transcriptional regulation mechanisms in GRNs in commonly used synthetic biology chassis	34
1.2.1.2.1	Transcriptional regulation mechanisms.	34
1.2.1.2.2	Applications of transcriptional regulation.	35
1.2.1.2.3	Post-transcriptional regulation mechanisms.	35
1.2.1.3	Synthetic GRNs in commonly used synthetic biology chassis	36
1.2.1.4	Oscillatory synthetic GRNs in commonly used synthetic biology chassis	36
1.2.1.4.1	Characteristics of proof-of-principle synthetic oscillator GRNs.	40
1.2.1.4.2	The scope of mathematical models in the development of synthetic oscillator GRNs.	40
1.2.1.5	Design principles for oscillatory synthetic GRNs	41
1.2.1.5.1	The role of negative feedback.	41
1.2.1.5.2	The role of time-delay.	42
1.2.1.5.3	The role of non-linear dynamics.	43
1.2.1.5.4	The role of positive feedback.	44
1.2.1.5.5	The role of noise.	44
1.2.1.6	Construction and characterisation of oscillatory synthetic GRNs	45
1.2.1.6.1	Construction of oscillator plasmids.	45
1.2.1.6.2	Insertion of oscillator networks into cells.	45
1.2.1.6.3	Characterisation, microscopy and cell imaging.	46

1.2.1.6.4	Microscopy imaging set-up.	47
1.2.1.6.5	Post-imaging analysis.	47
1.2.1.7	Mathematical modelling of oscillatory synthetic GRNs	48
1.2.1.7.1	Deterministic modelling of network dynamics.	48
1.2.1.7.1.1	Ordinary differential equations.	48
1.2.1.7.1.2	Simulating deterministic ODEs.	50
1.2.1.7.1.3	Assumptions of deterministic ODEs.	50
1.2.1.7.2	Stochastic modelling of network dynamics.	51
1.2.1.7.2.1	Master and Langevin equations.	51
1.2.1.7.2.2	Stochastic Simulation Algorithms.	52
1.2.1.7.2.3	The Gillespie Stochastic Simulation Algorithm. . .	53
1.2.1.7.2.4	Variations of the Gillespie algorithm.	54
1.2.1.7.3	Qualitative analysis of dynamics.	54
1.2.1.7.3.1	Bifurcation analysis.	54
1.2.1.7.3.2	Hopf bifurcations.	55
1.2.1.7.3.3	Stability and robustness analysis.	56
1.2.1.7.4	Parameter estimation of GRN biochemical reaction rates.	57
1.2.2	The value of broadening the range of synthetic biology chassis	57
1.2.2.1	The <i>Roseobacter</i> marine bacteria genus	59
1.2.2.1.1	Novel capabilities <i>Roseobacter</i> offers as a synthetic biology chassis.	59
1.2.2.1.2	The challenges of using <i>Roseobacter</i> bacteria in synthetic biology.	60
1.2.2.2	The <i>Trypanosomatida</i> order	61
1.2.2.2.1	<i>Leishmania tarentolae</i> in bioprocessing.	61
1.2.2.2.2	<i>Trypanosoma brucei</i> as model <i>Trypanosomatida</i>	62
1.2.2.2.2.1	The life cycle of <i>T. brucei</i>	62
1.2.2.2.2.2	Transcription in <i>T. brucei</i>	64
1.2.2.2.2.3	Translation in <i>T. brucei</i>	64
1.2.2.2.2.4	Variant surface glycoproteins as an antigen variation mechanism in <i>T. brucei</i>	64
1.2.2.2.3	Novel capabilities <i>T. brucei</i> offers as a synthetic biology chassis.	65
1.2.2.2.4	The challenges of using <i>T. brucei</i> as a chassis in synthetic biology.	66
1.3	Aims and Objectives	66
1.3.1	Aim 1: Establishing <i>Roseobacter</i> clade marine bacteria as a synthetic biology chassis by identifying reliable molecular biology procedures for construction of a synthetic gene network.	67
1.3.2	Aim 2: Establish <i>T. brucei</i> as a synthetic biology chassis by designing, modelling, constructing and measuring a Goodwin oscillator encoded by a novel synthetic gene network.	67
2	Materials and Methods	69
2.1	<i>Roseobacter</i> marine bacteria strains and culturing conditions	69
2.1.1	<i>Roseobacter</i> bacterial strains	69
2.1.2	<i>Roseobacter</i> culturing conditions	69
2.1.3	Glycerol stock preparation for <i>Roseobacter</i> strains	69
2.1.4	Antibiotics used for culturing of <i>Roseobacter</i> strains	70

2.2	Recombinant DNA techniques in <i>E. coli</i>	70
2.2.1	Plasmids used for construction of the oscillator networks	70
2.2.2	<i>E. coli</i> strains used in <i>Roseobacter</i> investigations and the construction of the oscillator networks	70
2.2.3	<i>E. coli</i> culturing conditions	71
2.2.4	Glycerol stock preparation for <i>E. coli</i> strains	71
2.2.5	Antibiotics used for culturing of <i>E. coli</i> strains	71
2.2.6	Generating chemically-competent <i>E. coli</i> W3110 cells	71
2.2.7	Chemical transformation of chemically-competent <i>E. coli</i> cells	72
2.2.7.1	Transformation of chemically competent <i>E. coli</i> W3110 cells	72
2.2.7.2	Transformation of chemically competent <i>E. coli</i> One Shot® TOP10 cells	72
2.2.7.3	Transformation of chemically competent <i>E. coli</i> XL10-Gold® cells	73
2.2.8	Plasmid DNA elution and purification from <i>E. coli</i> cultures, plasmid DNA digests, PCR reactions and gel electrophoresis runs using commercial kits	73
2.2.9	Plasmid DNA purification from plasmid DNA digests using phenol-chloroform extraction in preparation for <i>T. brucei</i> transfection	74
2.2.10	Plasmid DNA digests	74
2.2.10.1	Analytical DNA digests of less than 1 μ g of plasmid DNA	74
2.2.10.2	Preparative DNA digests of 1-10 μ g of plasmid DNA	75
2.2.11	Dephosphorylation of 5'-ends of plasmid DNA	75
2.2.11.1	Dephosphorylation using Calf Intestinal Alkaline Phosphatase enzyme (NEB)	75
2.2.11.2	Dephosphorylation using Calf Intestinal Alkaline Phosphatase enzyme (Promega)	75
2.2.11.3	Dephosphorylation using Shrimp Alkaline Phosphatase (NEB)	76
2.2.12	Agarose gel electrophoresis	76
2.2.12.1	Preparation and electrophoresis of agarose gel	76
2.2.12.2	Electrophoresis analysis	76
2.2.13	Polymerase Chain Reaction	76
2.2.13.1	Controls used in Polymerase Chain Reaction	77
2.2.14	Ligation of two or more DNA fragments	77
2.2.14.1	Directional, non-directional & three-fragment ligations using Quick T4 DNA ligase	77
2.2.14.1.1	Controls used in ligation.	78
2.2.14.2	Gibson assembly®	79
2.2.14.2.1	Controls used in Gibson assembly®	79
2.2.15	Sequencing Plasmid DNA	81
2.3	<i>T. brucei</i> strains and culturing conditions	81
2.3.1	<i>T. brucei</i> strains	81
2.3.2	BSF <i>T. brucei</i> culturing conditions	82
2.3.3	Cryopreservation and revival of BSF <i>T. brucei</i>	82
2.3.4	PCF <i>T. brucei</i> culturing conditions	82
2.3.5	Cryopreservation and revival of PCF <i>T. brucei</i>	83
2.3.6	Antibiotics used for the culturing of <i>T. brucei</i> strains	83
2.3.7	Tetracycline induction of <i>T. brucei</i> cells transfected with the oscillator plasmid	83
2.3.8	Blocking of <i>T. brucei</i> protein synthesis using cycloheximide	83
2.4	Recombinant DNA techniques in <i>T. brucei</i>	84
2.4.1	Stable DNA transfection of BSF <i>T. brucei</i>	84

2.4.2	Stable DNA transfection of PCF <i>T. brucei</i>	84
2.4.3	<i>T. brucei</i> immobilisation using chemical agents and microfluidic devices	85
2.4.3.1	Microfluidic laser cage devices	85
2.4.3.2	Agarose block trap devices	85
2.4.3.3	Agar and agarose solutions	86
2.4.3.4	Xanthan gum	86
2.4.3.5	Cygel™ thermoreversible gel	86
2.5	<i>T. brucei</i> protein analysis	86
2.5.1	<i>T. brucei</i> protein isolation	86
2.5.2	SDS-PAGE gel analysis of <i>T. brucei</i> proteins	87
2.5.2.1	Preparation of resolving and stacking gels	87
2.5.2.2	Running the SDS-PAGE gel	87
2.5.3	Western blotting	88
2.5.3.1	Semi-dry protein transfer	88
2.5.3.2	Blocking of non-specific binding at the membrane	89
2.5.3.3	Addition of antibodies for protein detection in western blots	89
2.5.4	Coomassie stain for protein visualisation following SDS-PAGE	90
2.5.5	Densitometry plot	90
2.5.6	Immunofluorescence assay	90
2.5.6.1	Cell fixation process	90
2.5.6.2	Addition of antibodies for protein detection in immunofluorescence assays	91
2.5.6.3	Visualisation	91
2.6	<i>T. brucei</i> live cell imaging via fluorescent and brightfield microscopy	91
2.6.1	DEFRA licensing for <i>T. brucei</i> live cell imaging	91
2.6.2	Imaging set-up using live <i>T. brucei</i> cells	91
2.6.3	Imaging conditions	92
2.6.4	Background correction	92
2.6.5	Quantification of fluorescence intensity profiles	92
2.7	Mathematical Modelling of the Goodwin oscillator GRN in a <i>T. brucei</i> host	93
2.7.1	The Gillespie Direct Method stochastic simulation algorithm	93
2.7.1.1	Copasi 4.12.65 for coding the oscillator GRN in SBML format	93
2.7.1.2	ABC-Sysbio 2.05 for running the Gillespie stochastic simulation algorithm	94
2.7.2	MATLAB R2012a for analysis of Gillespie simulations	95
2.7.3	Data smoothing by application of imputation or a moving average filter	95
2.7.4	Cubic spline data interpolation	95
2.7.5	Fitting of linear growth trends	96
2.7.6	Phase space reconstruction	96
3	Establishing <i>Roseobacter</i> marine bacteria as a synthetic biology chassis	97
3.1	Long-term viability of <i>Roseobacter</i> glycerol stocks	97
3.1.1	Preparation of glycerol stocks	97
3.1.2	Measurement of long-term viability of glycerol stocks using optical density and colony count measurements	98
3.1.3	The viability of <i>Roseobacter</i> glycerol stocks in liquid and static cultures	98
3.2	Growth of <i>Roseobacter</i> strains in response to varying Chloramphenicol concentrations in the culture medium	100

3.2.1	Preparation of <i>Roseobacter</i> strains in culture medium with varying Chloramphenicol concentrations	100
3.2.2	The Chloramphenicol concentration threshold for cell growth inhibition of <i>Roseobacter</i> strains	101
3.3	Transformation of <i>Roseobacter</i> clade bacteria using chemical transformation and electroporation methods	101
3.3.1	Transformation based on variations of electrocompetence and electroporation methods	103
3.3.1.1	Efficiency of electroporation-based transformation as a function of plasmid mass	103
3.3.1.1.1	Preparation and electroporation of electrocompetent cells.	103
3.3.1.1.2	Effectiveness of electroporation-based approach using varying plasmid mass for <i>Roseobacter</i> transformation.	105
3.3.1.2	Efficiency of electroporation-based transformation as a function of voltage	106
3.3.1.2.1	Preparation and electroporation of electrocompetent cells.	106
3.3.1.2.2	Effectiveness of electroporation-based approach using varying voltage for <i>Roseobacter</i> transformation.	106
3.3.1.3	Efficiency of electroporation-based transformation based on a modified electroporation protocol and a broader range of plasmid ori sequences	107
3.3.1.3.1	Preparation and electroporation of electrocompetent cells.	107
3.3.1.3.2	Effectiveness of modified electrocompetence and broader plasmid range approach to <i>Roseobacter</i> transformation.	108
3.3.2	Transformation using the Piekarski et al. electroporation method	108
3.3.2.1	Preparation and electroporation of electrocompetent cells	109
3.3.2.2	Effectiveness of the Piekarski et al. approach to <i>Roseobacter</i> transformation	109
3.3.3	Transformation using the heat shock method	111
3.3.3.1	Preparation and transformation using the heat shock method of chemically competent marine bacterial cells	111
3.3.3.2	Effectiveness of the heat-shock method to <i>Roseobacter</i> transformation	111
3.4	Discussion	112
4	Design and construction of the stable and unstable trypanosomal Goodwin oscillators	114
4.1	The stable and unstable Goodwin oscillators	114
4.1.1	The GRN which encodes the stable Goodwin oscillator	115
4.1.2	Emergence of oscillations in protein production in the stable Goodwin oscillator	116
4.1.3	Theoretical design principles implemented in the stable Goodwin oscillator which enable oscillations	117
4.1.4	The unstable Goodwin oscillator	117
4.1.4.1	Modification 1: Addition of a c-Myc tag peptide and modified Ubiquitin gene upstream of the <i>tetR</i> gene	118
4.1.4.2	Modification 2: Change of UTRs flanking <i>tetR</i>	120
4.2	Construction of the stable Goodwin oscillator	120
4.2.1	Outline of the principle cloning steps	120

4.2.2	Oscillator assembly strategy using non-directional ligation	122
4.2.2.1	Introduction	122
4.2.2.2	Restriction site mapping of the pUBeK backbone	124
4.2.2.3	Preparation of eGFP for insertion downstream of the Ubiquitin degron region	124
4.2.2.4	KpnI non-directional ligation of <i>tetR</i> into pRP _{eGFP} SIR2rp3	126
4.2.2.5	Appraisal of non-directional ligation strategy	127
4.2.3	Oscillator assembly strategy using three-fragment ligation	127
4.2.3.1	Introduction	127
4.2.3.2	Ligation 1: Assembly of pConstitutive plasmid by ligation of Ub-L and eGFP fragments into pRP _{eGFP} SIR2rp3 backbone	129
4.2.3.3	Ligation 2: KpnI non-directional ligation of the <i>tetR</i> cassette into pRP _{eGFP} SIR2rp3	131
4.2.3.4	Appraisal of three-fragment ligation strategy	134
4.2.4	Oscillator assembly strategy inserting the <i>tetR</i> cassette followed by the eGFP gene via Gibson assembly	134
4.2.4.1	Introduction	134
4.2.4.2	Ligation 1: Gibson assembly of ptetR_UbSir by insertion of the <i>tetR</i> cassette into pUbSir	136
4.2.4.3	Ligation 2: Gibson assembly of pStable plasmid by insertion of eGFP into ptetR_UbSir backbone	137
4.2.4.4	Appraisal of <i>tetR</i> -first ligation strategy	139
4.2.5	Oscillator assembly strategy inserting eGFP gene followed by the <i>tetR</i> cassette via Gibson assembly	139
4.2.5.1	Introduction	139
4.2.5.2	Ligation 1: Gibson assembly of pConstitutive by inserting eGFP gene into pUbSir	139
4.2.5.3	Ligation 2: Gibson assembly of pStable oscillator by inserting the <i>tetR</i> cassette into pConstitutive	141
4.2.5.4	Appraisal of eGFP-first ligation strategy	142
4.3	Construction of the unstable Goodwin oscillator	143
4.4	Transfection of <i>T. brucei</i> cells with DNA encoding the stable and unstable Goodwin oscillators	143
4.4.1	Transfection of BSF and PCF <i>T. brucei</i> cells with DNA encoding the stable Goodwin oscillator	144
4.4.2	Transfection of BSF and PCF <i>T. brucei</i> with DNA encoding the unstable Goodwin oscillator	146
4.4.3	Transfection of BSF <i>T. brucei</i> with the pConstitutive control plasmid	147
4.5	Discussion	148
5	Establishing an experimental framework for measurement of oscillating fluorescence levels in viable <i>T. brucei</i> cells	150
5.1	Microfluidic cage devices	153
5.1.1	Design and fabrication of microfluidic traps	153
5.1.2	Testing of microfluidic traps using live <i>T. brucei</i> cells	154
5.2	Agarose blocks for entrapping live <i>T. brucei</i> cells	155
5.3	The use of glass slides, petri dishes and multiwell slides as bases for experimental set-up to immobilise live <i>T. brucei</i> cells	155
5.4	Agarose solutions as live <i>T. brucei</i> cell immobilisation agents	158

5.4.1	Handling and viscosity properties of agarose solutions	158
5.4.2	Effect of agarose solutions on <i>T. brucei</i> cell motility and viability	159
5.5	Agar solutions as live <i>T. brucei</i> cell immobilisation agents	159
5.5.1	Handling and viscosity properties of agar solutions	159
5.5.2	Effect of agar solutions on <i>T. brucei</i> cell motility and viability	161
5.6	Xanthan gum solutions as live <i>T. brucei</i> cell immobilisation agents	162
5.6.1	Handling and viscosity properties of xanthan gum solutions	162
5.6.2	Effect of xanthan gum solutions on <i>T. brucei</i> cell motility and viability	162
5.7	Cygel TM as a live <i>T. brucei</i> cell immobilisation agent	163
5.7.1	Handling and viscosity properties of Cygel TM solutions	163
5.7.2	Effect of Cygel TM solutions on <i>T. brucei</i> cell motility and viability	164
5.7.3	Optimal cell density of live <i>T. brucei</i> cells in Cygel TM	165
5.8	Discussion	168
6	Characterisation of the unstable Goodwin oscillator phenotype in <i>T. brucei</i> cells	170
6.1	Preliminary investigations of the unstable Goodwin oscillator	170
6.1.1	Growth profile of <i>T. brucei</i> cells bearing the unstable Goodwin oscillator . .	170
6.1.1.1	Growth profile of BSF <i>T. brucei</i> cells bearing the unstable Goodwin oscillator	171
6.1.1.1.1	Use of Mann-Whitney statistical analysis for comparison of growth rate data.	171
6.1.1.2	Growth profile of PCF <i>T. brucei</i> cells bearing the unstable Goodwin oscillator	173
6.1.2	Functional activity of the TetR regulation - Expression dynamics of the unstable Goodwin oscillator expression cassette in <i>T. brucei</i> cells following tetracycline addition and removal	176
6.1.3	Half-life of the destabilised TetR and eGFP proteins in <i>T. brucei</i> cells bearing the unstable Goodwin oscillator	178
6.1.3.1	Half-life of the destabilised, Myc-tagged TetR protein in <i>T. brucei</i> cells bearing the unstable Goodwin oscillator	178
6.1.3.2	Half-life of the destabilised eGFP protein in <i>T. brucei</i> cells bearing the unstable Goodwin oscillator	178
6.1.4	Evaluation of the unstable oscillator phenotype following preliminary investigations	179
6.2	Characterisation of the phenotype of the unstable Goodwin oscillator	181
6.2.1	Phenotype of pConstitutive cells in BSF <i>T. brucei</i>	181
6.2.2	Phenotype of the unstable Goodwin oscillator in BSF <i>T. brucei</i>	184
6.2.2.1	Imaging of uninduced BSF <i>T. brucei</i> bearing the unstable Goodwin oscillator	184
6.2.2.2	Imaging of BSF <i>T. brucei</i> bearing the unstable Goodwin oscillator and induced with tc_{thres} to increase the probability of detecting oscillations	184
6.2.3	Phenotype of the unstable Goodwin oscillator in PCF <i>T. brucei</i>	185
6.2.4	Evaluation of the unstable oscillator phenotype following live cell microscopy imaging	185
6.3	Discussion	187
7	Characterisation of the stable Goodwin oscillator phenotype in <i>T. brucei</i> cells via experimentation and mathematical modelling	189

7.1	Preliminary investigation of the stable Goodwin oscillator	189
7.1.1	The growth profile of <i>T. brucei</i> transfected with the Goodwin oscillator	190
7.1.1.1	Growth profile of BSF <i>T. brucei</i> cells bearing the stable Goodwin oscillator	190
7.1.1.1.1	Use of Mann-Whitney statistical analysis for comparison of growth rate data.	190
7.1.1.2	Growth profile of PCF <i>T. brucei</i> cells bearing the stable Goodwin oscillator	193
7.1.2	The minimum tetracycline induction concentration, tc_{thres} , required to induce pStable in <i>T. brucei</i> cells	193
7.1.3	Functional activity of the TetR regulator - Expression dynamics of the stable Goodwin oscillator expression cassette in <i>T. brucei</i> cells following tetracycline addition and removal	198
7.1.4	Half-life of TetR and the destabilised eGFP protein in <i>T. brucei</i> cells having the stable Goodwin oscillator	198
7.1.4.1	Half-life of the TetR protein in <i>T. brucei</i> cells bearing the stable Goodwin oscillator	198
7.1.4.2	Half-life of the destabilised eGFP protein in <i>T. brucei</i> cells bearing the stable Goodwin oscillator	200
7.1.5	Evaluation of the stable oscillator phenotype following preliminary analysis	203
7.2	Preliminary quantitative mathematical modelling and analysis of protein expression dynamics within the trypanosomal Goodwin oscillator	204
7.2.1	Screening of the Goodwin oscillator's parameter space using Gillespie stochastic simulations	204
7.2.1.1	Development of the mathematical model representing protein expression dynamics within the trypanosomal Goodwin oscillator	205
7.2.1.2	Differential equation representation of the mathematical model	206
7.2.1.3	Assumptions made for the analysis of the stable Goodwin oscillator behaviour in <i>T. brucei</i> cells	209
7.2.2	Results and analysis of Gillespie time series simulations to inform screening analysis of the Goodwin oscillator parameter space	210
7.2.2.1	Time series trajectories of the Goodwin oscillator species	210
7.2.2.2	Time series trajectories for different starting conditions	213
7.2.2.3	Time series trajectories of TetR and eGFP proteins	213
7.2.2.4	Parameter analysis based on TetR ₂ dynamics	217
7.2.2.5	Parameter analysis based on eGFP dynamics	220
7.2.3	Evaluation of the Goodwin oscillator parameter space following screening analysis	224
7.3	Characterisation of the phenotype of the stable Goodwin oscillator via live cell imaging and qualitative mathematical pattern analysis	228
7.3.1	Phenotype of the stable Goodwin oscillator in live BSF <i>T. brucei</i>	228
7.3.2	Phenotype of the stable Goodwin oscillator in uninduced live PCF <i>T. brucei</i> cells	229
7.3.3	Phenotype of the stable Goodwin oscillator in live PCF <i>T. brucei</i> cells induced with tc_{thres}	230
7.3.3.1	Live cell imaging of the stable Goodwin oscillator in PCF <i>T. brucei</i> induced with tc_{thres}	230

7.3.3.2	Quantification and analysis of fluorescence oscillatory trajectories at the single cell level in <i>T. brucei</i> cells bearing pStable and induced with tc_{thres}	243
7.3.3.3	Phase space reconstruction of the stable Goodwin oscillator fluorescent profile in the <i>T. brucei</i> host	244
7.3.4	Evaluation of the stable oscillator phenotype following live cell microscopy imaging and qualitative mathematical analysis	247
7.4	Discussion	248
8	Conclusion	250
Appendices		256
A Ligation recipes and gel electrophoresis results from cloning during construction of the stable Goodwin oscillator		257
A.1	Oscillator assembly strategy using non-directional ligation	257
A.2	Oscillator assembly strategy using three-fragment ligation	261
A.3	Oscillator assembly strategy inserting the <i>tetR</i> cassette followed by the eGFP gene via Gibson assembly	265
A.4	Oscillator assembly strategy inserting eGFP gene followed by the <i>tetR</i> cassette via Gibson assembly	268
B	List of primers used during PCR amplification	272
C	List of primers used during plasmid sequencing	274
D	MATLAB code used to apply moving average smoothing technique	276
E	MATLAB code used for phase space reconstruction	278

List of Figures

1.1	Synthetic proof-of-principle oscillatory GRNs which have been constructed and modelled mathematically within synthetic biology using first generation chassis.	37
1.2	Diagrams representing positive and negative regulatory pathways used within proof-of-principle synthetic oscillators developed.	39
1.3	Graph of Hill function for different hill coefficient values	49
1.4	The life cycle of <i>T. brucei</i> , which changes morphology depending on the host (insect vector or mammalian host) and the state (proliferating or not) it is in.	63
2.1	Gibson Assembly Method. Diagram explaining the main steps involved in the preparation and Gibson ligation of two or more DNA fragments.	80
2.2	Agarose block trap device: Main components of the agarose block trap device used to immobilise viable and motile <i>T. brucei</i> cells.	86
3.1	Long-term viability of <i>R. denitrificans</i> and <i>O. indolicum</i> glycerol stocks stored at -80°C, revived in shaken and static cultures over a period of 12 weeks.	99
3.2	Growth of <i>Roseobacter</i> strains in liquid culture in response to varying Chloramphenicol concentrations.	102
3.3	Plot mapping the different electroporation and chemical transformations which were run to insert foreign plasmids into <i>D. shibae</i> and <i>E. coli</i> strains.	113
4.1	Diagram of a theoretical Goodwin oscillator showing the two main components of the GRN; <i>gene A</i> and the associated <i>operator</i>	114
4.2	Diagram of the plasmid encoding the stable Goodwin oscillator which is to be constructed showing the main molecular species in the network.	115
4.3	Diagram explaining how oscillatory protein production dynamics in the Goodwin oscillators occur as a result of the negative feedback loop within the GRN.	117
4.4	Diagram of the unstable Goodwin oscillator showing main molecular species and highlighting the differences between the stable and unstable constructs.	118
4.5	Diagram highlighting the different nucleotides between the original 228bp Ub-L sequence (excluding the start codon) found upstream of eGFP (top row of each line) and the codon optimised Ub-L sequence to be inserted upstream of the <i>tetR</i> ORF (bottom row of each line) in the unstable oscillator.	119
4.6	Diagram representing the silencing of genes by active genes at sub-telomeric loci via inhibitive cross-talk between homologous 3' UTRs.	120
4.7	Summary of the four Assembly strategies developed to construct the plasmid encoding the stable Goodwin oscillator.	121
4.8	Diagram showing the pRP _{eGFP} SIR2rp3 plasmid and the stable oscillator device highlighting the three main construction steps required for assembly.	122
4.9	Oscillator assembly strategy using non-directional ligation.	123

4.10 Analysis of the XbaI and BamHI restriction sites on pUBeK. Diagrams on the left show the pUBeK plasmid cut with XbaI (top), XbaI & BamHI (middle) and BamHI (bottom). The right panel shows digest results of XbaI and BamHI single and double digests.	124
4.11 Diagram showing primers used in the sequencing of pUBeK plasmid to map region surrounding the Ub-L gene and known restriction sites surrounding Ub-L gene prior to and post sequencing.	125
4.12 Assembly strategy using three-fragment ligation.	128
4.13 Assembly strategy using three-fragment ligation - Assessing success of ligation: Analytical double digests using HindIII/BamHI, XbaI/BamHI and HindIII/XbaI on clones picked from ligation plates to assess success of second HindIII/XbaI/BamHI three-fragment ligation to insert Ub-L and eGFP in pRP _{eGFP} SIR2rp3.	130
4.14 Assembly strategy using Gibson assembly to insert the <i>tetR</i> cassette followed by the eGFP gene into pUbSir.	135
4.15 TetR-first assembly strategy: Representation of Gibson primers used for the amplification of the <i>tetR</i> cassette to be inserted into the pUbSir backbone.	136
4.16 TetR-first assembly strategy - Assessing success of ligation: Diagram on the left shows ptetR_UbSir cut with KpnI to assess success of ligation. The right panel shows gel results for digest of DNA eluted from clones picked from ligation plates following the Gibson ligation to insert <i>tetR</i> into pUbSir.	137
4.17 TetR-first assembly strategy - Assessing success of ligation: Diagram on the left shows pStable cut with Apal to assess success of Gibson ligation to insert eGFP into ptetR_UbSir. The right panel shows gel results of the digest.	138
4.18 Assembly strategy to insert eGFP followed by the <i>tetR</i> cassette via Gibson ligation.	140
4.19 eGFP-first assembly strategy, Assessing success of ligation: The left panel shows pConstitutive digested with BsrI. The right panel shows gel results of digests of clones picked following Gibson ligation to insert eGFP into pUbSir to get pConstitutive.	141
4.20 eGFP-first assembly strategy - Assessing success of ligation: Diagram on the left shows pStable cut with KpnI to assess success of Gibson ligation to insert <i>tetR</i> into pConstitutive to get pStable. The right panel shows gel results of digest.	142
4.21 Diagram showing main cloning step required to construct the unstable oscillator plasmid, pUnstable, by inserting the UTR-Myc-Ub-L- <i>tetR</i> -UTR cassette at the KpnI site on the pConstitutive GRN in one step.	143
4.22 Assessing the strength of protein expression of eight different BSF <i>T. brucei</i> clones bearing the pStable plasmid and uninduced (-) or induced (+) with 1 μ g/ml tetracycline via western blot detecting for the eGFP signal and the corresponding coomassie stain.	145
4.23 Assessing the strength of protein expression of four different PCF <i>T. brucei</i> clones bearing the pStable clones uninduced (-) or induced (+) with 1 μ g/ml tetracycline via western blot detecting for the eGFP signal and the corresponding coomassie stain.	145
4.24 Assessing the strength of protein expression of eight different BSF <i>T. brucei</i> clones bearing the pUnstable plasmid and uninduced (-) or induced (+) with 1 μ g/ml tetracycline via western blot detecting for the eGFP signal and corresponding coomassie stain.	146
4.25 Assessing the strength of protein expression of four different PCF <i>T. brucei</i> clones bearing the pUnstable plasmid and uninduced (-) or induced (+) with 1 μ g/ml tetracycline via western blot detecting for the eGFP signal and corresponding coomassie stain.	147

4.26 Assessing the strength of protein expression of eight different BSF <i>T. brucei</i> clones bearing the pConstitutive control plasmid which express protein constitutively. Following protein preparation, analysis is carried out via western blot detecting for the eGFP signal and a corresponding coomassie stain.	147
5.1 Images taken at 30 second intervals showing how free-swimming <i>T. brucei</i> mounted on 5mm multiwell slides move across the plane and in and out of focus, which limits cell tractability when comparing images.	151
5.2 Images showing different BSF morphologies: a) Healthy viable cells, b) Unhealthy viable cells which clump together and c) Non-viable cells which remain isolated but shrink and lose their original morphology.	152
5.3 Microfluidic cage devices: CAD design of two different layouts of microfluidic cage devices used to trap trypanosomes.	153
5.4 Microfluidic cage devices: Images of different microfluidic cage devices constructed having a) a consecutive pattern with square width of 50 μ m, 100 μ m or 150 μ m or b) an alternating pattern with square width of 300 μ m. In both cases, laser mark speed is of 300mm/s.	154
5.5 Microfluidic cage devices: The image on the left shows the device with wells having square width of 50 μ m. The circled wells are amplified in the middle and right images using a magnification of 400X. Visibility is hampered due to the thick border around the perimeter of the well, the bubbles and small dark spots which are present.	155
5.6 Diagram representing the experimental set-up to test the optimal base for the maintenance and imaging of viable <i>T. brucei</i> cells. The set-up consisted of a glass slide, petri dish or 5mm multiwell slide, on top of which cells suspended in media were pipetted, covered with a cover slip and sealed to prevent spills.	156
5.7 Petri dish and glass slide set-ups: Plot showing the percentage of viable healthy cells, viable unhealthy (clumped/misshapen) cells and non-viable cells over time which were encased using a Petri dish or glass slide to mount cells on.	157
5.8 Agarose as an immobilisation agent: Plot showing the percentage of viable healthy cells, viable unhealthy (clumped/misshapen) cells and non-viable cells over time which were overlaid with 0.5% or 1% agarose.	160
5.9 Agar as an immobilisation agent: Plot showing the percentage of viable healthy cells, viable unhealthy (clumped/misshapen) cells and non-viable cells over time which were overlaid with 0.25% or 0.5% agar.	161
5.10 Xanthan gum as an immobilisation agent: Plot showing the percentage of healthy and unhealthy (clumped/misshapen) motile viable and non-viable cells over time which were overlaid with 0.25% or 0.5% xanthan gum.	163
5.11 Cygel TM as an immobilisation agent: Plot showing the percentage of viable healthy cells, viable unhealthy (clumped/misshapen) cells and non-viable cells over time which were overlaid with 40% or 80% Cygel TM	165
5.12 Cell density screening assay: Images of live BSF <i>T. brucei</i> embedded in 40% Cygel TM cell at different concentrations.	167
5.12 (Continued) Cell density screening assay: Images of live BSF <i>T. brucei</i> embedded in 40% Cygel TM cell at different concentrations.	168
6.1 Growth rate profiles and doubling times of induced/uninduced BSF <i>T. brucei</i> bearing pUnstable.	172
6.2 Growth rate profiles and doubling times of induced/uninduced PCF <i>T. brucei</i> bearing pUnstable.	175

6.3	eGFP expression levels over a 72-hour period of BSF <i>T. brucei</i> bearing pUnstable following 1 μ g/ml tetracycline induction and removal. Analysis is carried out via a western blot detecting for the eGFP signal, corresponding coomassie stain and densitometry plot quantifying the amount of eGFP protein detected in the western blot using the positive control as a reference point of intensity 1.0.	177
6.4	Deducing the half-life of the destabilised eGFP protein in BSF <i>T. brucei</i> bearing pUnstable and induced with 1 μ g/ml tetracycline, with analysis carried out via a western blot detecting for the eGFP protein, corresponding coomassie stain and densitometry plot.	180
6.5	Images of pConstitutive-transfected BSF <i>T. brucei</i> over five consecutive time-points taken using a 20X objective. The encircled cell is pConstitutive Cell 1 which was used to quantify variation in the fluorescent profile due to cell undulations.	182
6.6	Images of <i>T. brucei</i> cells bearing the pConstitutive plasmid using an objective of 20X and taken at different time-points.	183
6.7	Diagram explaining how different concentrations of tetracycline molecules affect protein dynamics in <i>T. brucei</i> cells bearing the Goodwin oscillator expression cassette, leading to either undetected fluorescence oscillations, detected fluorescence oscillations or increasing fluorescence levels.	186
7.1	Diagram showing the stable Goodwin oscillator, pStable. The main components within the network are the P_{RRNA} and P_{Ep1} promoters, the <i>tetR</i> cassette, the Ub-L-eGFP fusion gene, the tetracycline operator Tet Op; the 5' and 3' UTRs for both <i>tetR</i> and Ub-L-eGFP; the Hygromycin (Hyg) and Ampicillin (Amp) selectable markers and the rRNA gene targeting fragment having the NotI restriction site.	189
7.2	Growth rate profiles and doubling times of induced/uninduced BSF <i>T. brucei</i> bearing pStable.	191
7.3	Growth rate profiles and doubling times of induced/uninduced PCF <i>T. brucei</i> bearing pStable.	194
7.4	Titration assay carried out to deduce the tetracycline induction threshold, tc_{thres} , of the pStable network. Analysis was carried out via a western blot detecting for the eGFP signal, corresponding coomassie stain and densitometry plot of a BSF clone bearing the pStable network and induced with different levels of tetracycline in the 0-1 μ g/ml range.	196
7.5	Second titration assay carried out to deduce the tetracycline induction threshold, tc_{thres} , for pStable. Analysis is carried out via live cell imaging of a pStable-transfected BSF clone induced with different levels of tetracycline in the 1-2.5ng/ml range. . .	197
7.6	eGFP expression levels over a 72-hour period of BSF <i>T. brucei</i> bearing pStable following 1 μ g/ml tetracycline induction and removal. Analysis is carried out via a western blot detecting for the eGFP signal, corresponding coomassie stain and densitometry plot.	199
7.7	Deducing the half-life of the destabilised eGFP protein in BSF <i>T. brucei</i> bearing pStable and induced with 1 μ g/ml tetracycline, with analysis carried out via a western blot detecting for the eGFP protein, corresponding coomassie stain and densitometry plot.	201
7.8	Deducing the half-life of the non-destabilised eGFP protein in 2T1 BSF <i>T. brucei</i> , with analysis carried out via a western blot detecting for the eGFP protein, corresponding coomassie stain and densitometry plot.	202

7.9 Comparing the half-life of non-destabilised and destabilised eGFP proteins in induced control 2T1 and pStable-transfected BSF clones respectively, with analysis carried out via a western blot and densitometry plot detecting for the eGFP signal.	203
7.10 Diagram showing set of mathematically modelled components and biochemical reactions corresponding to: a) interactions between TetR ₂ components when TetR ₂ is not bound to Op, b) interactions between TetR ₂ components when TetR ₂ is bound to Op and c) interactions between eGFP components when TetR ₂ is not bound to Op.	208
7.11 Time series trajectories of 500 Gillespie simulations over a time-span of 1000 minutes for the species: a) mRNA _{TetR} , b) TetR, c) TetR ₂ , d) mRNA _{eGFP} and e) eGFP. Starting conditions are set to 0 for all species except Op=1.	211
7.12 Time series trajectories of 10 randomly selected Gillespie simulations over a time-span of 1000 minutes for the species: a) mRNA _{TetR} , b) TetR, c) TetR ₂ , d) mRNA _{eGFP} and e) eGFP. Starting conditions are set to 0 for all species except Op=1.	212
7.13 eGFP time series trajectories of 10 randomly selected Gillespie simulations over a time-span of 1000 minutes. The eGFP degradation rate was set to 0.02min ⁻¹ instead of 0.004min ⁻¹ . The axis show Number of molecules vs Time in minutes. The band in black represents the average number of species molecules over time. Starting conditions are set to 0 for all species except Op=1.	213
7.14 Starting molecule numbers set randomly from 1 to 20 for all species except Op=1, TetR₂.Op=0: Time series trajectories for 10 randomly selected Gillespie simulations over a time-span of 1000 minutes for the species: a) mRNA _{TetR} , b) TetR, c) TetR ₂ , d) mRNA _{eGFP} and e) eGFP.	214
7.15 Starting molecule numbers set to 50 for all species except Op=1, TetR₂.Op=0: Time series trajectories for 10 randomly selected Gillespie simulations over a time-span of 1000 minutes for the species: a) mRNA _{TetR} , b) TetR, c) TetR ₂ , d) mRNA _{eGFP} and e) eGFP.	215
7.16 Starting molecule numbers set to 0 for all species except TetR₂.Op=1: Time series trajectories for 10 randomly selected Gillespie simulations over a time-span of 1000 minutes for the species: a) mRNA _{TetR} , b) TetR, c) TetR ₂ , d) mRNA _{eGFP} and e) eGFP.	216
7.17 TetR₂ dynamics: Individual plots of 12 randomly selected TetR ₂ protein quantity trajectories simulated with the Gillespie SSA using different parameter sets. Starting conditions were set to 0 for all species except Op=1.	218
7.18 eGFP dynamics: Individual plots of 12 randomly selected eGFP protein quantity trajectories simulated with the Gillespie SSA using different parameter sets. Starting conditions were set to 0 for all species except Op=1.	219
7.19 Maximum number of molecules: Histogram plots of the maximum number of molecules attained during the 500 simulations for the protein trajectories of: a) TetR ₂ and b) eGFP.	220
7.20 Analysing the parameter space which results in low value (<3 molecules) TetR ₂ oscillation.	221
7.21 Analysing the parameter space which results in medium value (3-4 molecules) TetR ₂ oscillations.	222
7.22 Analysing the parameter space which results in high value (>4 molecules) TetR ₂ oscillations.	223
7.23 Analysing the parameter space which results in low value (<24 molecules) eGFP oscillations.	225

7.24 Analysing the parameter space which results in medium value (24-42 molecules) eGFP oscillations.	226
7.25 Analysing the parameter space which results in high value (>42 molecules) eGFP oscillations.	227
7.26 Images of uninduced pStable-transfected BSF and PCF <i>T. brucei</i> highlighting cell motility over 120 minutes. Images are taken using a 20X objective and brightfield microscopy.	229
7.27 Part 1/6 - Images of PCF <i>T. brucei</i> bearing the stable Goodwin oscillator and induced with tc_{thres} tetracycline, taken at 2 minute intervals and a 20X objective for a two hour period showing fluorescent background corrected (top level) and brightfield (bottom level) microscopy stills. The circled cells were taken forward for quantification analysis.	231
7.27 Part 2/6 - Images of PCF <i>T. brucei</i> bearing the stable Goodwin oscillator and induced with tc_{thres} tetracycline, taken at two minute intervals and a 20X objective for a two hour period showing fluorescent background corrected (top level) and brightfield (bottom level) microscopy stills. The circled cells were taken forward for quantification analysis.	232
7.27 Part 3/6 - Images of PCF <i>T. brucei</i> bearing the stable Goodwin oscillator and induced with tc_{thres} tetracycline, taken at two minute intervals and a 20X objective for a two hour period showing fluorescent background corrected (top level) and brightfield (bottom level) microscopy stills. The circled cells were taken forward for quantification analysis.	233
7.27 Part 4/6 - Images of PCF <i>T. brucei</i> bearing the stable Goodwin oscillator and induced with tc_{thres} tetracycline, taken at two minute intervals and a 20X objective for a two hour period showing fluorescent background corrected (top level) and brightfield (bottom level) microscopy stills. The circled cells were taken forward for quantification analysis.	234
7.27 Part 5/6 - Images of PCF <i>T. brucei</i> bearing the stable Goodwin oscillator and induced with tc_{thres} tetracycline, taken at two minute intervals and a 20X objective for a two hour period showing fluorescent background corrected (top level) and brightfield (bottom level) microscopy stills. The circled cells were taken forward for quantification analysis.	235
7.27 Part 6/6 - Images of PCF <i>T. brucei</i> bearing the stable Goodwin oscillator and induced with tc_{thres} tetracycline, taken at two minute intervals and a 20X objective for a two hour period showing fluorescent background corrected (top level) and brightfield (bottom level) microscopy stills. The circled cells were taken forward for quantification analysis.	236
7.28 Cell A, Part 1/6 - Images showing close-up of Cell A selected from the tc_{thres} - induced PCF <i>T. brucei</i> bearing the stable Goodwin oscillator.	237
7.28 Cell A, Part 2/6 - Images showing close-up of Cell A selected from the tc_{thres} - induced PCF <i>T. brucei</i> bearing the stable Goodwin oscillator.	238
7.28 Cell A, Part 3/6 - Images showing close-up of Cell A selected from the tc_{thres} - induced PCF <i>T. brucei</i> bearing the stable Goodwin oscillator.	239
7.28 Cell A, Part 4/6 - Images showing close-up of Cell A selected from the tc_{thres} - induced PCF <i>T. brucei</i> bearing the stable Goodwin oscillator.	240
7.28 Cell A, Part 5/6 - Images showing close-up of Cell A selected from the tc_{thres} - induced PCF <i>T. brucei</i> bearing the stable Goodwin oscillator.	241
7.28 Cell A, Part 6/6 - Images showing close-up of Cell A selected from the tc_{thres} - induced PCF <i>T. brucei</i> bearing the stable Goodwin oscillator.	242

7.29 Quantified fluorescent trajectories in tc_{thres} induced PCF cells bearing the pStable network: Plot showing raw quantified fluorescence data (in blue) and imputed, noise-filtered and smoothed fluorescence (in orange) time series for three PCF tc_{thres} -induced cells bearing the pStable network, which did not clump or move out of the plane of vision during the two hour experiment and had detectable fluorescence levels.	245
7.30 Smoothed and De-trended data: Diagrams showing smoothed de-trended time series profile (purple) for the fluorescent intensity in Cell A, Cell B and Cell C. Each plot also shows the imputed, noise-filtered, smoothed data which was then detrended (blue), the linear trend which was removed (orange) and the detrended mean which has value zero (black).	246
7.31 Reconstructed phase space of de-trended data: Diagrams showing reconstructed phase space of smoothed de-trended fluorescence signal trajectory of a) Cell A, b) Cell B and c) Cell C. d) shows the reconstructed phase space using all three trajectories. A delay of 8 was used.	247
A.1 Assembly strategy using non-directional ligation - Plasmid preparation: Two-step process to amplify eGFP from pRP _{eGFP} SIR2rp3 and add overhangs using complementary primers on pRP _{eGFP} SIR2rp3 in the first PCR step and mutagenic primers on previously amplified DNA segment in the second PCR step.	258
A.2 Assembly strategy using non-directional ligation - Plasmid preparation: Gel results showing the <i>tetR</i> cassette and pRP _{eGFP} SIR2rp3 during preparation steps for a KpnI non-directional ligation to insert the <i>tetR</i> cassette into pRP _{eGFP} SIR2rp3.	259
A.3 Assembly strategy using three-fragment ligation, Plasmid preparation: Gel results showing eGFP insert, Ub-L insert and pRP _{eGFP} SIR2rp3 backbone during the preparation process for the HindIII/XbaI/BamHI three-fragment ligation of Ub-L and eGFP into pRP _{eGFP} SIR2rp3.	261
A.4 TetR-first assembly strategy - Plasmid preparation: Gel results showing the pUbSir backbone and <i>tetR</i> cassette during preparation steps for a Gibson ligation to insert the <i>tetR</i> cassette into pUbSir.	265
A.5 TetR-first assembly strategy - Plasmid preparation: Gel results showing the insert and backbone during the preparation process for a Gibson ligation to insert eGFP into ptetR_UbSir to get pStable.	267
A.6 eGFP-first assembly strategy - Plasmid preparation: Purified eGFP insert and pUb-Sir backbone in preparation for the Gibson ligation to insert eGFP into pUbSir to get pConstitutive.	269
A.7 eGFP-first assembly strategy, Plasmid preparation: Gel results showing <i>tetR</i> and pConstitutive DNA amplified, digested and purified in preparation for a Gibson ligation to insert <i>tetR</i> into pConstitutive to get pStable.	270

List of Tables

2.1	Set-up and thermocycling conditions used for the PCR-amplification of DNA template using either complementary or mutagenic primers.	77
2.2	Controls used in standard, non-directional and three-fragment ligations. Each control is prepared using a combination of dephosphorylated cut purified backbone OR cut purified backbone OR uncut backbone, Quick T4 DNA Ligase (optional), 2X Buffer Ligase and dH ₂ O to add up to 21 μ l volume.	78
2.3	Controls used in a Gibson Assembly. Each control is prepared using a combination of PCR-amplified/cut, purified backbone OR uncut backbone, 2X Gibson Assembly Master Mix (optional) and dH ₂ O to add up to 20 μ l volume.	81
2.4	Camera settings used in brightfield and fluorescent microscopy to acquire consecutive images of viable cells.	92
3.1	The different Chloramphenicol concentrations which each of the three <i>Roseobacter</i> strains and <i>E. coli</i> were cultured in, in order to assess cell growth inhibition by Chloramphenicol.	100
3.2	BioBrick TM and engineered plasmids used in the chemical and electro transformation of <i>Roseobacter</i> strains.	103
3.3	Electroporation based transformation using different plasmid masses for electroporation. Reactions and controls carried out in electroporation transformation of <i>Roseobacter</i> and <i>E. coli</i> strains with the BioBrick TM pSB3C5.	104
3.4	Transformation using different output voltages in electroporation. Reactions and controls used in electroporation transformation of <i>D. shibae</i> with the BioBrick TM pSB3C5.	106
3.5	Transformation using electrocompetence protocol by Sambrook et al. (2001). Reactions and controls used to transform <i>D. shibae</i> using a modified electroporation protocol and a broader range of plasmid ori sequences	108
3.6	Transformation using the method by Piekarski et al. (2009). Reactions and controls used in the electroporation of <i>D. shibae</i> and <i>E. coli</i> cells with plasmids having a range of ori sequences	110
3.7	Transformation using the heat-shock method. Reactions and controls used to transform <i>D. shibae</i> with the pSB3C5 BioBrick TM using chemical instead of electro transformation.	112
4.1	Assembly strategy using three-fragment ligation - Dephosphorylation: Steps used in the dephosphorylation of pUbSir backbone using different dephosphorylation enzymes.	132
4.2	Assembly strategy using three-fragment ligation: Recipes for ligation reactions and controls used in study to assess efficiency of different dephosphorylation enzymes to prepare the pUbSir backbone for non-directional ligation.	133

5.1	Summary of the different set-ups, solutions and technology used to develop a framework for the measurement of fluorescence levels in viable <i>T. brucei</i> cells.	152
5.2	Agarose as an immobilisation agent: Handling properties of different concentrations of liquid agarose solutions at room temperature at 0, 3 and 7 minutes post preparation.	158
5.3	Agar as an immobilisation agent: Handling properties of different concentrations of liquid agar at room temperature at 0, 3 and 7 minutes post preparation.	160
5.4	Cygel TM as an immobilisation agent: Handling and pipetting properties of different concentrations of Cygel TM solutions when placed on ice, after two minutes of incubation at room temperature and after seven minutes of incubation at room temperature.	164
5.5	The cell density assay: Analysis of the expected and observed cell counts of <i>T. brucei</i> embedded in 40% Cygel TM , depending on the cell density being used.	169
6.1	List of preliminary characterisation assays carried out to investigate function of the main components in the pUnstable oscillator in its <i>T. brucei</i> host.	170
6.2	Comparison of growth rates of BSF <i>T. brucei</i> cells bearing the uninduced/induced unstable Goodwin oscillator expression cassette and unmodified wildtype cells via the Mann-Whitney U test, which tests whether cumulative growth rates for transfected cells are significantly different from wildtype BSF cells.	174
6.3	Comparison of growth rates of PCF <i>T. brucei</i> cells bearing the uninduced/induced unstable Goodwin oscillator expression cassette and unmodified wildtype cells via the Mann-Whitney U test, which tests whether cumulative growth rates for transfected cells are significantly different from wildtype PCF cells.	176
6.4	The pConstitutive and pUnstable-transfected BSF and PCF <i>T. brucei</i> clones which were imaged using live-cell fluorescent and brightfield microscopy to observe fluorescent patterns and characterise the phenotype of pUnstable. 'tc _{thres} ' is the tetracycline concentration above which protein expression across the cell population is induced.	181
6.5	Mean fluorescent signal intensity and percentage of standard deviation of three pConstitutive-transfected BSF <i>T. brucei</i> cells over five time-points. The fluorescent signal has no units.	184
7.1	List of preliminary characterisation assays carried out to investigate function of the main components of the pStable oscillator network in transfected <i>T. brucei</i>	190
7.2	Comparison of growth rates of BSF <i>T. brucei</i> cells bearing the uninduced/induced stable Goodwin oscillator expression cassette and unmodified wildtype cells via the Mann-Whitney U test, which tests whether cumulative growth rates for transfected cells are significantly different from wildtype BSF cells.	192
7.3	Comparison of growth rates of PCF <i>T. brucei</i> cells bearing the uninduced/induced stable Goodwin oscillator expression cassette and unmodified wildtype cells via the Mann-Whitney U test, which tests whether cumulative growth rates for transfected cells are significantly different from wildtype PCF cells.	195
7.4	The species/components from the plasmid encoding for the Goodwin oscillator which were selected for preliminary quantitative modelling and trajectory simulations.	206
7.5	The biochemical reactions selected for quantitative modelling of network dynamics. The species are as described in Table 7.4, while the parameter reaction rates k_i , o_i and d_i are as described in Table 7.6.	207

7.6	The biochemical reaction rate parameters and their distributions, used in the modelling of the TetR and eGFP family of components within the trypanosomal Goodwin oscillator mathematical model.	207
7.7	The pStable-transfected BSF and PCF <i>T. brucei</i> clones which were imaged using live-cell fluorescent and brightfield microscopy to observe fluorescent patterns and characterise the phenotype of pStable. ‘ tc_{thres} ’ is the tetracycline concentration above which protein expression across the cell population is induced.	228
A.1	Assembly strategy using non-directional ligation - Ligation: Set-up of ligations and controls for the KpnI non-directional ligation of the <i>tetR</i> cassette into pRP _{eGFP} SIR2rp3, using Quick T4 DNA ligase.	260
A.2	Assembly strategy using three-fragment ligation - Ligation: Set-up of ligations and controls for the HindIII/XbaI/BamHI three-fragment ligation using increased backbone and insert mass to insert Ub-L and eGFP into pRP _{eGFP} SIR2rp3 to get pConstitutive, using Quick T4 DNA ligase.	262
A.3	Assembly strategy using three-fragment ligation - Ligation: Set-up of ligations and controls for an XbaI/BamHI directional ligation to insert eGFP into pRP _{eGFP} SIR2rp3 as an alternative to the three-fragment ligation.	263
A.4	Assembly strategy using three-fragment ligation - Ligation: Set-up of ligations and controls for the KpnI non-directional ligation of the <i>tetR</i> cassette into pRP _{eGFP} SIR2rp3, using Quick T4 DNA ligase.	264
A.5	TetR-first assembly strategy - Ligation: Recipes for reactions and controls used in Gibson ligation to insert the <i>tetR</i> cassette into pUbSir to get ptetR_UbSir.	266
A.6	TetR-first assembly strategy - Ligation: Recipes for reactions and controls used in Gibson ligation to insert eGFP into ptetR_UbSir to get pStable.	268
A.7	eGFP-first assembly strategy - Ligation: Recipes for reactions and controls used in Gibson ligation to insert eGFP into pUbSir to get pConstitutive.	270
A.8	eGFP-first assembly strategy - Ligation: Recipes for reactions and controls used in Gibson ligation to insert <i>tetR</i> into pConstitutive to get pStable.	271
B.1	List of primers used to PCR-amplify and edit DNA segments throughout the construction process of the Goodwin oscillator.	273
C.1	List of primers used in sequencing reactions following standard and Gibson ligations in order to confirm success of reactions.	275

Abbreviations

eGFP	Enhanced Green Fluorescent Protein
GRN	Genetic Regulatory Network
ODE	Ordinary Differential Equation
SSA	Stochastic Simulation Algorithm
<i>tetR</i>	Tetracycline Repressor gene
TetR	Tetracycline Repressor protein
Ub-L	Ubiquitin protein mutated at the 76th amino acid
Ub-L-eGFP	Fused Ub-L and eGFP proteins
UTR	Untranslated Coding Region

Chapter 1

Introduction

1.1 Synthetic biology and its application

A principal aim of synthetic biology is to build up our understanding of nature (Stephanopoulos, 2012). While this remains an integral part of the vision of synthetic biology, it has now evolved into a much more ambitious enterprise which calls for increasingly complex, non-linear and layered applications from the proof-of-principle networks (Elowitz & Leibler, 2000; Gardner et al., 2000) which were developed at the dawn of synthetic biology.

Projects in synthetic biology are now looking towards industrialisation (Kitney & Freemont, 2012) and the development of real-world applications in areas as diverse as bioprocessing (Peralta-Yahya et al., 2012), therapeutics (Lienert et al., 2014), medicine (Ye et al., 2013; Ruder et al., 2011), energy generation (Vetter et al., 2014; Peralta-Yahya et al., 2012) and biomaterials (Bryksina et al., 2014). Synthetic biology has the potential to optimise bioprocessing and metabolic engineering methodologies in industry through the development of cell factories - 'wet nano-robots' - (Hasty et al., 2002b) which can replace current culturing practices which are resulting in high production costs (Kim et al., 2012; Holtz & Keasling, 2010; Anesiadis et al., 2008).

The wide scope of research within synthetic biology has attracted the attention of public and private funding bodies. For example, the UK government has listed synthetic biology as one of eight main technologies to be invested in (Willets, 2013), resulting in an investment of over £130 million in multiple research centres, grants, training and collaboration schemes (BBSRC, 2014, 2012; Department for Business, Innovation and Skills, 2013). The Engineering and Physical Sciences Research Council (EPSRC) alone has increased its funding almost six-fold since 2011, increasing the value of the number of grants it funds from £12m in 2011 (ESPRC, n.d) to £71m in 2014 (ESPRC, 2014). An analysis on patent applications related to synthetic biology reveals a growing trend, pointing to the growing role industry is playing in this field (van Doren et al., 2013).

As synthetic biology develops and attracts more investment, the challenges it faces become increasingly pressing. The research being carried out is not short on ambition. However, the field is currently based on an *ad hoc* network design process (Tabor, 2012) and lacks a standardised, modular practice which is prevalent in other engineering areas. This is fuelling research towards building a solid foundation and the development of better tools, platforms and techniques to work with (Cameron et al., 2014). For example, projects are being undertaken to establish the

design, construction and use of standardised DNA-plasmids. The most successful of these is the Registry of Standard Biological Parts project (www.partsregistry.org/). This is a database of DNA-plasmids known as BioBricks™ which satisfy specific design requirements (Muller & Arndt, 2012). Other public registries include the ATCC (www.lgcstandards-atcc.org/) and Addgene (www.addgene.org/) databases. However, as of yet, none of these appear to be used by the majority of academia and research bodies (Kahl & Endy, 2013).

Other challenges which are being tackled include the development of alternative host cells (McIsaac et al., 2014), financially competitive products (Church et al., 2014), orthogonal and consistent networks which avoid molecular crosstalk (Arkin, 2013) and quicker development processes. Moreover, developments are being made in order to be able to modify network parts with minimal effort (Arkin, 2013; Kwok, 2010) and for synthetic biology to gain acceptance by the public and favourable and permissive policies from administrative bodies (Hayden, 2014; Erickson et al., 2011), to name but a few. Multiple administrative and funding bodies are now coming together to draw up strategy frameworks (ERASynBio, 2014; Khalil et al., 2012; Board, 2012; Win et al., 2009; Endy, 2007). A search of available literature shows an increase in the past three years on the awareness of, discussion of and funding for dealing with challenges in synthetic biology.

1.1.1 Synthetic biology adopts a multidisciplinary approach: The role of wet lab and mathematical analysis in synthetic biology

Synthetic biology aims to manipulate molecular systems on which available knowledge is often limited. There can be missing information on the parameter rates being used, the stability, optimal culturing conditions, strength, range of function and component compatibility, amongst others. In order to address these gaps, synthetic biology seeks to draw knowledge from multiple fields via collaborations with biologists, chemists, biochemists, mathematicians, computer scientists and engineers (Vinson & Pennisi, 2011) in order to enrich the field with a diverse problem-solving skill-set. Since it is in essence an engineering principle, synthetic biology aims to incorporate theoretical with practical work to complement and inform the other.

Mathematics, in particular, is utilised to develop realistic theoretical models of Genetic Regulatory Networks (GRNs) and to carry out quantitative and qualitative analysis, such as time series simulations, steady state analysis and sensitivity analysis. Multiple hypotheses and different conditions can be tested out mathematically, which would otherwise be infeasible to carry out practically (Chandran et al., 2008). Kaznessis (2007) proposes that modelling should be done on two levels; that of network topologies, where modules in a GRN and their interactions are modelled quantitatively, such as in Stricker et al. (2008), and that of qualitative molecular sensitivity and characterisation analysis, as seen in Francois and Hakim (2004). Endler et al. (2009) suggest that a study which is based on an iterative designing-testing-validation cycle which includes modelling will tend to lead to an economic practice.

The role of mathematics within synthetic biology has been described as “indispensable” (Zheng & Sriram, 2010) and has been shown to bring an element of rational design when dealing with complex, layered cellular systems (Gramelsberger, 2013). A clear example of this can be found in Stricker et al. (2008). Here, the authors develop an oscillator and analyse it mathematically. When carrying out standard time-series simulations, they found that, non-intuitively, oscillations occurred in two different regions of parameter values, rather than the expected single region. The

necessity of mathematical models becomes increasingly evident when looking at studies which show that even when dealing with simple synthetic networks, complex, rich and unexpected dynamics evolve due to non-linearities within the system and results in the whole being greater than the sum of its parts (Borg et al., 2014).

1.1.2 The democratisation of synthetic biology and the DIYbio movement

As public awareness of synthetic biology has increased, a different development has been taking place (Kean, 2011) in parallel to the increased interest from the government, industry and funding bodies. This is the evolution of the do-it-yourself (DIY) synthetic biology movement, which is the participation of citizen scientists in synthetic biology. Citizen scientists have participated in astronomy (Littmann & Suomela, 2014), conservation (Jordan et al., 2011) and ecology projects (Hochachka et al., 2012), and this approach to science has more recently spread to other fields, including biology and genetics (Kawrykow et al., 2012).

Known as 'bio-hackers', 'garage scientists' and 'DIYbiologists', they take advantage of readily-available online software, easy-to-use technology such as the Arduino prototyping platforms, new advanced technology such as 3D-printing (Pearce, 2012) and cheap and second-hand equipment (Ledford, 2010) to build improvised machinery and set up 'community' laboratories at minimised costs outside of an academic or industrial setting. The community laboratories, of which there were 37 registered as of 2013 (Landrain et al., 2013) such as the London Hackspace (www.london.hackspace.org.uk), San Francisco's BioCurious (www.biocurious.org) and New York's Gen-Space (www.genspace.org), are non-profit organisations which offer a space for citizen scientists to pursue their scientific interests.

The DIYbio movement has resulted in a variety of ongoing projects, from genetic disease testing (Alper, 2009) to the designing of water quality monitoring devices (Jorgensen & Grushkin, 2011). Few such initiatives have led to publications or influenced policy-making (Freitag & Pfeffer, 2013). That being said, citizen scientists can bring a different outlook and problem-solving skill set to the field precisely because they are not conventionally trained and are not constrained with deadlines and grants and thus have the freedom and creativity to carry out experiments in a way academics tend not to. Comparisons are often made between synthetic biology and computer science. Many suggest that just like citizen computer scientists revolutionised their field through creativity, the participation of many and an enterprising spirit, so can citizen synthetic biologists contribute to and enrich their respective field (The Economist, 2010).

One step towards the democratisation of synthetic biology which would greatly benefit citizen scientists is the move towards a more open source practice (Nelson, 2014; Pearce, 2012). This remains a controversial topic as the overall benefit of it to synthetic biology is open to questioning (Saukshmya & Chugh, 2010). Moreover, this highlights the ethical issues and biosecurity fears concerning DIYbio and the development and release of mutant strains by individuals, either due to malicious intent (Jorgensen & Grushkin, 2011) or due to a lack of education (Seyfried et al., 2014) which leads to sub-standard handling and disposal of strains.

The importance of having a standardised, modular synthetic biology practice in line with engineer-

ing principles has been mentioned previously. The benefits of this progress will be felt nowhere more strongly than in the DIYbio movement, as this will allow citizen scientists to follow protocols and construct plasmids with ease, in turn allowing for an increase in the quality of DIYbio based science. Projects such as the 3A assembly technique (Shetty et al., 2011) which is a standardised cloning method using BioBrick™ cassettes, OpenPCR (<http://openpcr.org/>) which is an open-source thermocycler which can be assembled in a lab, online funding campaigns (Scudelari, 2013), accessibility to the Registry of Standard Biological Parts (Kean, 2011) and the popular iGEM competition (www.igem.org/) are a step in the right direction towards democratisation of synthetic biology.

1.2 The choice of chassis for synthetic GRNs

Within synthetic biology, networks tend to be developed on a chassis to chassis basis (Danchin, 2012), and networks developed for use in a specific chassis tend to function differently in another host cell (Heidorn et al., 2011). The chassis is defined as the compartment within which the synthetic plasmid is placed and which provides “a metabolic environment, energy sources, transcription, and translation machinery, as well as other minimal cellular function” (Kelwick et al., 2014). Choice of chassis is important since it determines the level of protein complexity which can be synthesised along with downstream processes required. It also influences the rate of output, range of potential applications and stability of the network, as will be seen below.

The use of chassis-specific GRNs has resulted in synthetic biology being at risk of becoming limited by the platforms available for use. Looking at synthetic GRNs being developed, it quickly becomes clear that while the scope and function of synthetic GRNs is not lacking, the choice of host cell to integrate the network in is more restricted.

1.2.1 First generation synthetic biology chassis

Within this thesis, the standard, established chassis within synthetic biology are considered to be bacterial *Escherichia coli* (*E. coli*), mammalian and yeast *Saccharomyces cerevisiae* cells. *E. coli* was the first platform to be widely employed within the field, with mammalian and yeast synthetic networks being developed at a later date. A case-in-point is the development of proof-of-principle synthetic toggle-switches and oscillators: the first networks in *E. coli* cells were developed in 2000, but the first mammalian toggle switch device was constructed in 2004 (Kramer et al., 2004) while the first mammalian oscillator was presented in 2009 (Tigges et al., 2009).

E. coli was and remains a very common platform (Markson & Elowitz, 2014) for the development of both basic components and advanced networks. For example, over 100 regulatory parts having been established for *E. coli* (Nielsen et al., 2013a) and advanced systems such as pathogen biosensing-and-destroying *E. coli* cells (Gupta et al., 2013) have also been constructed. *E. coli* is a common host cell choice due to several reasons, including but not limited to, well-established cloning techniques (Voigt, 2011), extensive troubleshooting know-how (Cameron et al., 2014), rapid doubling time (Markson & Elowitz, 2014) and strain robustness (Liu et al., 2013).

However, as synthetic biology moved from the development of proof-of-principle networks towards real-world applications and more complex functions were demanded of the host cell, the restrictions of *E. coli* became increasingly apparent. Of the more prominent of these limitations is the inability to carry out post-translational modifications of many eukaryotic proteins, the lack of compartments allowing for sophisticated function (Liu et al., 2013) and incompatibility with potential human hosts (Bacchus et al., 2013).

The use of mammalian chassis has gained momentum more recently (Leonard, 2014) for the production of clinically relevant complex proteins (Lienert et al., 2014; Bacchus et al., 2013; Aubel & Fussenegger, 2010). This is because higher eukaryotes, in contrast to prokaryotes, can produce complex proteins, make use of advanced regulatory tuning of network dynamics (Cameron et al., 2014) and can be used to insert multicomponent network systems (Auslander & Fussenegger, 2013).

However, mammalian cells present a considerable set of challenges, especially in a bioprocessing and pharmaceutical industry where product demand and drive for profit is consistently high. Cultures are expensive to maintain (Hirokawa et al., 2013), protein production is a lengthy process and the available parts toolbox is much more limited than that for *E. coli* (Markson & Elowitz, 2014). Moreover, mammalian cells have low growth rates; while *E. coli* double every 20 minutes in optimal conditions, Chinese Hamster Ovary (CHO) cells only double every 14-17 hours given optimal conditions. Research & Developments projects are also time-consuming, with cell line selection processes taking up to one year (Birch, 2005). Throughout the practice of mammalian synthetic biology to date, there remains a lack of standardisation and engineering practices (Auslander & Fussenegger, 2013).

In parallel to the proof-of-principle studies carried out in mammalian cells, yeast (especially *S. cerevisiae*) based networks were also developed. Attempts to construct proof-of-principle systems (Becskei et al., 2001) and cell-cell communication networks (Chen & Weiss, 2005) have been successful, with studies looking into the manipulation of the chassis prior to any mammalian-based studies (Park et al., 2003). Moreover, there is an ongoing project for the development of a synthetic yeast genome, Sc2.0 (<http://syntheticyeast.org/>). Notably, the production of the synthetic artemisinin based anti-malarial medication has been implemented within *S. cerevisiae* (Paddon et al., 2013). However, the use of yeast chassis does not appear to have become as popular as the use of bacterial or mammalian platforms, except for use within genome engineering studies (David & Siewers, 2015) and metabolic engineering studies related to biofuel production (Nielsen et al., 2013b).

Yeast chassis are favoured because culturing and cloning methods, based on the homologous integration of foreign DNA into the genome, are well-established and robust strains are available for use (David & Siewers, 2015). Moreover, yeast is insensitive to phage contamination, has a low growth temperature and high by-product tolerance (Liu et al., 2013). Yeast cultures face limitations because their post-translational modification processes diverge from mammalian ones, making them unsuitable for the production of complex, higher-eukaryote proteins and applications within therapeutics (Kelwick et al., 2014). Moreover, production rates remain low when compared to bacterial counterparts (Liu et al., 2013) and the integration of synthetic pathways is reported to require extensive remodelling of native systems (David & Siewers, 2015).

1.2.1.1 Gene structure, expression and regulatory mechanisms in prokaryote and eukaryote chassis

First generation synthetic biology platforms are used to maintain and express plasmids encoding GRNs, which typically feature many types of genes, associated promoters and other modules such as operators and binding sites. GRNs are self-regulating and can be designed so as to interact with other GRNs and pathways through positive and/or inhibitive feedback. Basic GRNs function as a toggle switch (Gardner et al., 2000), amplifier (Nistala et al., 2010) or oscillator (Elowitz & Leibler, 2000). Their basic structure, expression patterns which allow for a gene to be transcribed to RNA which in turn is translated into protein, and regulatory patterns which control protein quantities vary depending on whether they are placed within a prokaryotic or eukaryotic platform, as will be discussed below.

1.2.1.1.1 Initiation of transcription and transcription. Transcription involves the process of reading and copying a gene sequence of DNA in order to make single-stranded RNA. This process is carried out by RNA polymerase enzymes which catalyse the formation of phosphodiester bonds between RNA nucleotides (Alberts et al., 2010).

Initiation of transcription differs between eukaryotes and prokaryotes. In prokaryotes, an RNA polymerase recognises the promoter via the use of a sigma factor protein to which it is attached. The promoter is a short conserved sequence found at the -35 and -10 positions of the start of the transcription site, taken to be the start of the gene. The sigma factor is released after this function. Given recognition, the polymerase binds to the gene's DNA downstream of the promoter, serving to unwind the double helix and initiate transcription of one of the DNA strands which acts as a template. Transcription occurs one nucleotide at a time. The polymerase serves to bind an incoming ribonucleotide (adenine (A), guanine (G), cytosine (C) or uracil (U), the latter which replaces the thymine (T) nucleotide in DNA). The bound ribonucleotide is complementary to the nucleotide being transcribed in the DNA strand. Thus, the single-stranded RNA chain is elongated one nucleotide at a time. Following transcription of a DNA nucleotide, it is released by the polymerase, allowing the DNA strand to wind back with its complementary DNA strand. This continues until the polymerase encounters a stop signal at which point both the DNA and RNA strands are released and the polymerase can attach to another free sigma factor to start the transcription process all over again. Multiple transcripts can be in development at the same time, at different sections of the DNA template.

In eukaryotes, transcription initiation is similar but includes extra features. For example, eukaryotes have three different types of RNA polymerases, of which RNA polymerase II transcribes genes which lead to protein, while the remainder transcribe genes which lead to tRNA, rRNA and other RNAs which have catalytic functions. Moreover, the RNA polymerase is not enough for transcription to be initiated within eukaryotes. Alongside it, a body of general transcription factors must assemble to form the transcription initiation complex on the promoter. This complex serves to recognise a sequence of T and A nucleotides upstream of the transcription start site known as the TATA box, bind to it and in turn bind the polymerase to the promoter, allowing for initiation of transcription. After attaching the polymerase to the DNA strand, the complex disassembles via a phosphorylation reaction of the polymerase. Following transcription the polymerase goes through a dephosphorylation reaction which enables it to start another round of transcription.

1.2.1.1.2 Post-transcriptional modification. Unlike prokaryotes, genes in eukaryotes tend to be fragmented with sequences of non-coding DNA known as introns placed between the coding parts of a gene which lead to protein, which are known as exons. This means that the two sets of chassis undergo different post-transcriptional processes prior to translation. In prokaryotes, mRNAs transcribed from genes in plasmids are free to begin the translation process even while transcription is ongoing. The ribosome latches on to the free end of the mRNA strand and initiates the translation process simultaneously.

In eukaryotes, transcribed mRNA is not yet mature and undergoes several reactions while transcription is ongoing. RNA processing for mRNA transcripts includes capping, polyadenylation and splicing. Capping is the addition of a guanine nucleotide having a methyl end to the 5' end of the RNA which facilitates transportation of the RNA chain to the cytoplasm. Polyadenylation, on the other hand, involves the processing of the 3' end. Two enzymes serve to cut off the RNA at a specific sequence towards its 3' end and then add a tail which is composed of hundreds of adenine nucleotides. These modifications help to stabilise the mRNA, flag it up for translation and also aid its transfer out of the nucleus. Thirdly, following capping, the introns are removed via splicing which also serves to attach the remaining exons together. This is carried out via a spliceosome which is a complex of small nuclear ribonucleic proteins (snRNPs). Finally, given that the network of interest was inserted within the genome, post-transcriptional processing involves transportation of the mature and processed mRNA from the nucleus via pores in order for ribosomes, which are found in the cytoplasm, to have access to it. Transport of mRNA is only enabled when certain proteins recognise and bind parts of the mRNA, such as the 5' cap, and flag up the structure for transport.

Note that following transcription, RNA can also be modified via folding. This means that RNA is not only an important middle step in the development of proteins to act as a template during translation but that it can also function as an end in itself by folding into a 3D structure to act as a catalyst. For example, ribosomal RNAs (rRNAs) are the backbone of the ribosome while transfer RNAs (tRNAs) are important to the translation process and in the handling of amino acids for attachment to a protein chain as it is being synthesised. Other small RNAs are used for house-keeping duties.

1.2.1.1.3 Initiation of translation, translation and post-translational modifications. Given the above, translation of the mRNA goes ahead to lead to protein. While in prokaryotes, one mRNA can code for several proteins (polycistronic mRNA), in eukaryotes one mRNA codes for one protein (monocistronic mRNA). At this stage, nucleotides on the mature mRNA are read in sets of three to lead to the addition of a single amino acid to the protein chain. This process is carried out via the ribosome. Prokaryotic and eukaryotic ribosomes are similar. They are a complex made of proteins and rRNA. The RNAs form the core while the proteins are found on the surface. The ribosome is made of a small and larger subunit, with each serving a different purpose. The smaller subunit matches tRNAs to the codon triplet being read while the larger subunit develops a covalent bond between amino acids. The ribosome contains three principal sites to which the tRNA binds in succession; the A-, P- and E- sites. Whereas in prokaryotes, circa 20 amino acids are read per second, the rate of translation is lower in eukaryotes, with circa two amino acids

being read per second. tRNAs are important at this stage because they allow the amino acid to bind to the tRNA's 3' end and the corresponding triplet of RNA nucleotides to bind to the tRNAs anti-codon site, as explained further below.

Translation is initiated at a start site, represented by the codon AUG. In prokaryotes, upstream of the start codon is a ribosome binding site (RBS) to which the ribosome can bind directly and initiate translation. In eukaryotes, an initiator tRNA is required. The tRNA binds tightly to the middle P- site on the small ribosomal subunit, along with a set of initiation factor proteins. The mRNA is then bound to the ribosomal subunit from its 5' end which is flagged up by the mRNA's cap and the AUG site located. At this stage, the initiation factors unbind, making way for the larger ribosomal subunit to assemble and form the ribosomal complex.

This allows for the progression of translation which involves reading the mRNA, capturing tRNAs and elongating the polypeptide chain. A tRNA initially binds to the A- site by pairing with a triplet of nucleotides, leaving no empty nucleotides between the previously translated nucleotides and the present ones. Only a tRNA which has the correct amino acid which corresponds to the nucleotide triplet in the A-site can bind. The polypeptide chain then uncouples from the tRNA on the P- site and instead forms a bond with the amino acid on the A- site's tRNA. The large subunit then shifts onwards, moving this tRNA from the A-site to the P- site. As the small subunit also shifts onwards, the process is repeated and at the end of the next round the tRNA would be at the E- site, at which stage it is ejected. Thus, the polypeptide chain is developed from the N- to the C-terminus.

The end of translation is similar in prokaryotes and eukaryotes. The stop codon is not bound by any tRNA but instead binds to a release factor protein which catalyses the addition of a water molecule to the polypeptide chain and hence the release of the chain. At this stage the mRNA is also released, as are the large and small ribosome subunits. Following release, the protein is still not complete and often requires folding and other post-translational modifications. There are over a hundred covalent and non-covalent modifications which can be carried out, the most common being glycosylation and phosphorylation.

1.2.1.1.4 Transcriptional regulatory mechanisms in prokaryote and eukaryote chassis.

Regulation of protein production can occur at any of the transcription, translation and post-translation steps discussed above, although the use of transcriptional regulatory remains the more common method.

In prokaryotes, transcriptional regulation makes use of regulatory DNA sequences known as operators, which are located upstream of the transcription initiation site. The operator can vary in size up to thousands of nucleotides, with eukaryote sites tending to be larger in size. Transcriptional regulator proteins can bind to the operator at multiple locations, resulting in a strong bond. A common DNA binding motif in eukaryotes is zinc fingers. The transcriptional regulator can act as a repressor or as an inducer, either inhibiting or allowing for transcription to take place. A repressor blocks the binding of RNA polymerase. Activators, on the other hand, work by enabling the binding of RNA polymerase which otherwise would not recognise the promoter site. An operator can be bound by both repressor and activator proteins.

In eukaryotes, a regulatory protein can be bound several thousand nucleotides upstream of the promoter. This is possible because the DNA sequence appears to bend to allow easier contact between the bound protein and promoter. Moreover, eukaryotic cells tend to require combinations of proteins to regulate gene expression, known as combinatorial control, as opposed to regulation in prokaryotes wherein one protein can perform the function. Moreover, due to the structure of prokaryotic genes, a single regulatory protein can have an effect on the expression of multiple genes by having these genes within the same operon. This level of co-ordination of expression of multiple genes can also be implemented in eukaryotes but requires the use of combinatorial control which makes use of a common trigger protein.

Positive and negative transcriptional feedback loops can be used for autonomous regulation of a gene being expressed or for exerting an influence on other genes being expressed. Alberts et al. (2010) describe feedback loops as a memory mechanism within cells, because they can propagate specific expression patterns indefinitely.

1.2.1.1.5 Post-transcriptional regulatory mechanisms in prokaryote and eukaryote chassis. Post-transcriptional regulation can also be implemented in order to regulate the amount of protein which is synthesised. There are multiple mechanisms in place for this purpose, starting with processes used during the mRNA maturation stage. In prokaryotes, mRNAs can be regulated autonomously by containing riboswitches. These are sequences of RNA which change structure in the presence of certain metabolites, leading to a stop in the transcription process, resulting in incomplete mRNA. Alternative splicing within eukaryotes, as another example, is a form of post-transcriptional regulation which can lead to a transcript maturing into different products depending on the splicing pattern.

Moreover, the half-life of different mRNAs affects the quantity of protein translated, with mRNAs with short half-lives allowing for only a small amount of proteins to be translated prior to degradation when compared to mRNA with similar length but longer half-lives. The mRNA half-life is in part regulated by the 3' end untranslated region (UTR) lying between the end of the coding sequence and the polyadenylation tail.

Regulation can also occur at the initiation of translation. At this stage, a protein which blocks the RBS can be utilised to prevent translation. In other cases, proteins which bind to RBS can serve the opposite purpose, wherein translation can only be initiated if a protein is bound to the site, allowing the ribosome to locate the AUG site. Moreover, in eukaryotes, the same principle can work by having proteins bind to the 5' cap on mRNA chains and inhibit or activate translation by blocking or aiding the ribosome.

RNAs, for example, micro RNAs (miRNAs) can be used instead of proteins to regulate translation. miRNA can bind to a set of proteins to form the RNA-induced silencing complex (RISC complex) which in turn binds to mRNA. The miRNA in the complex can only bind to complementary mRNA, with the complementary sequence tending to be located in the mRNAs' UTR. This results in the mRNA being degraded by nucleases present in the complex, at which stage the RISC body is released. This results in the cell having no template to use for translation.

RNA interference is a mechanism which draws from the above process and is nowadays engineered to control translation. It is naturally employed to remove viruses. The mechanism works via the use of a nuclease which cleaves double stranded RNA into small interfering RNAs (siRNA) which are then incorporated into a RISC complex, which discards one of the strands. The remaining siRNA strand then acts as a template for other RNA, serving the same purpose as the miRNA described above. Thus, siRNA which recognise any of a number of templates can be used to control gene expression.

Given mature protein, regulation can be carried out via the protein's half-life and stability. Proteases serve to degrade proteins by cutting the peptide bonds between amino acids. Eukaryotes use whole complexes of proteases to degrade proteins, known as proteasomes, which recognise ubiquitin proteins which are covalently linked to other proteins. The process of ubiquitin getting attached to the protein is known as ubiquitylation, wherein enzymes add a ubiquitin train to the protein on recognition of certain motifs. They serve to flag up the proteins for the proteasome to break down.

1.2.1.2 Transcriptional and post-transcriptional regulation mechanisms in GRNs in commonly used synthetic biology chassis

1.2.1.2.1 Transcriptional regulation mechanisms.

Systems which regulate transcription in prokaryotic DNA-plasmids and prokaryotic/eukaryotic operons are based on the principle that regulatory proteins and transcription factors (TFs) bind to corresponding operators and DNA-binding domains (Khalil & Collins, 2010), on the same (Gardner et al., 2000) or neighbouring (Elowitz & Leibler, 2000) operons. This serves to activate or repress transcription by either allowing for or hindering the recruitment or binding of RNA polymerase at the promoter. The effects of the regulatory proteins can often be reversed through signals or chemicals to which the proteins has a greater affinity than the operator (Hillen et al., 1983).

For example, the *tet* operon to which the family of Tetracycline Repressor (TetR) proteins bind is widely employed in synthetic transcriptional control, as are the bacterial λ phage operon, inducible *lac* operon, repressible *trp* operon and inducible/repressible *ara* operon which is equipped with two promoters (P_C and P_{BAD}). Although many of these simple regulation mechanisms are based in prokaryote organisms, they have also been implemented in eukaryotes (Khalil et al., 2012; Wu et al., 1997; Hu & Davidson, 1987).

Apart from the use of single operon systems, combinatorial systems have also been engineered, wherein operators and promoters from different operons are placed together to form a 'mosaic' operon, as developed by Lutz and Bujard (1997). For example, Elowitz and Leibler (2000) use the ' $P_L lacO1$ ' promoter-operator system, which combines the strong P_L promoter from the λ phage with the *lac* operator to which LacI binds and inhibits transcription. This strategy can be seen in multiple other studies, such as Stricker et al. (2008), Basu et al. (2004) and Gardner et al. (2000). Thus, with the manipulation of a limited number of promoters and operators acting either as one unit or as a cascade, a variety of devices, including logic gates, have been developed (Moon et al., 2012; Tamsir et al., 2011).

1.2.1.2.2 Applications of transcriptional regulation. In the Registry of Standard parts, three of the top ten most widely used parts from the database are TetR- and LacI-controlled promoters (Registry of Standard Biological Parts, 2014). Similarly, of the ten most used promoters, seven are inducible and only three are constitutive. Moreover, of the synthetic proof-of-principle oscillators which are discussed later on, the majority use LacI, CI, TetR or Ara-based transcriptional control (see Section 1.2.1.4 for further information). This occurs in spite of a wide variety of options for alternative post-transcriptional regulation mechanisms which are available.

Developments have led to more diverse methods of regulating transcription via add-on features to the relatively simpler operon systems. For example, Ye et al. (2011) use extracellular light signals to power protein production via transcriptional pathways while Weber et al. (2009) use an electronic current as the input signal in a transcriptionally-controlled network. More recently, the construction of synthetic zinc fingers (Khalil et al., 2012; Beerli et al., 2000) and transcription activator-like effectors (TALEs) (Moore et al., 2014) has allowed for more tailored transcriptional control while CRISPR (clustered regularly interspaced short palindromic repeats) DNA has been re-purposed to interfere with gene transcription (Larson et al., 2013).

Transcriptional regulation mechanisms have also enabled cell-cell communication efforts, known as quorum sensing, to be implemented across a population of cells (Danino et al., 2010; Basu et al., 2005; Kobayashi et al., 2004). A popular transcription-based quorum sensing mechanism is the well characterised system extracted from the Gram-negative bacterium *Vibrio fischeri*, based on the *lux* operon, which allows for these marine bacteria to bio-luminesce when population density is high (Gray et al., 1994). This is just one example of a prokaryotic quorum-sensing mechanism, with several other systems based on AHL (N-Acyl homoserine lactones) and non-AHL based quorum sensing systems having been uncovered (Hong et al., 2012; Pappas et al., 2004; Whitehead et al., 2001), although their use in synthetic biology remain limited. Eukaryotic transcription-based cell-cell communication phenomena have also been developed via synthetic biology techniques (Chen & Weiss, 2005; Neddermann et al., 2003; Weber et al., 2003). Intracellular signalling systems have been used in cellular pattern formation (Liu et al., 2011), logic gates (Brenner et al., 2007), population regulation (You et al., 2004) and virus biosensors (Tabor et al., 2009; DeAngelis et al., 2007).

1.2.1.2.3 Post-transcriptional regulation mechanisms. Post-transcriptional regulation mechanisms which can occur at the mRNA maturation, translational or post-translational level vary widely in their scope. A common approach in translational mechanisms is to regulate protein production via mRNA stability and splicing (Auslander & Fussenegger, 2013), the use of riboregulator/riboswitches molecules (Carothers et al., 2011; Bayer & Smolke, 2005; Isaacs et al., 2004) and synthetic RNA designs to prevent the binding of the ribosome (Na et al., 2013). Others use another form of RNA molecules; antisense RNAs, which can be released via the use of trigger molecules to bind to mRNA and inhibit translation (Tigges et al., 2009; Wagner & Simons, 1994).

Post-translational regulation, on the other hand, can be effected through synthetic scaffolds protein (Park et al., 2003) to regulate signalling pathways or proteolysis repression, wherein a protease rapidly degrades an expressed protein, causing a decrease in protein levels (Maldonado et al., 1999). To this, Liernert et al. (2014) add the manipulation of signal pathways, wherein protein receptors are modified so as to respond differently to intra- and extra-cellular signals, as seen in

Dueber et al. (2003).

Research is now at the stage where hybrid regulatory mechanisms with both transcriptional and post-transcriptional control features are being developed, such as the study in Auslander et al. (2012). In fact, the mammalian oscillators discussed further on (see Section 1.2.1.4) use a hybrid approach by employing a combination of anti-sense expression, siRNAs and inducible promoters (Tigges et al., 2010, 2009). However, as Auslander and Fussenegger (2013) discuss, whilst this diverse custom-made regulatory toolkit appears to enrich the field of synthetic biology, it is also representative of the fact that the field still lacks a set of standardised, modular methods to regulate product synthesis post-transcriptionally.

1.2.1.3 Synthetic GRNs in commonly used synthetic biology chassis

Synthetic GRNs are developed via wet lab techniques and mimic and optimise GRNs found in the natural world. Studies on natural networks date back to the 1950s, when it was argued that the body houses 'internal' mechanisms (Menaker, 1996). Although theoretical studies on synthetic networks followed shortly afterwards (Goodwin, 1963), the first synthetic GRNs were designed, constructed and characterised at a much later stage with the development of a bistable toggle switch (Gardner et al., 2000) and oscillator (Elowitz & Leibler, 2000), both of which are based on transcriptional regulatory mechanisms.

Following this, multiple proof-of-principle synthetic GRNs were constructed, some as simple as the single-gene auto-inhibiting, stabilising GRN (based on the TetR and Lac transcriptional regulatory loops) (Becskei & Serrano, 2000), others more complex such as the inverter network (based on the Lac regulatory loop) (Yokobayashi et al., 2002), the mammalian oscillator (Tigges et al., 2009) and the 'calculator' GRN (Friedland et al., 2009). With the maturation of different branches of synthetic biology, applications based on synthetic GRNs, such as biosensor capsules used for artificial insemination in cows (Kemmer et al., 2011), a partially synthetic pathway for producing the anti-malarial drug artemisinin (Paddon et al., 2013) and a light-controlled cascade for controlling glucose levels in mice (Ye et al., 2011), also saw fruition.

1.2.1.4 Oscillatory synthetic GRNs in commonly used synthetic biology chassis

GRNs which function as oscillators are networks wherein protein is produced periodically, based on the interaction of negative feedback loops between the network's components. Several natural systems are known to function as oscillators, although underlying design principles remain only partially understood (Agapakis & Silver, 2009). The most widely studied system is the circadian rhythm, which is the body's natural clock. The circadian rhythm is a dual-feedback GRN with a 24-hour period that is responsible for wake-sleep cycles in mammals, bacteria and plants (McClung, 2006). It is of interest, especially to synthetic biologists, because of its robust (consistent) behaviour in spite of functioning within a noisy environment (Sayut et al., 2007). Other natural biochemical oscillations are found in insulin secretion (Koseska et al., 2011), ATP concentrations during glycolysis, mitosis processes (Novak & Tyson, 2008) and AMP production (McKane et al., 2007) to name but a few. Improper function of these networks is associated with health problems such as diabetes, epileptic seizures, and sleep disorders (Aziz et al., 2010; Danino et al., 2010;

Weldemichael & Grossberg, 2010; Menolascina et al., 2012).

Synthetic biology comes into play because it can be used to gain knowledge and understand clock mechanisms by constructing synthetic networks which function as clocks, as seen in the repressilator design (Elowitz & Leibler, 2000), which is a three-gene network based on the circadian rhythm (which also has three central clock-keeping genes) (Khalil & Collins, 2010). The drive for the forward-engineering of synthetic oscillators as a means of understanding natural systems is echoed by Fung et al. (2005) in the development of the first metabolic oscillator, by Tigges et al. (2010, 2009) in the construction of the first mammalian oscillators and by Danino et al. (2010) in the design of the first quorum-sensing oscillator, amongst others. See Figure 1.1 for a list of the main proof-of-principle synthetic oscillators which have been developed *in-vivo* and Figure 1.2 for an overview of the network designs used in these studies.

Prior to this, the first oscillator which was widely studied, albeit *in silico*, was the Goodwin oscillator (Figure 1.2a) through which the basic design principles required for oscillations to occur in a molecular network were deduced. The Goodwin oscillator is a single gene network which auto-represses gene regulation (Goodwin, 1963). It has often been employed in the study of circadian rhythms, as seen in Gonze et al. (2005) and Ruoff et al. (2001), amongst others. In spite of its simplicity, there were no efforts to develop a Goodwin oscillator *in vivo* prior to Stricker et al. (2008).

Different synthetic oscillators developed to the proof of principle devices presented above include those by Perry and Ninfa (2012) wherein the synthetic oscillator was integrated into the bacterial chromosome. This is different to the approach usually taken in synthetic GRN integration in bacterial hosts, where the network is introduced as a plasmid via chemical or electrical transformation. Chromosomal integration increased plasmid stability and allowed authors to keep track of plasmid copy number, which can not be done accurately if the network exists as a circular DNA-plasmid. Kim and Winfree (2011), on the other hand, constructed a synthetic oscillator *in vitro*, bypassing the use of host cells and making use of just two commercially-available enzymes (bacteriophage T7 RNA polymerase and ribonuclease H). Mondragon-Palomino et al. (2011) developed on the work by Stricker et al. (2008) by focusing on the entrainment of a population of engineered cells into which the oscillator developed by Stricker et al. was inserted. The study is unusual in that it chooses to build on the developments of a previous project. The majority of other studies are original and isolated designs.

There is also a large number of *in silico* studies available on oscillators, focusing on design algorithms (Chang et al., 2013; Chen & Chen, 2010) and specific oscillator features, such as the impact on the network of cell division (Gonze, 2013), growth rate (Osella & Lagomarsino, 2013), nutrients available, intra- and extra-cellular conditions (O'Brien et al., 2012) and quorum sensing

Figure 1.1 (following page): Synthetic proof-of-principle oscillatory GRNs which have been constructed and modelled mathematically within synthetic biology using first generation chassis. 'Org' stands for the organism in which the network was inserted, 'Osc' stands for percentage of oscillating cells, 'Amp' is amplitude of oscillations, 'Dur' is the duration of time over which oscillations were observed, 'CHO' stands for Chinese Hamster Ovary cells, 'S.G' is systematic growth, 'Var' stands for varying, 'f.u' stands for fluorescent units, 'r.a' is relative amplitude, 'a.u.' stands for GFP arbitrary units, 'Ref.' stands for Reference.

	Org	Osc	Period	Amp	Dur	Regulation	Model	Features
1	<i>E. coli</i>	40%	160min ±40	S.G.	c. 10hrs	Negative feedback loop using LacI, λ and TetR systems	Deterministic & Stochastic. Parameters estimated	Known as the 'represillator'; Oscillatory period 3x cell cycle; Cells remain correlated for 95 ± 10 min; Significant cell-to-cell variation; Proteins oscillate in anti-phase. Ref: Elowitz et al. (2000). See Figure 1.2b.
2	<i>E. coli</i>	N.A	11-20hrs	N.A	40-70hrs	Transcriptional, positive and negative (known as amplified negative) feedback loop using Lac and Ntr systems.	Deterministic	Damped oscillations; Relaxation profile for oscillations. Ref: Atkinson et al. (2003). See Figure 1.2c.
3	<i>E. coli</i>	60%	45min ±10	Var.	+4hrs	Transcriptional, amplified negative feedback loop with metabolism controlled by acetyl phosphate.	Deterministic & Stochastic	The first metabolic synthetic oscillator, Known as the 'metabolator'; Tunable response to external stimulus (carbon); Model not robust (sensitive to glycolytic rates, acetate concentration). Ref: Fung et al. (2005). See Figure 1.2e.
4	<i>E. coli</i>	90%	c. 40min	N.A	+4hrs	Transcriptional, amplified negative feedback loop based on Ara and Lac systems, using araC gene to positively regulate network and lacI gene to inhibit transcription.	Deterministic & Stochastic	Use of microfluidic device to control environment; Robust (to arabinose levels); Tunable period (by IPTG, temperature, doubling time). Following parameter analysis via mathematical model authors construct new minimal oscillator based on a singular negative feedback loop - Weak oscillations are observed. Ref: Stricker et al. (2008). See Figure 1.2d.
5	CHO	N.A	170min ±71	1.8±2f.u	N.A	Transcriptional, positive regulation of tTA and PIT transcriptions factors. Post-transcriptional, negative feedback via PIT-driven anti-sense expression of tTA.	Deterministic. Parameters estimated	The first mammalian synthetic oscillator; Tunable (by gene dosage & reporter stability); Robust (to protein and mRNA degradation variations); Cell-cell variability reported (20% variation of GFP expression). Ref: Tsigas et al. (2009). See Figure 1.2c.
6	CHO	12%	25.8hrs ±7.6	34.8 ±22.5 r.a	+100 hrs	Transcriptional, positive regulation via auto-induction of tTA. Post-transcriptional, negative feedback via siRNA mediated silencing of tTA.	Deterministic. Parameters estimated	Oscillations have low frequency; Oscillations repressed by antibiotics; Robust (to different plasmid amounts transfected.). Ref: Tsigas et al. (2010). See Figure 1.2c.
7	<i>E. coli</i>	N.A	55-90min	30±54a.u	N.A	Transcriptional, positive feedback loop using AHL-based quorum sensing components. Post-transcriptional, negative regulation via triggered degradation of AHL enzyme.	Deterministic	The first bacterial synthetic oscillator incorporating quorum sensing; Use of microfluidic device for characterisation; Synchronised oscillations; Period is roughly proportional to enzymatic protein decay time. Ref: Danino et al. (2010). See Figure 1.2d.

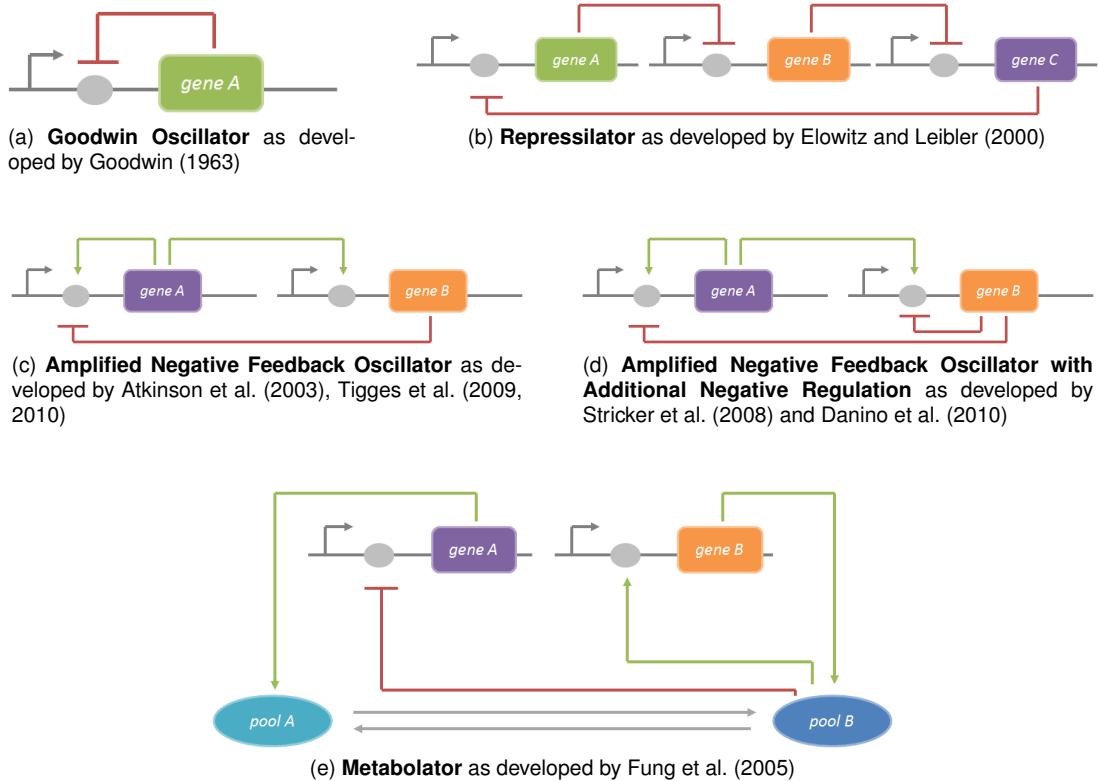


Figure 1.2: Diagrams representing abridged versions of positive and negative regulatory pathways used within proof-of-principle synthetic oscillators developed: a) Goodwin Oscillator as proposed by Goodwin (1963), b) Repressilator as developed by Elowitz and Leibler (2000), c) Amplified Negative Feedback Oscillator as developed by Atkinson et al. (2003), Tigges et al. (2009, 2010), d) Amplified Negative Feedback Oscillator with Additional Negative Regulation as developed by Stricker et al. (2008) and Danino et al. (2010) and e) Metabolator as developed by Fung et al. (2005). Red pathways represent negative regulation while green pathways represent positive regulation. Grey pathways in 1.2e represent metabolic flux.

(Chen & Hsu, 2012; Lang et al., 2011; Garcia-Ojalvo et al., 2004).

Most oscillators are constructed using non-standardised parts. Moreover, characterisation and analytical work is not consistent across the networks. For example, some authors chose to report on oscillatory quantitative dynamics such as the amplitude (Danino et al., 2010; Tigges et al., 2010, 2009) whereas others do not (Stricker et al., 2008; Fung et al., 2005). Even when amplitude readings are given, different units of measurement are used such as 'GFP arbitrary units' and 'fluorescence units', making comparison between them difficult. Having no consistent quantification of the amplitude also means that there is no information available on the strength of the GRN. Progress of synthetic biology as an engineering field will require that genetic networks have the same set of characterisation data available, similar to the data information sheets proposed by Canton et al. (2008), which show that abstraction and modularity is important not only in the construction stages but also in the data communication process.

Throughout, noise plays an important role within synthetic oscillators. In spite of mimicking the circadian rhythm which is extremely robust, the synthetic copy by Elowitz and Leibler (2000) is prone to stochasticity and a varying amplitude (Drubin et al., 2007) and other oscillators similarly exhibit damped or varying dynamics which are attributed to noise (Fung et al., 2005; Atkinson et al., 2003). As it became clear that the role of noise within natural and synthetic oscillators could not be left out of the equation, there was an increase in the number of studies which analyse and debate the role of noise within networks (such as Raser and O'Shea (2005, 2004) and Elowitz et al. (2002) to mention but a few).

1.2.1.4.1 Characteristics of proof-of-principle synthetic oscillator GRNs. The majority of synthetic oscillators constructed have been developed for *E. coli* chassis. Moreover, proof-of-principle networks tend to be relatively simple constructs, both in terms of the size and number of genes used and the type of feedback implemented. In spite of this architectural simplicity, resulting dynamics can be very diverse. As a case in point, proof-of-principle oscillators have been constructed with periodic cycles lasting from 13 minutes (Stricker et al., 2008) to 26 hours (Tigges et al., 2010). In addition, the oscillations have different profiles, from non-sinusoidal relaxation oscillations characterised by a steep amplitude rise and gradual decrease (Tigges et al., 2010; Atkinson et al., 2003) to standard sinusoidal oscillations (Danino et al., 2010; Stricker et al., 2008).

1.2.1.4.2 The scope of mathematical models in the development of synthetic oscillator GRNs. For most synthetic oscillator constructs, deterministic or stochastic mathematical models were developed in tandem. A deterministic model ignores random events and is simpler, since it avoids use of probability measures. However, this can result in shortcomings since variability can have significant effects on the system. In fact, studies where both a deterministic and stochastic model were developed generally resulted in different observations being made. In both bacterial (Purcell et al., 2010) and mammalian (Tigges et al., 2010) dual feedback networks, oscillations in the stochastic model occurred over a larger set of conditions and replicated *in-vivo* dynamics more faithfully than deterministic counterparts. As with choice of characterisation experiments carried out, depth and scope of mathematical analysis varies between studies. Danino et al. (2010), Tigges et al. (2009), Fung et al. (2005) and Elowitz and Leibler (2000) use their model to carry out parameter space analysis while Atkinson et al. (2003) use modelling to scan multiple GRN

designs and select those networks which function as oscillators. It appears that only the minority of studies (Tigges et al., 2010; Stricker et al., 2008) employ mathematical models in an iterative cycle of designing and testing GRNs. Stricker et al. (2008) are one such case, using mathematical analysis to locate a second region in the parameter space where oscillations occur to extract a minimal oscillator from their original device, based on a singular dominant negative feedback loop.

In summary, proof-of-principle synthetic oscillator GRNs which have been developed for use within first generation synthetic biology host cells have been discussed with respect to the variety of dynamics displayed, mathematical models developed, the stochastic environment and standardisation efforts, amongst other. The data shows that proof-of-principle networks are simple in structure and predominantly use transcriptional regulation which has nonetheless led to a rich variety of oscillatory dynamics.

1.2.1.5 Design principles for oscillatory synthetic GRNs

The following section adopts a more theoretical viewpoint and discusses design principles required to construct an oscillatory synthetic GRN which is also autonomous. Networks that can only be made to function as an oscillator via an external periodic input are classed as non-autonomous. Even though there is little in terms of standardisation and the synthetic GRNs presented so far vary in their use of parts, choice of parameters and environment conditions, certain basic theoretical considerations are common to all oscillatory synthetic and natural GRNs. This section will show that fundamentally, oscillations occur due to a series of activation and inactivation biochemical regulatory processes (Novak & Tyson, 2008).

Of the features discussed, the combination of negative feedback, time-delay and non-linearity is necessary and sufficient to cause instabilities which result in oscillations, or in mathematical terms, a Hopf bifurcation (Erneux, 2009). Positive feedback is an optional design feature which can make oscillations more robust (Stricker et al., 2008). It remains unclear as to whether noise features within the systems have a constructive effect on oscillations. These design considerations can not be considered in isolation without reference to each other.

Although theoretical, this section will not adopt a strong mathematical undertone. Mathematically, the design principles backing the development of oscillations have long been elucidated (Tyson & Othmer, 1978; Griffith, 1968). It was only more recently that these started being incorporated in *in vivo* biological studies. For a detailed review of mathematical work on oscillators, see Purcell et al. (2010) and Novak and Tyson (2008).

1.2.1.5.1 The role of negative feedback. Synthetic oscillators have a dominant negative feedback loop which acts at the transcriptional or translational level. Mathematically, it is defined as follows:

If there is a path from the i^{th} node, p_i , of an interaction graph to itself, $p(i, i) = (i = p_1) \xrightarrow{e_{p_2, p_1}} p_2 \dots p_{l-1} \xrightarrow{e_{p_l, p_{l-1}}} p_l = i$ for $l > 2$, where e_{p_k, p_j} is the interaction between the nodes p_k and

p_j , then this path is said to be a feedback loop. In addition, this feedback loop is said to be negative if $\prod_{m=1}^{l-1} s_{m+1,m} = -1$ where $s_{k,j} = 1$ if the interaction between the j th and k th component is positive and $s_{k,j} = -1$ if the interaction between the j th and k th component is negative (Wang et al., 2005)

It has been proven mathematically that a system without negative feedback can not perform oscillations and moreover, the negative feedback must be dominant for a system to perform stable oscillations, as shown in Muller et al. (2006) who looked at n -gene repressilators. These are a generalisation of the three-gene repressor by Elowitz and Leibler (2000), which can be defined as a ring of genes where gene l inhibits the expression of gene $l + 1$ and the last gene n inhibits gene 1.

In addition to the strength of the negative feedback within the GRN, the type of negative feedback used (whether it is transcriptional or post-transcriptional) influences network dynamics, as seen in Guantes and Poyatos (2006), where two dual-feedback oscillators were developed, with sigmoidal translational repression and linear post-translational repression respectively. Oscillations occurred in both models, but were triggered by different biochemical processes. In spite of similar parameter choices, the network with translational repression had larger amplitudes, more tunable frequencies and increased sensitivity to external stimuli than the oscillator with post-transcriptional repression.

1.2.1.5.2 The role of time-delay. *Delayed* negative feedback (Tigges et al., 2010, 2009; Stricker et al., 2008) prevents a synthetic oscillator from attaining homeostasis (mathematically referred to as a stable steady state), which occurs if plain negative feedback is present in the system. A time-delay within a negative feedback pathway can be defined as the time interval τ between when a reaction r in the pathway is initiated and terminated. For example, a time-delay within a synthesis pathway can be defined as the time interval τ between when a gene a starts being transcribed and when the product protein A starts exerting an influence on genes within the GRN. This delay occurs due to the time taken for transcription, translation, post-translation and translocation processes. Thus, time-delay is implemented naturally and automatically in *in vivo* networks, especially in eukaryote platforms. In fact, Purcell et al. (2010) attribute the robustness of oscillations in the mammalian devices by Tigges et al. (2009) and Tigges et al. (2010) to natural time-delay caused by synthesis in eukaryotic platforms.

It has been proved mathematically (through use of the Bendixson's Negative Criterion) that no two-gene network can produce sustained oscillations without a sufficient element of time-delay or equivalent (Xiao & Cao, 2008). A time-delay prevents protein levels from converging to a stable steady state and instead introduces an element of instability, acting as a bifurcation parameter and forcing the system to undershoot and overshoot the homeostatic state. This forms a loop around this state known as a limit cycle, which represents oscillatory dynamics (Wang et al., 2005).

If insufficient time-delay is present in the GRN architecture, additional time-delay can be generated by introducing supplementary genes within the feedback path so as to generate a negative feedback cascade (Novak & Tyson, 2008). This has the additional advantage of making the network more insensitive to noisy fluctuations, since effects of noise are averaged out along the cascade (Andrianantoandro et al., 2006). An alternative way of generating sufficient time-delay is

to modify protein synthesis by, for example, slowing protein folding (Orosz et al., 2010).

Time-delay can have additional benefits such as increasing robustness of oscillations (Orosz et al., 2010; Fung et al., 2005). It can be used to tune period and amplitude (Wang et al., 2005; Chen & Aihara, 2002) and prevent damped oscillations (Tigges et al., 2010). However, time-delay can work against network stability if it is too long. It can cause oscillations to become irregular or even to stop altogether (Fung et al., 2005). This can be explained by the fact that the time-delay is inherently related to the value of biochemical parameter rates (for example, the longer the time-delay in transcription, the lower the transcription rate). Thus, suboptimal time-delay values can result in corresponding parameter values which fall outside the parameter range which represents the range of possible values of the k parameters in a GRN region which allows for oscillations and Hopf bifurcations to occur. In fact, it has been shown that different parameter combinations within the same network architecture can cause damped oscillations, undamped oscillations and bistability (Atkinson et al., 2003).

There is an equivalent way of considering time-delay; via ratios of activator modules' to repressor modules' biochemical reaction rates. The activator module refers to the set of genes which source products which activate synthesis. Similarly, the repressor module is the set of genes which when expressed, produce proteins which inhibit expression. Research has found that for oscillations to occur, the timescales of *activator* synthesis and degradation rates should be higher than the respective *repressor* rates (Guantes & Poyatos, 2006; Hasty et al., 2002a). This allows for an accumulation of activator molecules and proteins prior to inhibitive proteins becoming dominant, thus introducing instability and non-linearity. This accumulation produces the same effect as a time-delay. Zhou et al. (2008) suggest that the main difference between having a time-delay and specific activator-to-repressor ratios is that whereas a time-delay leads to sinusoidal oscillations, fast-activator-slow-repressor dynamics lead to relaxation oscillations. In order to ensure correct ratios, several modifications can be carried out. One of the more common is the fusion of proteins to small stable RNA A (ssrA) tags to shorten decay rates, as seen in Goodwin (1963) and Elowitz and Leibler (2000).

1.2.1.5.3 The role of non-linear dynamics. A third design requirement for an oscillator device is to have a source of non-linearity within the network (Purcell et al., 2010; Chen & Aihara, 2002) which prevents the system from achieving equilibrium, in line with the role of time-delay. Non-linearity is best described and assessed mathematically (Teschl, 2012):

If protein level variations of a system of gene products $\mathbf{x} = (x_1, x_2, \dots, x_n)$ are represented by the set of first-order, differential equations $\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(t - \tau, \mathbf{x})$, the system is said to be linear if for $i = 1, \dots, n$ it can be represented as $\frac{dx_i(t)}{dt} = g_i(t - \tau) + \sum_{j=1}^n a_{i,j}(t - \tau) \cdot x_j$, and is said to be non-linear otherwise.

It is manifested via feedback and the level of co-operativity within the GRN (Kaern et al., 2003; Goldbeter, 2002). Several synthesis and metabolic processes, such as substrate and product inhibition and diffusion are non-linear reactions wherein, in abstract terms, the output is not proportional to the amount of causative agent which is present. This means that products of reactions

do not vary at a constant rate. This is often represented by the Michaelis-Menten enzyme theory (Wilkinson, 2011), with corresponding reactions being represented as a biochemical rate by using the Hill equation (Bernot et al., 2013).

Non-linear reactions are also found within bistable systems (Goldbeter, 2002). This highlights how the different design requirements must be combined together to lead to the required oscillatory output. In fact, Becskei and Serrano (2000) developed a 1-gene repressilator which led to homeostasis in spite of dominant negative feedback. Stricker et al. (2008) developed a similar design, but based the negative loop on the *lac* operon, repressed by the LacI *tetramer* rather than the *tetR* operon, repressed by the TetR *dimer*. Their network exhibited oscillations. Purcell et al. (2010) attribute the onset of oscillations in the Stricker et al. design to a stronger element of non-linearity within the system due to the extra polymerisation reactions required. Non-linearity, although implicitly present in the choice of genes, time-delay and feedback within the network, is rarely discussed by authors.

1.2.1.5.4 The role of positive feedback. A positive feedback loop activates gene expression and results in increased protein levels. Similarly to negative feedback, it can act at the transcriptional or post-transcriptional level. Although direct positive feedback is not a necessary component for oscillations to occur (see Stricker et al. (2008) and Elowitz and Leibler (2000)), many choose to incorporate it into the GRN design (Danino et al., 2010; Tigges et al., 2010; Atkinson et al., 2003). Positive feedback loops and pathways are found in several natural systems, including circadian clocks (Guanes & Poyatos, 2006), cell cycles and pacemaker neurons (Pomerening et al., 2005). This suggests that positive feedback has distinct benefits on network dynamics and system stability. In fact, in *E. coli* (Stricker et al., 2008) and *Xenopus laevis* (Tsai et al., 2008) synthetic oscillators, positive feedback allowed for a more tunable frequency, increased robustness and more regular oscillations.

1.2.1.5.5 The role of noise. GRNs function within a noisy environment caused by extrinsic sources (neighbouring cells, temperature gradients, metabolites, ribosomes, cell-cycles) and intrinsic sources (biochemical synthesis and metabolic processes, protein leakage) (Strelkowa & Barahona, 2010). The effects of noise is felt strongly when gene copy numbers are low (Tigges et al., 2009; Yoda et al., 2007). Larger numbers of gene and mRNA copies, on the other hand, even out the effect of noise. Some suggest that extrinsic noise sources have a larger impact on gene networks since they cause fluctuations which are not synchronised with the network (Andrianantoandro et al., 2006).

The role of noise is not well understood, particularly as to whether it has a constructive or destructive role on network dynamics. Bates et al. (2014) show that different levels of noise can have a varied effect on GRN dynamics, with optimal noise levels leading to stochastic resonance in the intracellular genetic perceptrons being analysed. In a study showcasing how noise can have a constructive role, Bratsun et al. (2005) carried out both deterministic and stochastic analysis on a time-delayed negative feedback network and showed that while no limit cycles were observed in the deterministic case, the instabilities caused by noise resulted in oscillations. In multiple other GRNs, noise has led to stochastic coherence (Borg et al., 2014; Purcell et al., 2010). Guantes and Poyatos (2006) claim that stochasticity increases survival chances of cells within a constantly

changing environment. This is exhibited in the analytical work of Stricker et al. (2008), wherein a noise component retains oscillations mostly stable by preventing damping of oscillations.

Noise can also have a detrimental effect and result in a varying amplitude and loss of uniform distribution of genetic material during cell division, as seen in Fung et al. (2005) and Elowitz and Leibler (2000). Robustness in the presence of noise remains a major challenge within synthetic biology (Tigges et al., 2010), partly because it is unintuitive (Guantes & Poyatos, 2006). As Elowitz and Leibler (2000) postulate, progress needs to be made to at least distinguish between noise and unknown intrinsic causes, including chaos. For example, only a minority of cells transformed with Elowitz's repressilator showed oscillations, which they only account in part to stochasticity (Purcell et al., 2010). Similarly, Fung et al. (2005) are hesitant in their conclusions regarding noise, saying that noise *may* account for dynamics observed. Although there are multiple papers looking at sources and effects of noise, fewer discuss options on how to control them. Danino et al. (2010) offer one solution, arguing that quorum sensing has the potential to lessen and overcome effects of internal stochasticity. To this, Purcell et al. (2010) and Barkai and Leibler (2000) add that the effects of noise can be minimised by selecting specific parameter rate combinations.

In summary, dynamics and dynamic features in networks, such as oscillations, robustness, sensitivity, amplitude, period and tunability depend on the optimal combination of the design requirements discussed above (delayed negative feedback within a non-linear system) and the levels of noise which are present.

1.2.1.6 Construction and characterisation of oscillatory synthetic GRNs

1.2.1.6.1 Construction of oscillator plasmids. The first step in the development of GRNs prior to insertion into a first generation chassis, is to design a DNA plasmid consisting of promoter sites, operators, genes, selective markers plus additional features such as restriction sites. The layout of the plasmid allows for the implementation of the theoretical considerations discussed in the previous section. For example, Stricker et al. (2008) use a higher copy number of the activator cassette than the repressor cassette, to enable an increase in protein levels before inhibition becomes effective, in line with time-scale and non-linearity requirements.

Preliminary work often requires other modifications to strains, genes and backbones which are to be used. These include altering synthesis and degradation rates such as fusing of the short 11 amino acid ssRA tags which flag up the protein for degradation. This was seen in Danino et al. (2010), Stricker et al. (2008), Fung et al. (2005) and Elowitz and Leibler (2000), who used the same machinery to destabilise both functional and reporter genes. Other common modifications include replacing the start codon GTG with ATG in order to increase transcription efficiency (as seen in Stricker et al. (2008), Fung et al. (2005), Elowitz and Leibler (2000)).

1.2.1.6.2 Insertion of oscillator networks into cells. Transformation processes are not discussed in detail by any of the authors and therefore it is not possible to deduce the reasons for their choice of transformation. It appears that standard transformation (Elowitz & Leibler, 2000), co-transformation of multiple plasmids (Stricker et al., 2008), electroporation (Atkinson et al., 2003)

or transfection (Tigges et al., 2009) procedures are employed. Fung et al. (2005), on the other hand use the CRIM (conditional replication integration and modular) method by Haldimann and Wanner (2001) to integrate their plasmid into the bacterial chromosome using site-specific recombination. For bacterial oscillators, the use of plasmid ensures a high copy number. This is useful because high copy numbers of a GRN may reduce the effect of noise (Yoda et al., 2007; Tigges et al., 2009). For oscillatory GRNs within mammalian and yeast cells, chromosomal integration is the most straightforward route for gene insertion, which helps to lead to a more stable construct which is more feasible to maintain than cells transformed with plasmid DNA which require a constant supply of selective agents (Gu et al., 2015). Moreover, chromosomal integration allows for tighter control over the copy number.

Following plasmid insertion with any of the methods mentioned above, cells which were successfully transformed were selected using antibiotics, with Kanamycin and Ampicillin being the most widely used, followed by Chloramphenicol and Gentamycin. Culturing of cells is not discussed in depth, either, with authors mentioning use of media, culturing conditions and equipment only briefly. Culturing conditions are important because they can be used as a tuning measure. For example, a different source of carbon can result in varying oscillation characteristics, as seen in Fung et al. (2005). Only Atkinson et al. (2003) make specific reference to the turbidity, stirring and aeration conditions of the cell culture in the chemostat used.

1.2.1.6.3 Characterisation, microscopy and cell imaging. By looking at oscillator characterisation efforts, it is evident that emphasis is placed on live single cell phase contrast and fluorescent microscopy and subsequent analysis of images acquired during this step. Authors do not provide detail on characterisation assays, choosing instead to discuss microscopy and flow cytometry efforts and whether oscillations are observed during this stage. This is relatively easy to do via fluorescent microscopy due to the use of fluorescent reporters as the GRN output (Shaner et al., 2005; Kain et al., 1995) which make for simple visual detection of oscillations in the protein quantities. The reporter gene is set up within the plasmid or system, such that while it does not influence regulatory genes, they affect its expression via the use of feedback loops and environment-responsive promoters. Thus, the reporter protein is able to reflect activation or inhibition events occurring within the plasmid. In the case of the oscillators, variants of the green fluorescent protein (GFP) are the most widely employed reporter protein, with only Tigges et al. (2010) (from the studies reviewed above) using a variant of the yellow fluorescent protein (YFP).

In order to detect oscillations in protein levels, authors use a combination of time-lapse fluorescent (Tigges et al., 2010), bright-field (Fung et al., 2005) and phase contrast (Stricker et al., 2008) microscopy to observe cells. Preparation, mounting devices, exposure times and frequency vary from study to study, however, it is interesting to note that images were taken every 2-5 minutes in all cases with the exception of Tigges et al. (2010) where images of CHO cells bearing the synthetic oscillator were taken every 20 minutes. Apart from single-cell microscopy, Stricker et al. (2008) also carry out flow cytometry to extract population-level information on the oscillations.

The number of cells/colonies observed and the duration of observations varies. For example, Elowitz and Leibler (2000) observed more than 100 cell lineages in each of three micro-colonies transformed with the repressilator. Tigges et al. (2010), on the other hand, transfected 35,000 cells but carried out studies on less than 600 cells in each case. Authors tend to not provide

the rationale behind their decisions. Efficiency of cell transformations, percentage of cell viability and the type of cells analysed (for example, all cells or viable cells only) is frequently not stated.

Moreover, few discuss replication and validation issues although Fung et al. (2005) mention that triplicate measurements were carried out, while Stricker et al. (2008) develop a device to allow for parallel analysis of three isolated colonies. Fung et al. (2005) and Elowitz and Leibler (2000) report using a reporter-only plasmid, using different promoters and adding inducer molecules, amongst others, to analyse control plasmid dynamics.

1.2.1.6.4 Microscopy imaging set-up. Authors emphasised the importance of retaining a closed system with stable conditions and insulation throughout the duration of live cell imaging. Danino et al. (2010) placed the microscope within a plexiglass incubation chamber, Elowitz and Leibler (2000) used a Peltier device, which is a device wherein heat is transferred between two different materials, while Fung et al. (2005) use an objective heater.

Interestingly, whereas most authors mount cells on glass slides in some form or another, Danino et al. (2010) and Stricker et al. (2008) make use of a microfluidic devices to culture and observe cells. In the Stricker et al. (2008) design, the device has a well with a height of $1\mu\text{m}$ which allows a monolayer of *E. coli* to grow while being constantly supplied with nutrients. By developing three such wells, three populations could be analysed simultaneously. The authors also incorporated heater channels in the device through which to pass heated water to maintain temperatures and a switch to allow for induction experiments to be carried out. The design, based on a yeast chemo-stat device, takes into account how to remove cells on the edge of the growing population which are spilling out of the well. The device was built using standard soft lithography techniques and polydimethylsiloxane (PDMS) material.

The design of Danino et al. (2010) was based on that of Stricker et al. (2008), with modification such as a higher well height ($1.65\mu\text{m}$ instead of $1\mu\text{m}$) to allow for better nutrient supply. In their review, Gulati et al. (2009) describe how microfluidics has the potential to be a foundational technology in synthetic biology, and proof of this can be seen in the above smart designs which allow for the culturing of populations of cells and single-cell imaging within the same set-up.

1.2.1.6.5 Post-imaging analysis. Following the imaging process, analysis had to be carried out to quantify the fluorescence observed. Again, one can see a variation here in the methods described and the software programs used by the authors. Stricker et al. (2008), Fung et al. (2005), Atkinson et al. (2003) and Elowitz and Leibler (2000) chose to analyse cells manually by picking a number of cells on the last image and back-tracking through the previous images to build the cells' fluorescent trajectories. Elowitz and Leibler (2000) then analyse their data using Fast Fourier Transforms. Tigges et al. (2010) develop their own custom-built software within MATLAB to segment, track and select oscillating cells.

1.2.1.7 Mathematical modelling of oscillatory synthetic GRNs

Mathematical modelling of GRNs is a powerful tool and encompasses a wide variety of theories and techniques with which countless studies have been carried out. This section will focus on those techniques which are made use of for the modelling of synthetic genetic oscillators. These methods serve three purposes; carrying out quantitative analysis, qualitative analysis and parameter estimation. Quantitative analysis leads to the development of a deterministic or stochastic mathematical model from which time series profiles of protein expression dynamics within the GRN can be simulated. Qualitative analysis, on the other hand, can be used to analyse a network's innate properties such as stability and robustness. Parameter estimation is used in the deduction of biochemical rates, which are often unknown and which often are the foundation for the models.

In the study of network dynamics, a mathematical model serves the purpose of describing protein turn-over due to feedback signals under different environmental conditions (Chandran et al., 2008; de Jong, 2002; Smolen et al., 2000). This reduces the need for time- and cost-expensive experiments and provides a “rigorous, systematic, and quantitative linkage between molecular and microscopic phenomena on one hand and macroscopic process performance on the other” (Bailey, 1998). A model can be either deterministic or stochastic. The former ignores all sources of noise within and surrounding the GRN under consideration. The latter, although more biologically relevant, is more tasking to solve and analyse. Non-linear, first-order ordinary differential equations (ODEs) are often used to represent synthetic gene oscillators, although alternative representations can be used. These include discrete representation techniques such as directed graphs (Yongling & Su-Shing, 2006), Bayesian and boolean networks (Saadatpour & Albert, 2013) which make use of logic networks to represent network topology and dynamics, and continuous representation techniques such as partial and stochastic differential equations (Das, 2009).

1.2.1.7.1 Deterministic modelling of network dynamics.

1.2.1.7.1.1 Ordinary differential equations. Deterministic, non-linear, first-order, continuous ordinary differential equations (ODEs) are often used to represent protein and mRNA kinetics in synthetic gene networks as seen in Danino et al. (2010), Tigges et al. (2010, 2009), Stricker et al. (2008) and Elowitz and Leibler (2000). A differential equation shows the interplay between production rates and consumption rates of a product (protein or mRNA) x_j , based on feedback and biochemical processes within the network such as transcription, translation and degradation (Smolen et al., 2000). ODEs which represent gene oscillatory networks often incorporate a time-delay τ to take into account time taken up by ‘long’ biochemical processes such as transcription and translation. In general, a first order non-linear ODE with time-delay τ for a gene product x_j is represented by:

$$\frac{d x_j}{d t} = f(t - \tau, \mathbf{x}) - g(t - \tau, \mathbf{x}); \tau > 0$$

where $\mathbf{x} = (x_1, \dots, x_n)$ represents the concentrations of n gene products S_1, \dots, S_n , t represents the independent variable (often taken to be time), $\frac{d x_j}{d t}$ is the change in x_j over time, $f(\cdot)$ is the function representing rate laws which lead to an increase in x_j and $g(\cdot)$ is the function representing biochemical rate laws which lead to a decrease in x_j quantities. By incorporating a delay τ ,

Danino et al. (2010) call the system of ODEs a system of delay differential equations (DDEs). Technically, the system can be solved through integration or numerical methods to give a solution $x_j(t)$, however, this is often a very complex process and simulations are used instead to visualise the products' trajectories over time without having to solve the ODE. These are discussed later on.

Mass action kinetics, first-order and second order rates, and Michaelis-Menten kinetics are used to represent $f(t-\tau, x)$ and $g(t-\tau, x)$ in terms of the different biochemical reactions occurring in the GRN. Michaelis-Menten equations, in particular Hill functions (Figure 1.3), are used to represent positive and negative regulatory reactions (Danino et al., 2010; Tigges et al., 2009; Fung et al., 2005) and binding reactions (Chandran et al., 2008). Given a gene product x_j , the Hill function for positive feedback is defined as:

$$h^+(x_j, \theta_j, n) = \frac{x_j^n}{x_j^n + \theta_j^n}$$

where θ is the disassociation constant/threshold which represents the value x_j for which the Hill function has value 1/2 (Cinquemani et al., 2008) and n is the Hill coefficient, also referred to as the steepness/co-operativity coefficient, which influences the gradients of the Hill curve (de Jong, 2002). For $n < 1$, the system is said to have negative co-operativity, indicating lack of affinity. For $n > 1$, there is positive co-operativity or multimerisation, depending on the context (Chandran et al., 2008). At $n = 1$, there is no co-operativity. In order to model negative feedback, the function $1 - h^+(x_j, \theta_j, n)$ is used (de Jong, 2002).

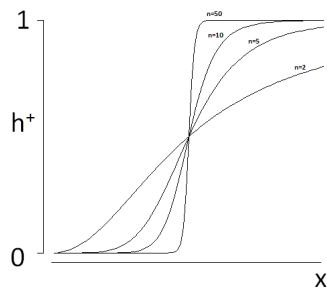


Figure 1.3: Hill function for different values of the hill coefficient, $n = 2, 5, 10, 50$. The x-axis represents the value of the input x , while the y-axis represents the value of the Hill function, h^+ . The dissociation constant is set at 1.

Authors use ODEs, one for each product of interest, to represent the GRN system to a different degree of complexity. Elowitz and Leibler (2000) choose to model proteins and their respective mRNA in terms of the transcription, translation and degradation rates. Fung et al. (2005), on the other hand, choose to model transcription and translation as one main synthesis rate. Stricker et al. (2008), at the other end of the spectrum, develop a very detailed model, accounting for transcription, translation and post-translation processes such as folding, dimerisation and tetramerisation of the proteins involved. Both Tigges et al. (2010) and Elowitz and Leibler (2000) develop second generation models which incorporate a greater level of complexity than the basal simpler model they start with. For example, Elowitz and Leibler (2000) choose to model the binding and unbinding of the protein to the operator in their second model, while Tigges et al. (2010) distinguish between whether the regulatory molecule (tetracycline) was intra- or extracellular. This allowed them to extract a previously undetected relation between oscillation frequency and gene dosage.

1.2.1.7.1.2 Simulating deterministic ODEs. As a system of ODEs increases in complexity and size, direct integration to solve the equation becomes infeasible. Instead, analytical methods such as numerical integration can be used to simulate product trajectories indirectly without explicitly solving the equation. One option is to use the family of single-step, finite difference methods, where the continuous time domain $[0, T]$ is discretised and broken down into a set of small time-points t_i . The value of the ODE is then calculated in iterations. At a time t_i , the value of the ODE (i.e. the change in product quantity) is calculated based on its values at time t_{i-1} . One such example is the Runge-Kutta method as used in Fung et al. (2005). Other methods include the Forward and Backward Euler methods. Alternatively, multi-step finite difference techniques such as the Adams-Basforth algorithms can be utilised, which use concentration levels at multiple previous time steps. However, care needs to be taken when using these methods because they are less stable than single-step ones (Schwartz, 2008).

As an alternative to discrete methods, spectral methods can be used to represent a system of ODEs, as seen in Fung et al. (2005) and Elowitz and Leibler (2000). A Fourier series can be used to represent the concentration of a gene product, $x(t)$, in terms of the frequency ω , rather than time. However, this technique is mathematically demanding.

1.2.1.7.1.3 Assumptions of deterministic ODEs. The representation of GRNs by deterministic non-linear, first-order ODEs is based on a set of assumptions which serve to simplify the system that is under consideration. This is necessary within a biological setting, where explicit modelling of the intra- and extracellular environment, cell cycle and metabolic and regulatory process complexity would quickly render a mathematical model infeasible to design and analyse.

The most common assumption is that there is spatial, thermal and structural homogeneity across the cell and the cell population (Zheng & Sriram, 2010). This allows for simplification of the cell system, since processes such as diffusion can be ignored. It also means that the system is assumed to have only one independent variable, which in most cases is taken to be time (Schwartz, 2008). When this assumption fails, partial differential equations have to be used instead to represent both time and volume as independent variables. Thus, for example, in tissue and embryonic development and diffusion processes, a set of PDEs is used to represent the dynamics and to study pattern formation across both time and space (Zheng & Sriram, 2010; de Jong, 2002). A second common assumption is the ‘continuum assumption’ which states that product concentrations vary continuously and deterministically over time (de Jong, 2003). This allows for changes in protein quantities to be represented by ODEs, which are continuous functions. In actuality, gene product numbers vary discreetly and non-deterministically, which is why discrete stochastic modelling is considered to be more biologically relevant (albeit more challenging).

Other assumptions made to reduce degrees of freedom and variables are specific to the model under consideration. For example, Atkinson et al. (2003) assume that “the rate of translation is directly proportional to the concentration of mRNA”, while Guido et al. (2006) assume that chemical rates, such as transcription factor binding and releasing rates are at equilibrium with each other. Whereas assumptions appear to cast doubt on the validity of the model and its connection to the biological world, Danino et al. (2010), Tigges et al. (2009), Stricker et al. (2008) and Fung et al. (2005), to mention but a few, have reported good agreement between results obtained in experiments and modelling. This shows that assumptions can serve the purpose of simplification

while remaining loyal to the molecular model.

1.2.1.7.2 Stochastic modelling of network dynamics. Stochastic modelling is used instead of deterministic modelling to model network dynamics while introducing an element of randomness and noise to the network to represent intra- and extracellular noise sources (Zheng & Sriram, 2010). Stochastic models represent biological systems more accurately, especially when each cell harbours the network in low gene copy numbers such that natural biological inconsistencies are felt more strongly than in large populations when the impact of noise is averaged out (El Samad et al., 2005). This is known as the ‘finite-number effect’ and is considered to be the main source of intrinsic stochasticity (Kaern et al., 2005, 2003).

Deterministic models can be considered as average representations of stochastic network dynamics in a unitary stable-state system (Kepler & Elston, 2001; Stricker et al., 2008). The stochastic model becomes equivalent to a deterministic one at the hypothetical thermodynamic limit or in systems of a large size (Gillespie, 1976). This does not imply that stochastic models can be bypassed in favour of deterministic ones. In their stochastic model, Elowitz and Leibler (2000) observe variable oscillation amplitudes and a finite autocorrelation time between oscillations, which features were not observed in deterministic simulations. Similarly, in Tigges et al. (2009), deterministic and stochastic simulations yielded trajectories with the same trends but different values.

A crucial difference between a deterministic and stochastic model is that the former considers protein concentration levels over time, whereas the latter considers *the probability* of protein values varying over time. More specifically, a stochastic model represents how the probability of being in a (discrete) state \mathbf{X} , $P[\mathbf{X}(t)]$, changes with time, where a state $\mathbf{X} = [X_1, X_2, \dots, X_n]$ represents the number of molecules of products/species S_1, S_2, \dots, S_n (which can be proteins or mRNA molecules). Moreover, deterministic models are reproducible. Stochastic models, on the other hand, are unlikely to follow the same trajectory twice even when starting from the same initial starting point (de Jong, 2003), although they still tend to follow a general trend.

A selection of the most relevant stochastic modelling terms are defined below:

- the set of possible reactions, R_1, \dots, R_M , which can occur in the GRN. An example of a reaction is the translation or degradation of a protein.
- the vector $\mathbf{v}_i = (v_{i1}, v_{i2}, \dots, v_{in})$, given a reaction R_i , which represents the changes in molecular numbers occurring to the n species being modelled. If reaction R_i occurs, the state $\mathbf{X} = \mathbf{x}$ becomes $\mathbf{x} + \mathbf{v}_i$. The $M \times n$ matrix, \mathbf{v} , is referred to as the stoichiometric matrix.
- the propensity function, $a_i(x)$, which is defined such that $a_i(x)dt = P[R_i \text{ occurs in } [t, t + \delta t] | X(t) = x]$ (Gillespie, 1976).

1.2.1.7.2.1 Master and Langevin equations. The Master Equation is the equivalent of the deterministic ODE and models the probability of a Markov process $X(t)$ being in a state x at time t . Thus, instead of a system of n deterministic ODEs for each species in the network under consideration, one Master Equation is used to represent the entire network system state. The Master Equation, same as a set of ODEs, can be solved through direct integration or Stochastic Simulation Algorithms (SSAs). Alternatively, approximations such as the Fokker-Planck equation can be

used to simplify and solve the Master Equation. However, the Master Equation representation is rarely used in studies because it is not practical to work with due to its complexity. An alternative stochastic representation of the GRN dynamics is the Langevin Equation.

The Langevin equation, unlike the Master Equation, does not model probabilities explicitly and instead models the discrete number of molecules of a species S over time. It is a stochastic differential equation, in that a normally distributed stochastic term is added to a first-order, non-linear (deterministic) ODE, in order to represent intrinsic and extrinsic fluctuations (Fung et al., 2005). The Langevin equation can be derived from the Master Equation or independently. The Chemical Langevin equations for each gene products S_1, \dots, S_n with X_1, \dots, X_n molecules are given by:

$$\frac{dX_i(t)}{dt} = \sum_{j=1}^M v_{ji} a_j(t) + \sum_{j=1}^M v_{ji} \sqrt{a_j(t)} W_j(t)$$

where the first term represents a drift term proportional to the time increment dt , while the second term represents a diffusion term proportional to \sqrt{dt} . $W(t)$ represents white noise as a theoretical notion of stochasticity.

Effectively, the Langevin equation is the summation of deterministic terms with an additional stochastic summation. Thus, the Langevin equation deals with a continuous, Markov process, as opposed to a (discrete) jump, Markov process, which is generally the case in a Master Equation. The Langevin formula is effective given that a time increment dt is chosen such that the propensity function $a_j(t') \simeq a_j(t)$ for $t' \in [t, t + dt]$ is true i.e. dt is small enough so as to not affect propensity values a_j significantly and so that events occur independently of each other according to a Poisson distribution. Moreover, it is assumed that there are multiple occurrences of each reaction R_i in the time interval $[t, t + dt]$ such that the Central Limit Theorem can be employed to approximate the Poisson distribution by a Normal one. These assumptions are satisfied if the system is large, but may be incompatible otherwise (Gillespie, 2000).

The Langevin equation can be solved through numerical algorithms such as the Euler-Maruyama algorithm or simulated using SSAs (Mistry, 2007; Tuttle et al., 2005). The latter are a much more powerful and intuitive tool to use. In fact, studies on synthetic GRNs often bypass the development of a Master and Langevin Equation, and use SSA profiling directly to generate time series trajectories of the products without solving equations directly. Fung et al. (2005) is one exception, using the Langevin equation to represent the metabolator stochastically via the addition of a Gaussian white noise element to the parameters within the system. The equation is then solved directly via the analytical weak Euler scheme.

1.2.1.7.2.2 Stochastic Simulation Algorithms. As already mentioned above, SSAs are used in order to simulate trajectories of a set of species S_1, \dots, S_n for which the change in quantity over time is represented by differential equations showing the biochemical processes which cause the variation. SSAs make use of Monte Carlo (MC) sampling and simulation techniques to map a species S_i 's trajectory over time. This bypasses the use of integration to solve the differential equation.

MC algorithms work by iteratively generating random numbers which in this case are taken to represent change in the dependant and independent variables. This serves to build the time series profile of the species incrementally, using time as the independent variable. When run several times from the same initial condition, an average trajectory output can be extracted (Gillespie, 1976). In addition, statistics such as the mean number of proteins and the autocorrelation function can be generated to give an overview of protein kinetics.

1.2.1.7.2.3 The Gillespie Stochastic Simulation Algorithm. The most widely used family of SSAs in simulation of oscillator GRNs is the Gillespie algorithm and its derivatives. They have been used in Tigges et al. (2009), Stricker et al. (2008), Guantes and Poyatos (2006), Fung et al. (2005) and Elowitz and Leibler (2000) with good agreement with experimental results, even though the original Gillespie algorithm does not take time-delayed reactions specifically into account.

The Gillespie algorithm is based on the ‘inversion method’ MC simulation technique. Given that it can be assumed that the system has thermal and spatial homogeneity, a pair (τ, j) is randomly generated. Given the system is at a time t , τ represents when the next reaction, taken to be R_j , will take place at time $t + \tau$. The protein and mRNA numbers are then updated with v_j from the stoichiometric matrix at this time-point according to the reaction R_j which takes place, and the iteration is repeated until an upper time limit T is reached. The Gillespie SSA makes use of a joint probability density function (pdf), $p(\tau, j|x, t)$ for the time increment τ and reaction R_j . This is a normalised, continuous function over $[-\infty, \infty]$. The pdf is defined as the probability that a reaction R_j will be the next and only reaction to occur in $[t + \tau, t + \tau + \delta\tau]$ (Gillespie, 1977) where δ is a small increment. The probability has value equal to $a_j \exp[\sum_{i=1}^M a_i \tau]$. Thus, the number of molecules $X_1(t), \dots, X_n(t)$ of species S_1, \dots, S_n which are dependant on the reactions R_1, \dots, R_m can be simulated over time.

The algorithm consists of the following steps (Gillespie, 1976):

Step 1: Given thermal and spatial homogeneity, set the time variable $t = 0$. Store initial values for the n variables $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$. Store M propensity values a_1, a_2, \dots, a_M of the M reactions. Calculate the probability density function $p(\tau, j)$. Define the final time point, T .

Step 2: Randomly generate a pair (τ, j) using the Direct Method (see below). This entails generating two numbers r_1 and r_2 from the uniform distribution $U(0, 1)$ and calculating the pair:

$$\tau = \frac{1}{a} \ln\left(\frac{1}{r_1}\right)$$

$$j \in \{1, \dots, M\} \text{ s.t. } \sum_{i=1}^{j-1} a_i < ar_2 \leq \sum_{i=1}^j a_i$$

Step 3: Update t to $t + \tau$ and x_i to $x_i + v_{ji}$ for $i = 1, \dots, n$. Recalculate the probabilities a_j .

Step 4: Store new values X_1, X_2, \dots, X_n . If $t > T$ or $a_j = 0$ for all possible j , Stop. Else Repeat from **Step 2**.

In Step 2, the Direct Method is used to first generate τ , then j , based on the fact that $p(\tau, j) = p_1(j|\tau)p_2(\tau)$ where $p_2(\tau) = a \exp(-a\tau)$ and $p_1(j|\tau) = a_j/a$, using $a = \sum_{j=1}^M a_j$ (Gillespie, 1976).

1.2.1.7.2.4 Variations of the Gillespie algorithm. The Gillespie algorithm assumes that processes (reactions) have similar time-scales (El Samad et al., 2005). It has been suggested that a biological system can have fast-slow dynamics (Cinquemani et al., 2008) and that 'fast' dynamics, such as dimerisation cause the Gillespie algorithm to become inefficient (Tuttle et al., 2005). These and computational efficiency issues have been addressed through the development of modified algorithms which are based on the original Gillespie algorithm and are either more computation and time efficient or more precise.

Alternative exact (where it is assumed that only one reaction occurs per iteration) and approximate (where it is assumed that more than one reaction occurs per iteration) simulation algorithms have been developed. These include the Next Reaction Method (Gillespie, 1977) which assesses when each and every reaction will next take place and then picks the one which occurs closest to the current time, the Tau Leaping Method which uses a pre-defined time-step τ (Gillespie, 2001), the Next Reaction Method which can be used when rates are time-dependant (Gibson & Bruck, 2000), the Optimised Direct Method and the Slow Scale SSA which takes advantage of fast-slow reactions to simplify the set of reactions under consideration (Cai & Wang, 2007). The Delay Stochastic Simulation Algorithm (DSSA) (Barrio et al., 2006) and the binomial Tau-leap DSSA (Leier et al., 2008) are used when the system is non-Markovian and take into account time-delay due to biochemical reactions.

Alternatively, hybrid SSAs can be used (Salis & Kaznessis, 2005). In these algorithms, both deterministic and stochastic techniques are implemented. One such hybrid model, developed for prokaryotic biological systems, models biochemical processes such as transcription as stochastic processes while others such as degradation are represented by deterministic processes (Cinquemani et al., 2008). Similarly, in another hybrid model, faster reactions are assumed to be continuous and are modelled using Langevin equations while the slower reactions are modelled using differential Jump equations (Tuttle et al., 2005).

1.2.1.7.3 Qualitative analysis of dynamics. Following quantitation of network dynamics, in-depth qualitative analysis is often carried out in order to assess network properties such as stability and robustness. The three main methods used in analysis of oscillatory GRNs are briefly discussed.

1.2.1.7.3.1 Bifurcation analysis. Bifurcation analysis is the study of bifurcation parameters and the changes in network dynamics which they cause. A bifurcation of equilibrium occurs when there is a qualitative change in the dynamics of the system such as the onset of oscillations or bistability due to a change in stability or the type of equilibrium. This is the result of a change in the value of a key parameter known as a bifurcation parameter which crosses a threshold value and, for example, causes the GRN to start exhibiting oscillations as in the case of a supercritical Hopf bifurcation. In synthetic GRNs, a common bifurcation parameter is time-delay. In light of this, it is easier to understand why changing time-delay within a GRN's architecture influences protein expression: it can push the system towards or away from stable equilibrium and can be used in

order to tune GRN dynamics (Guanes & Poyatos, 2006). Thus, bifurcation analysis is carried out in order to understand how critical parameters affect protein dynamics.

In order to study bifurcations it is necessary to study the equilibria (steady-states, critical points) of the network. Mathematically, an equilibrium is defined as a trajectory or point, $\phi \in \mathbf{C}^+$, where $\phi = \bar{x}$ for some $x \in \mathbf{R}^{+n}$ satisfying $f(\mu, \bar{x}) = 0$ which is a system of non-linear first order ODEs for species S_1, \dots, S_n with quantities represented by x_1, \dots, x_n (Wang et al., 2005). Multiple analysis techniques such as vector fields (Leite & Wang, 2010), Poincare maps and use of the centre manifold approach (Kuznetsov, 2004) are available to analyse equilibria. However, a more standard practice is to analyse the system's characteristic polynomial and eigenvalues. The latter can be deduced by using the Jacobian matrix, J or through use of the Taylor series (Xiao & Cao, 2008). Based on the values of the eigenvalues, the type of bifurcation which takes place can be deduced.

1.2.1.7.3.2 Hopf bifurcations. There are multiple bifurcations which bring a qualitative change in network kinetics, including a saddle node, pitchfork and flip bifurcations. A Hopf bifurcation causes a non-linear system to transition to or from oscillations as seen in repressilators (Strelkowa & Barahona, 2010; Elowitz & Leibler, 2000), the one-component Goodwin oscillator (Goodwin, 1963), dual-feedback oscillators (Guanes & Poyatos, 2006; Atkinson et al., 2003) and the metabolator (Fung et al., 2005). It should, however, be noted that in theoretical work, oscillations have been found to occur through other bifurcations, notably the saddle node bifurcation (Guanes & Poyatos, 2006). In order to define a Hopf bifurcation mathematically, the following theorem is used (Seydel, 1994):

Theorem 1. In a two-dimensional system, a birth of a limit cycle occurs at equilibrium (\bar{x}, μ) if, at the critical value $\frac{d\bar{x}}{dt} = 0$, also known as the Hopf bifurcation point/threshold:

1. a pair of the Jacobian matrix's simple conjugate eigenvalues $\sigma_1 = \bar{\sigma}_2 = \varepsilon + i\eta$ for $j = 1, 2$ have $\varepsilon = 0$, with no other eigenvalue having zero real part
2. $\frac{d\varepsilon}{dt} \neq 0$.

This is known as a Hopf bifurcation. The initial period is $\bar{t}_0 = \frac{2\pi}{\eta}$.

In an n -dimensional system, for $n > 2$, a Hopf bifurcation occurs when, in addition to the above conditions, the remaining eigenvalues have negative real parts. Bifurcations can be visualised using bifurcation diagrams (Stricker et al., 2008) and phase portraits (Fung et al., 2005). Note that given non-zero real parts, the system would result in a stable or unstable focus for negative and positive real parts respectively, wherein the trajectory either spiral towards or away from a node. When the real part has value 0, the spiral transforms into a limit cycle (Lakshmanan & Rajaseekar, 2012; Izhikevich, 2007).

There are two types of Hopf bifurcation; supercritical and subcritical (Byrne et al., 2014). In a supercritical bifurcation, a stable equilibrium bifurcates to a stable limit cycle, while in a subcritical bifurcation, an unstable limit cycle disappears in the equilibrium. The difference between a super- and subcritical bifurcation is whether the system is non-linearly stable or unstable at $\varepsilon = 0$. In order to analyse the onset of limit cycles following a Hopf bifurcation, it can be more convenient

to represent the system of differential equations using the normal or polar-form representation, although this approach is less conventional.

1.2.1.7.3.3 Stability and robustness analysis. Stability analysis complements bifurcation analysis when studying equilibria. In the case of synthetic oscillators, stability analysis serves to assess whether limit cycles and periodic solutions are stable (unchanging over time) or sensitive to changes in parameter values. Thus, stability analysis is concerned with the long-term dynamics of the model under study and the likelihood of the model changing dynamics following small perturbations. This translates to the study of homeostasis and robustness. Mathematically, a solution $\overline{x(t)}$ is **stable** if for each $\varepsilon > 0$ and $t_0 \in \mathbf{R}$ there exists $\delta = \delta(\varepsilon, t_0) > 0$ such that if $x(t)$ is a solution and $|x(t_0) - \overline{x(t_0)}| < \delta$ then $|x(t) - \overline{x(t)}| < \varepsilon$ for all $t > t_0$ (Grant, 1999) i.e the path does not diverge, irrespective of small perturbations within parameter values.

The simplest form of steady state analysis can be carried out by solving for $\frac{dx_j}{dt} = 0$ in a set of ODEs. This was done in Elowitz and Leibler (2000) to assess the stability of different parameter values, including the basal translation rate and the level of co-operativity in the Hill equation for gene transcription. Floquet theory is also widely used to assess the stability and quality of periodic solutions. It has been used in studies by Strelkowa and Barahona (2010) and Fung et al. (2005) to deduce quality of oscillations in the n -gene repressilators and the metabolator respectively.

In conjunction, or as an alternative, Poincare maps can be utilised to deduce results on stability. As with bifurcations, analysis is carried out on the model's eigenvalues. In this case, the eigenvalues of the Monodromy matrix (which will not be defined here) are calculated and analysed. Using this technique, by varying one parameter (the glycolytic rate), Fung et al. (2005) were able to assess the glycolytic conditions under which the network is more stable. The authors do not develop this analysis further, stating that it may not help in the explanation of network dynamics. This highlights the importance of using mathematical analysis as a means to an end, rather than an end in itself.

In line with stability analysis, the aim of robustness analysis is to check whether oscillatory dynamics are maintained for all parameter combinations. In most studies, analysis is carried out by simulating time-series wherein a parameter is perturbed by a small amount, ϵ , which can be random or pre-defined. By assessing the generated time-series using the modified parameters, a deduction can be made as to whether the network is susceptible to changes in a particular parameter or not (Tigges et al., 2009; Stricker et al., 2008).

Another technique for the analysis of robustness uses temporal logic to represent network behaviour as a set of conditions and check whether combinations from the parameter space lead to oscillations which satisfy these conditions. In the method developed by Batt et al. (2007), robustness conditions, ϕ , are represented by a set of atomic propositions for each species S_i in terms of quantities x_i ($x_i > a(t)$, $x_i < b(t)$) and logic operators (*and*, *or*, *not*) (for example, $\phi = x_i > a \vee x_j > b$). Then, given an n -dimensional parameter space P , which represents the ranges within which parameters can vary, a system is said to be robust if P is valid for ϕ , meaning that every parameter combination in P satisfies the condition (behaviour) ϕ . In order to carry out a validity check, numerical methods are used. The first step is to discretise the n -dimensional state space P , based on the solutions of the differential equations and to divide it into sections (Batt

et al., 2007). Afterwards, points from each section are selected and sampled, and if each sample satisfies the conditions, the network is said to be robust.

1.2.1.7.4 Parameter estimation of GRN biochemical reaction rates. Prior to the development of a model based on counteracting biochemical rates, all biochemical rate parameters being used in the model must be defined. It is rare for all of the model's rate parameters such as the transcription rate, folding rate and degradation rate to be known. Often, researchers use reported values and assumptions in literature sources to select parameter values (Stricker et al., 2008; Atkinson et al., 2003). Alternatively, parameter values can be determined via experiments, although this is a time-costly exercise.

As a more efficient measure, parameter estimation can be carried out through the use of iterative algorithms. Once the observed dynamics are quantified, computational protein trajectory data generated from the estimated parameter values can be compared to them via a difference measure, and if the two trajectories are sufficiently agreeable, the selected parameters are retained. Otherwise, a new parameter set is selected iteratively until the generated dynamics only differ from observed dynamics by a pre-specified difference, ϵ .

Tigges et al. (2009) initially make use of assumptions to minimise unknown parameter values, such as assuming that the degradation rates of all mRNA molecules are the same. Following this, an MC-based parameter estimation algorithm is applied to find parameter values which result in oscillations of a specific period (15 hours). The algorithm is based on an evolutionary strategy based on weighted least squares of the difference between observed and computational data. In cases where knowledge on parameters is very limited, substitute approaches such as the black box approach (Zheng & Sriram, 2010) can be used, wherein parameters are rendered dimensionless and abstract and no longer represent biochemical kinetics.

There is a wide range of parameter estimation algorithms available, including evolutionary algorithms, gradient search optimisation algorithms and scatter searches (Zheng & Sriram, 2010; Dasika & Maranas, 2008; Feng et al., 2004). Algorithms based on Bayesian parameter estimation are a more recent development (Liepe et al., 2014; Filippi et al., 2013; Barnes et al., 2011b, 2011a). These are MC based searches which, rather than inferring one value for an unknown parameter, estimate a range of values for each parameter, making the estimated data set more biologically realistic. Bayesian statistics are based on the proviso that a probability density function is fitted to the parameter of interest. Over subsequent iterations, this density function is refined and narrowed until a range of probable values for the parameter is extracted. In particular, the Barnes group use the sequential MC Approximate Bayesian Computation algorithm (ABC SMC), which they use to estimate parameter values in toggle switches, oscillators and large-size networks.

1.2.2 The value of broadening the range of synthetic biology chassis

As synthetic biology gains momentum, research is being carried out to address the lack of choice there is with regards to network chassis via two main approaches. The first is the engineering and

manipulation of established chassis in order to perform required functions and make for a more efficient culturing process (Church et al., 2014; Xiao et al., 2014). This requires extensive chassis engineering. The second option is to look beyond the traditional hosting platforms towards the rich domain of life in the natural world and search for organisms which may already possess more desirable physiological features which allow for the desired function than established host cells. This implies that less genetic manipulation is needed to achieve greater robustness and optimal performance within the cell (Zhu et al., 2012).

With regards to optimising traditional host cells, research is looking into the minimisation and 'stripping-down' of organisms by deleting long stretches of the genome to leave only the essential components and machinery (Xiao et al., 2014; Gao et al., 2010a; Keasling, 2008; Glass et al., 2006). However, due to the lack of knowledge which is available on genome structure and function, the process of genome minimisation remains empirical (Hirokawa et al., 2013).

An alternative approach is the development of cell-free systems, such as the *in vitro* cell-free oscillator by Kim and Winfree (2011). The driving principle is to pick the required replication and metabolic machinery (enzymes and molecules) from cells and use them within an *in vitro* setting (Forster & Church, 2006). Proponents of this approach suggest that this allows for more direct and easier control of the GRN (for example, adding inducer molecules) and bypasses the use of host cells which gain no evolutionary advantage by incorporating the synthetic network (Hodgman & Jewett, 2012; Kwok, 2010). The use of cell-free networks has a long history in biocatalysis, and has been used in the production of hydrogenases (Boyer et al., 2008) (which are of interest in biofuel research) and vaccine fusion proteins (Kanter et al., 2007). However, cell-free systems are not financially feasible when compared to *in vivo* systems, due to the necessity of providing expensive energy sources and metabolites (Smith et al., 2014).

With regards to the diversification of the choice of available host cells, there has recently been developmental work on the use of plant chassis within synthetic biology for application in medicine, biofuel and bioprocessing (Medford & Prasad, 2014; O'Connor & Brutnell, 2014). Efforts have also been made to establish a standardised practice within plant hosts using BioBricks (Boyle et al., 2012). There are multiple other ongoing studies to establish novel organisms as synthetic biology chassis. These include the minimal mycoplasma *Mesoplasma florum* (Baby et al., 2013; Panke, 2005) and *Mycoplasma genitalium* (Heinemann & Panke, 2006), the minimal plant species *Marchantia polymorpha* (<http://synbio.org.uk/marchantia/>), the gram-negative bacterium *Pseudomonas putida* (Nikel et al., 2014), the gram-positive bacterium *Bacillus subtilis* (Radeck et al., 2013) and microalgae strains (Gimpel et al., 2013) and cyanobacteria strains (Berla et al., 2013) for the generation of biofuel. There has even been an iGEM project looking into the use of the frog species *Xenopus tropicalis* (<http://2012.igem.org/Team:Evry>) as a host. Thermophile organisms such as the *Thermoanaerobacter* and *Clostridium* species have recently garnered attention for the production of proteins which would be toxic in other hosts and for their ability to withstand extreme bioprocessing and environment conditions (Menezes et al., 2015; Berla et al., 2013).

All of the organisms discussed above face challenges and shortcomings. For example, thermophiles retain a low protein yield (Lin & Xu, 2013), cyanobacteria present challenges on how to gather (solar) energy efficiently (Berla et al., 2013) and plant cells may be very sensitive to the networks inserted (Yang et al., 2013a).

There is no definitive list of requirements which a host cell should be able to meet, although Smith et al. (2014) suggest five different criteria; enabling control of GRN and host cell behaviour, straightforward monitoring of synthetic pathway dynamics, withstanding high-throughput optimisation techniques, broad application and cost-competitiveness, the latter of which is emphasised in several reviews (Medford & Prasad, 2014; Berla et al., 2013). As further research is carried out, Kelwick et al. (2014) propose that since extensive data has been gathered on model organisms from across all the tree of life, these could be tested as host platforms within synthetic biology.

1.2.2.1 The *Roseobacter* marine bacteria genus

Despite the vast bacterial diversity in the world's oceans, the majority of identified marine bacteria fall into nine major clades (Buchan et al., 2005). The *Roseobacter* clade comprises 20% of coastal and 15% of mixed-layer ocean bacterio-plankton communities (Buchan et al., 2005; Gonzalez et al., 1999). *Roseobacter* bacteria are extremely versatile and have been detected in the most diverse of marine environments, from oceanic upper mixed layers to deep sea, polar sea ice, hypersaline lakes, saline soil and even gold mines (Brinkhoff et al., 2008; Buchan et al., 2005). The genomes of the species vary from 3.1Mbp to 5.4Mbp and plasmids range in size from 4.3Kb to 821.7Kb (Brinkhoff et al., 2008). Petersen et al. (2013) report that over 40 strains in this clade have had their genome sequenced, and work in this area is ongoing (Buddruhs et al., 2013; Riedel et al., 2013). In fact, these bacteria are so diverse and flexible that studies are still looking to establish a solid taxonomic ranking (Newton et al., 2010).

In spite of the diverse environments these organisms have been found in, strains share over 89% of the 16S rRNA gene and there are common physiologies pertaining to the majority of the *Roseobacter* strains, as summarised by Moran et al. (2007) and Buchan et al. (2005) and on which information is also still being gathered (Hahnke et al., 2013; Lenk et al., 2012). These common traits include but are not limited to: the ability to generate energy via an aerobic, anoxygenic, phototrophic process which gives them a competitive advantage over neighbouring organisms; the ability to break down organic and inorganic sulfur compounds which may have an impact on global sulfur cycles; the ability to oxidise carbon monoxide and degrade aromatic compounds to use as primary growth substrates; the ability to survive as free-living organisms but also to form symbiotic relations with algae, cephalopods, fish and dinoflagellates (note that they have also been implicated as pathogens in oyster and coral diseases); and the ability to produce secondary metabolites such as poison, antibiotics and quorum sensing signals for defence and cell-cell signalling purposes. The *Roseobacter* clade use quorum sensing based on the luxI-luxR mechanism (Zan et al., 2014), which is also widely used in synthetic biology.

1.2.2.1.1 Novel capabilities *Roseobacter* offers as a synthetic biology chassis. The above highlights how, compared to traditional synthetic biology platforms, *Roseobacter* use a different set of pathways and mechanisms in several metabolism and energy-generation processes, even bypassing the Calvin cycle in some strains (Swingley et al., 2007). This results in the ability of *Roseobacter* batch cultures to flourish and use mixotrophic metabolism in spite of substrate limitations (Sato-Takabe et al., 2014; Swingley et al., 2007).

This has led to interesting research developments which showcase the potential of the *Roseobacter* clade in biotechnology and bioengineering. *Roseobacter* are among the most widely studied marine bacteria, are easily cultivated and cultured and are naturally programmed to uptake foreign DNA and integrate it into their own genome (Brinkhoff et al., 2008; Buchan et al., 2005). Commercially available culturing medium is on the market, and culturing conditions for the model strain *Roseobacter denitrificans* require a 30°C temperature (Piekarski et al., 2009). This strain has a doubling time of 2.2 hours with the doubling time of other strains such as *Parvularcula bermudensis* as low as 1.2 hours (Christie-Oleza et al., 2012). The versatility of *Roseobacter* and ability to survive in extreme and non-replete conditions gives them an advantage over the widely used *E. coli* bacteria.

Following the 2010 Deepwater Horizon oil spill in the Gulf of Mexico, research carried out on bacterial communities showed *Roseobacter* strains to be resistant to oil poisoning since cells contained hydrocarbon-degrading genes (hydrocarbons are predominantly found in crude oil) (Lamendella et al., 2014; Lu et al., 2012). Research has also highlighted the ability of *Roseobacter* bacteria to form homologous biofilms (Mitra et al., 2014; Elifantz et al., 2013) and they have been shown to precipitate phosphate through the formation of crystals from waste waster (phosphate purification is a necessary and costly process in waste water treatment) (Rivadeneyra et al., 2014).

The role which the *Roseobacter* clade and marine bacteria in general can play within bioprocessing and biotechnology is starting to garner more attention within academia and industry. There are discussions pertaining to the use of marine bacteria in therapies to tackle infectious diseases (Williams, 2009), the production of antibacterial compounds (Bruhn et al., 2007) and the production of bioenergy via hydrogen gas generation (Mirza et al., 2013). Moreover, due to the natural marine and extreme habitats and their various metabolic properties of *Roseobacter*, they can be used to address both ocean bioremediation challenges which synthetic biology is seeking to tackle (Zhang & Nielsen, 2014; Dash et al., 2013; Schmidt, 2010), and challenges in geo-engineering, a nascent field (Calvert et al., 2014) which aims to moderate climate change via the use of technology (Rayner et al., 2013). Discussions on the potential benefits and impacts of geo-engineering are gaining momentum (Hardman-Mountford et al., 2013; IPCC, 2012). For example, the 2012 iGEM project by University College London (UCL) (http://2012.igem.org/Team:University_College_London) focused on the bioremediation of oceans. 50-80% of sea-debris stranded on beaches, floating on the ocean surface and on the seabed is made up of plastic (Hidalgo-Ruz & Thiel, 2013; Barnes et al., 2009). The team proposed that this issue be tackled by engineering *Roseobacter* marine bacteria to host polyethylene degradation enzymes such as laccase (Santo et al., 2013), enabling the organisms to function as detection-and-degradation ocean patrollers.

The above discussion on innate processes and features of *Roseobacter* strains serves to bring to attention the potential of *Roseobacter* bacteria as chassis within synthetic biology. As of yet, this potential has not been tested.

1.2.2.1.2 The challenges of using *Roseobacter* bacteria in synthetic biology. Protocols have been established to allow for cloning (Irie et al., 2010), transfection (Piekarski et al., 2009) and genome modification of strains (Jaschke et al., 2011). However, in consideration of the potential application of *Roseobacter* organisms within synthetic biology and the principles of the field,

the applicability of these protocols across the *Roseobacter* spectrum is unknown. Moreover, no investigations have been carried out into whether these marine bacteria will stably uptake synthetic plasmids and retain them over subsequent generations, although it appears that *Roseobacter* can integrate DNA into their genome to enable adaptability (Brinkhoff et al., 2008).

Since research being carried out on *Roseobacter* stems from the field of marine biology, there are currently no engineering principles governing *Roseobacter* research. Thus, for example, the origins of replications used in GRNs which are compatible with *Roseobacter* are unknown, as is the plasmid size threshold for synthetic plasmids which can be stably inserted, the orthogonality of synthetic plasmids with respect to the cell, and whether protocols can be simplified and abstracted as to be applicable to all strains within the clade. Moreover, there appears to be no quantitative or qualitative modelling of *Roseobacter* systems. These questions need to be answered before *Roseobacter* bacteria can be established as a synthetic biology platform.

1.2.2.2 The *Trypanosomatida* order

Trypanosomatida are a highly divergent eukaryote order consisting of diverse unicellular protozoan species (Calvo-Alvarez et al., 2015). The order includes genera of parasites found in humans, animals, insects, birds and even plants (Jackson, 2014).

1.2.2.2.1 *Leishmania tarentolae* in bioprocessing. *Leishmania tarentolae* (*L. tarentolae*) is a member of the *Trypanosomatida* order which is pathogenic to gecko lizards (Raymond et al., 2012) and which has been commercialised as a host cell in bioprocessing. It combines the robustness of *E. coli* with the complex post-translational processes of higher eukaryotes (Sugino & Niimi, 2012). The commercialised *L. tarentolae* expression system is a product of Jena Bioscience GmbH, and can be used to express recombinant ‘constitutive or inducible, intracellular or secretory ... proteins’ (Jena Bioscience GmbH, n.d.). Since *L. tarentolae* is not pathogenic to mammals, it has a Level 1 Biosafety grading and does not require specialised apparatus. Suppliers claim that the whole process from construction to purification of protein takes six weeks (Breitling, 2013). The system has only been tested with LEXSY plasmid vectors, which are also developed by the company. These are expression systems which are *L. tarentolae* oriented and can be used *in vivo* or even *in vitro* (Jena Bioscience GmbH, n.d.).

The LEXSY platform has been used in several projects which include the expression of antibody fragments (Jorgensen et al., 2014), the expression of the FVII protease which has blood coagulation functions (Mirzaahmadi et al., 2011) and the expression of a *Trypanosoma cruzi* (another strain within the *Trypanosomatida* order) apurinic/apirimidinic endonuclease within a *T. cruzi* host (Sepulveda et al., 2014), showing that the LEXSY system may be applicable across the *Trypanosomatida* spectrum. LEXSY plasmid vectors have also been used in research on how the immune system responds to *Leishmania* parasites (Chamakh-Ayari et al., 2014), how *Leishmania* parasites respond to environmental cues (Dacher et al., 2014), on the use of *L. tarentolae* as a vaccine-producing alternative host to hen eggs (Pion et al., 2014) and proof-of-principle projects to clone a tetracycline-dependent system into the *Leishmania mexicana* strain (Kraeva et al., 2014) and to use fluorescent reporters for facilitated and rapid screening purposes (Vacchina & Morales, 2014). Although these were done primarily with the aim of being applied within the field

of immunology, they pave the way for use within bioengineering.

L. tarentolae and *Trypanosomatida* have characteristics and culturing properties which makes them amenable to bioprocessing. Protocols and GMP guidelines have already been developed for *Leishmania* strains (Melville, 2010). Additionally, protozoa benefit from growth to high cell densities (up to 1×10^9 cells), control of a neutral internal pH while experiencing pH 5-7.4 externally, low shear stress sensitivity (Fritsche et al., 2007), a rapid growth rate (circa 6.7 hours) both in liquid media and on agar plates (Basile & Peticca, 2009) and the use of cheap fully-defined media supplemented only with hemin (Simpson et al., 1991).

Cultures produce up to 500mg/l of recombinant protein (Breitling, 2013). In addition, cultures are capable of retaining stability for more than 50 passages (Fritsche et al., 2007). Several *Trypanosomatida* strains have been shown to have mammalian-like and homogeneous glycosylation processes (Basile & Peticca, 2009), which allows for mammalian recombinant proteins and pharmaceutical quality level proteins to be expressed (Breitling et al., 2002; La Flamme et al., 1995).

1.2.2.2.2 *Trypanosoma brucei* as model *Trypanosomatida*. *Trypanosoma brucei* (*T. brucei*) is a unicellular, motile, protozoan species with a single flagellum which moves in a corkscrew-like motion (Langousis & Hill, 2014). They belong to the *Trypanosomatida* order, which has ten different genera (Podlipaev, 2001) including *Leishmania*, discussed above, and *Trypanosoma*, to which *T. brucei* belongs. *T. brucei brucei* is non-pathogenic to humans due to the presence of trypanosome lytic factor (TLF) in human blood, but is pathogenic to livestock (Samanovic et al., 2009).

T. brucei are considered to be a model organism for the research of biological processes (Serrichio & Butikofer, 2011) since they are the most widely studied organism within the *Excavata* eukaryote superdomain (one of the six to have been classified) in which the *Trypanosomatida* are nested. Forward and reverse genetic techniques have been successfully used on *T. brucei* (Verner et al., 2010). Trypanosome parasitology has also resulted in established protocols to culture and manipulate *T. brucei* via standard genetic engineering cloning techniques (Clayton & Shapira, 2007; Clayton, 1999). An efficient trypanosomal cloning system has recently been developed to facilitate quick plasmid construction (Batista et al., 2010), along with techniques for genome modification, gene over-expression and RNA interference (Bouvier et al., 2013; Mansfield & Paulnock, 2008).

Trypanosomes are distinct from higher eukaryote and mammalian systems, both in their structure and the expression mechanisms used (Daniels et al., 2010; Basile & Peticca, 2009; Gruszynski et al., 2006; Palenchar & Bellofatto, 2006). This was made clearer with the completion and publication of the *T. brucei* genome (Berriman et al., 2005).

1.2.2.2.1 The life cycle of *T. brucei*. *T. brucei* switches between two predominant morphological forms, depending on whether they are within their insect carrier or host, as seen in Figure 1.4. *T. brucei* switches between a proliferating short, stumpy form known as the **procyclic** form when it exists in a benign form inside the (carrier) insect host gut, and a second long, slender and proliferating **bloodstream** form when it exists as a pathogen within blood vessels and the central

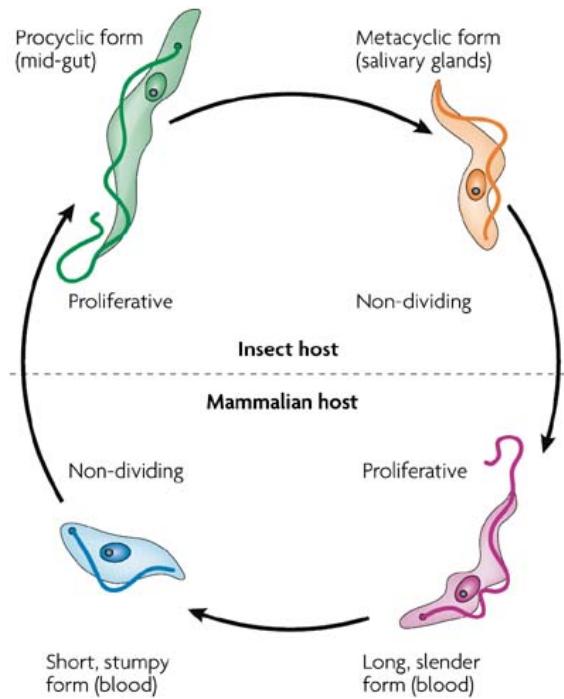


Figure 1.4: The life cycle of *T. brucei*, which changes morphology depending on the host (insect vector or mammalian host) and the state (proliferating or not) it is in. When in the insect host's mid-gut, the trypanosome is in its proliferative procyclic form, which morphs into a non-proliferative, metacyclic form as the trypanosome moves to the insect's salivary gland in preparation for transfer into the mammalian host's bloodstream when the vector insect bites the host. Once in the host, the trypanosome changes to a proliferative, bloodstream form which then switches to a non-dividing, stumpy form as it prepares to be transported back to another vector host via the bloodstream when the vector bites the human host. The image is reproduced from Lee et al. (2007).

nervous system in its host (Langousis & Hill, 2014). During the differentiation process between procyclic and bloodstream forms, *T. brucei* also differentiates into two intermediate forms (referred to as metacyclic and non-dividing forms) as it prepares to either pass from the host to the vector while in the blood vessels or vice-versa as it prepares to be transported from the vector to the host by relocating to the vector's salivary glands (Clayton, 1999). In both cases the transfer from host to vector or vice-versa occurs when the vector bites the host. Trypanosomes can replicate both during the bloodstream and procyclic form stages, but not during the intermediate stages (Gruszynski et al., 2006).

The two main life stages not only lead to a different morphology but also to the expression of different sets of protein. Thus, whereas metabolic enzyme expression increases during the procyclic stage, expression of surface immune-evasion proteins is up-regulated during the bloodstream form stage (Jensen et al., 2014). Whereas bloodstream form trypanosomes express variant surface glycoproteins to coat themselves, procyclics express EP and GPEET proteins to coat themselves (Matthews, 2005). Moreover, bloodstream form cells makes use of quorum sensing mechanisms in preparation of transmission into the carrier host (Langousis & Hill, 2014) and as the number of cells increase to prevent over-population (Matthews, 2005). Energy generation is also based on different sources and mechanisms; whereas bloodstream form trypanosomes use glycolysis, procyclic trypanosomes used a respiratory system based on the mitochondrion (Matthews, 2005).

1.2.2.2.2.2 Transcription in *T. brucei*. Transcription in trypanosomes is polycistronic, allowing for the transcription of multiple genes, similarly to prokaryotes (Breitling et al., 2002). Mammalian cells, by contrast, tend to have monocistronic gene clusters (Blumenthal, 2004). *T. brucei* genes within the same cluster do not necessarily share the same patterns (Clayton, 1999) or expression levels (Jensen et al., 2014), unlike bacterial counterparts. This is indicative of post-transcriptional regulatory processes, discussed below.

Whereas protein-coding genes in higher eukaryotes are generally expressed using RNA polymerase II (Palenchar & Bellofatto, 2006), trypanosomes use a combination of RNA polymerase I, II and III. Most importantly, the former transcribes proteins specific to the life cycle stage and all variant surface glycoproteins (VSGs) (Park et al., 2011) which are surface proteins which coat trypanosomes as part of the immune evasion strategy, as discussed below. The use of RNA polymerase I is believed to enable the rapid switching of these surface coats, giving it an advantage over the host's immune system (Park et al., 2011). Unique to *T. brucei* is an Expression Site Body (ESB) which is primarily involved in the polymerase I mediated transcription of trypanosome surface proteins and which allows for strict and rapid switching of proteins being expressed (Palenchar & Bellofatto, 2006). By contrast, RNA polymerase II transcribes mRNA and polymerase III transcribes tRNAs, 5S RNA and snRNAs, as seen in other organisms (Palenchar & Bellofatto, 2006).

Transcription is generally followed by 5' trans-splicing (Breitling et al., 2002) and 3' polyadenylation (Freire et al., 2014) to allow for the division of pre-mRNAs from the polycistronic transcribed bodies into monocistronic mature mRNA. Note that in *S. cerevisiae* and higher eukaryotes, co-transcriptional splicing is more common than 5' trans-splicing (Park et al., 2011).

1.2.2.2.3 Translation in *T. brucei*. Post-transcriptional reactions are important within trypanosomes since protein regulation occurs predominantly at this level (Bouvier et al., 2013). Translational efficiency also varies between the different life stages of the trypanosome, further implying that translational processes play an important role (Jensen et al., 2014) in the regulation of protein synthesis. Thus, whereas transcription mechanisms are similar to those in prokaryote systems, translation mechanisms are more similar to eukaryote systems (Vasquez et al., 2014).

Gene regulation within trypanosomes is governed by a combination of mechanisms. For example, upstream ORFs within 5' UTRs, some of which include an ATG codon, have been shown to regulate translation (Siegel et al., 2011) while RNA stability regulating motifs have also been identified (Jensen et al., 2014), as has a pattern in codon usage which is implicated in translational efficiency (Vasquez et al., 2014). In addition, transcripts have half-lives which differ over an order of two magnitudes due to the influence of the 3' UTR (Siegel et al., 2011) and several proteins which control translation initiation have been identified (Freire et al., 2014).

1.2.2.2.4 Variant surface glycoproteins as an antigen variation mechanism in *T. brucei*. Trypanosomes use an antigen variation mechanism to avoid detection by the host (Rudenko, 2011). Trypanosome cells are covered with a mono-layer of a variant surface glycoprotein (VSG). A mono-layer can have over 5×10^6 dimers (Smith et al., 2009). The VSG is chosen from a reservoir of over 1000 VSGs and pseudo-genes at the disposal of the trypanosome. The VSGs

are similar in structure but antigenically different (Hutchinson et al., 2007). The reservoir is much larger in size than that of other organisms which employ antigen variation mechanisms (Stockdale et al., 2008), accounting for over 10% of protein within *T. brucei* (Breitling et al., 2002). The trypanosomes can even piece together parts of different pseudogenes in order to synthesise novel functional genes known as 'mosaic' genes (Marcello & Barry, 2007).

The VSG mono-layer is continuously replaced through a low frequency selection process, in order to repeatedly avoid neutralisation by the host's immune system (Hutchinson et al., 2007). A selection of the VSGs are placed in expression sites found at the telomere via recombination or transcription-based switching (Morrison et al., 2009; Stockdale et al., 2008). However, it appears that only one of these sites is active and bound to the ESB extranucleolar site which contains transcription and RNA processing machinery (Aitcheson et al., 2005). When bound, the site actively represses the remainder of the sites such that only one VSG is fully expressed (Horn & McCulloch, 2010). The polymerase I mediated transcription mechanism transcribes VSG genes approximately 50 times more quickly than β tubulin genes (Park et al., 2011), which goes towards explaining why *T. brucei* can switch its coating so efficiently. The VSG coat is crucial to the viability of the cell such that inhibition of expression results in cell cycle arrest (Smith et al., 2009).

1.2.2.2.3 Novel capabilities *T. brucei* offers as a synthetic biology chassis. *T. brucei* are highly divergent eukaryotes that make use of mechanisms which are found in both prokaryotes and eukaryotes. This combination may give *T. brucei* a competitive edge over more traditional chassis in terms of doubling times, yields and ability to synthesise complex proteins. While research has mainly focused on the parasitic nature of trypanosomes, the work carried out on *L. tarentolae* within bioprocessing has shown that *Trypanosomatida* are worth investigating as chassis for the synthesis of complex proteins within industrial and biotherapeutics.

T. brucei, as a model organism about which information available is more detailed than for other strains within the order, would be the first organism to be tested. Moreover, the tools and techniques which will allow for the necessary investigative studies are already in place. Research has led to the development of protocols to culture and manipulate *T. brucei* via standard genetic engineering cloning techniques (Clayton & Shapira, 2007; Clayton, 1999), thus allowing for genome modification, gene over-expression and RNA interference (Bouvier et al., 2013; Mansfield & Paulnock, 2008) during the investigation. Work has also been done to streamline the cloning and protein production process as seen in Batista et al. (2010) and Dortay and Mueller Roeber (2010). Calvo-Alvarez et al. (2015) and Bouvier et al. (2013) developed a series of modular fluorescent and tagging proteins to bypass antibody technology, while Ma et al. (2012) developed a vector to allow for rapid degradation of fused genes. The *T. brucei* genome has been sequenced, which will allow for better selection and manipulation of genes (Berriman et al., 2005). Insights from these studies may be applied in synthetic biology and bioprocessing. For example, the VSG and ESB regulatory mechanism, which allow for the transcription of a number of VSG genes but only the full synthesis of a single protein, can be applied within genome engineering.

A step towards the systematic study of trypanosomes has been the development of TrypanoCyc (www.metexplore.fr/trypanocyc/), which is a database of metabolic pathways for *T. brucei* built using algorithms based on literature and genome sequence data. Plans have also been drawn to develop an in-depth *in silico* mathematical model of a trypanosome, a project known as the

‘silicon trypanosome’ (Bakker et al., 2010). Moreover, a metabolic pathway of bloodstream form trypanosomes has been developed (Bakker et al., 1997) and used to model the effect of drugs and compounds on the pathway and the glycosylation process, leading to effective ‘model-driven experimental design and hypothesis-driven systems biology’ (Bakker et al., 2010).

The above has also shown that trypanosomes have unique innate mechanisms which positions them as prime candidates for use within biotechnology, bioengineering and bioprocessing. Trypanosomes use polycistronic transcription and sophisticated post-transcriptional capping and regulation, have a two-stage life cycle which allows for the up-regulation of different sets of genes, different culturing conditions and antigen variation mechanisms based on switching between VSG coat mono-layers. Several other interesting features can be found. For example, glycolysis within trypanosomes occurs in a specialised organelle, known as the glycosome, resulting in a metabolism pathway which is unique to these organisms (Lee et al., 2007).

1.2.2.4 The challenges of using *T. brucei* as a chassis in synthetic biology. Trypanosomes have not yet been used as host cells in synthetic biology projects. For trypanosomes to become a viable platform option within synthetic biology, a lot of groundwork will need to be laid. Important technical data and techniques which are required in synthetic biology is missing. As of yet, trypanosome mathematical models still need to rely heavily on parameter estimation, due to a lack of data regarding biochemical parameters. The robustness of trypanosome strains has not been established. Previous work on trypanosomes which used fluorescent proteins did not require live, single-cell analysis, as is often necessary in characterisation processes in synthetic biology. Thus, there are no protocols in place to track single cells.

The establishment of *T. brucei* as a novel synthetic chassis will require both developments to the existing technology and increased and more in-depth knowledge of *T. brucei* mechanisms.

1.3 Aims and Objectives

The aim of this project is to establish *Roseobacter* clade marine bacteria and *T. brucei* as synthetic biology chassis. This work addresses the gap within synthetic biology resulting from the limited choice of host cells which are available for use in practice. Specifically, this will be done via three main channels: by developing synthetic bacterial and trypanosomal GRNs which function as an oscillator, by developing the necessary protocols and set-ups to allow for the characterisation of the GRN within the organism, and by analysing the GRN’s dynamics within the host. This will set the foundation for the use of the two organisms as chassis.

Roseobacter marine clade bacteria are naturally found in various diverse oceanic habitats, making them an ideal chassis for the application of synthetic biology to bioremediation challenges in the world’s oceans. *Roseobacter* also have an important ecological role in balancing global carbon fixation levels and as such, hold potential to be used as a tool in geo-engineering activities.

Protozoa are beginning to be exploited as host cells for industrial production of biotherapeutics

due to their fast doubling times and mammalian-like post-translational glycosylation of host proteins. As an established model organism for studying the protozoa, *T. brucei* provides a useful test case for establishing synthetic biology in this phylum for the first time. *T. brucei* is highly divergent from eukaryotes commonly used in synthetic biology to date and possesses a sophisticated genomic machinery to evade host immune systems by rapid switching of surface antigen expression. The establishment of standard synthetic biology approaches in mathematical modelling and synthetic gene network design in *T. brucei* will underpin the future application of synthetic biology to both enhance the industrial capability of the protozoa as a class of host chassis and to probe the pathobiology of this organism.

1.3.1 Aim 1: Establishing *Roseobacter* clade marine bacteria as a synthetic biology chassis by identifying reliable molecular biology procedures for construction of a synthetic gene network.

The objectives of this investigation take a bifocal approach, namely to:

1. Establish straight-forward molecular biology procedures in *Roseobacter* clade bacteria.

The objective is to investigate the tractability of *Roseobacter* marine bacteria and to establish whether different strains from the *Roseobacter* marine prokaryote clade can be modified, cloned and chemically/electrically transformed via the use of standardised synthetic plasmids, BioBricks™ and via reproducible, standard protocols which are applicable across the whole clade. Moreover, the objective is to ensure that the developed protocols are easy to follow and allow for the involvement of both professional and citizen scientists in synthetic biology.

2. Characterise and model an *E. coli* oscillator plasmid in *Roseobacter*. Upon establishment of robust molecular biology techniques, a plasmid encoding the standard Goodwin oscillator, following the scheme reported by Stricker et al. (2008), is used to transform competent *Roseobacter* cells. The oscillatory phenotype will then be characterised and modelled mathematically.

This investigation will lead to insights on the potential of *Roseobacter* to be used in addressing challenges in bioremediation and geo-engineering. Secondly, the development of standardised cloning techniques will enable the participation of do-it-yourself research communities (DIYbio) and foster the democratisation of synthetic biology.

1.3.2 Aim 2: Establish *T. brucei* as a synthetic biology chassis by designing, modelling, constructing and measuring a Goodwin oscillator encoded by a novel synthetic gene network.

The objectives of this investigation take a multi-focal approach, namely to:

1. Design a synthetic trypanosomal Goodwin oscillator. Develop a synthetic network based on the autonomous TetR negative feedback loop (Hillen & Berens, 1994). The network uses modified Green Fluorescent Protein (eGFP) as a fluorescent reporter output. The network

incorporates the main design features which are required for oscillations: dominant negative feedback as per the TetR loop, and an element of time-delay and non-linearity, implemented in the synthesis process.

- 2. Assemble DNA fragments to encode a synthetic trypanosomal Goodwin oscillator.** Construct a trypanosomal synthetic auto-inhibiting tetracycline-inducible oscillator which mimics the Goodwin oscillator. This will be done through standard and Gibson cloning (Gibson & Bruck, 2000) techniques within *E. coli*. The network will then be inserted via transfection into both bloodstream and procyclic form *T. brucei*.
- 3. Parameterise gene network components and network dynamics to inform modelling and characterise the network via live cell imaging.** Characterise the GRN with initial preliminary analysis carried out via protein detection techniques in order to confirm function of individual network components. Following this, live single cell imaging using fluorescent and brightfield microscopy will be carried out to determine whether the GRN functions as an oscillator and to quantify observed dynamics.
- 4. Develop an experimental set-up for live-cell imaging of single motile trypanosome cells.** Develop a mechanical set-up to allow for live-cell imaging of single motile trypanosome cells, as no solution is currently available on the market and in literature sources. Due to the motility of cells, a movement-constricting device needs to be developed without compromising cell viability and which allows for the tracking and imaging of a single cell over a large span of time. This work will draw from microfluidics, plant and parasite biology.
- 5. Mathematically model the synthetic Goodwin oscillator.** Develop a mathematical model of the GRN which complements characterisation work. Prior to analytical characterisation efforts, a system of differential equations which represent protein production patterns will be developed and sensitivity analysis carried out via stochastic time-series Gillespie simulations in order to generate protein production trajectories and establish the likelihood of observing oscillations and the feasibility of going ahead with characterisation studies. Following live cell imaging, qualitative analysis will be carried out in order to assess stability within observed dynamics.

This investigation will provide insight into the establishment of *T. brucei* as a synthetic biology platform which can be exploited in the industrial and biotherapeutic sectors.

Chapter 2

Materials and Methods

2.1 *Roseobacter* marine bacteria strains and culturing conditions

2.1.1 *Roseobacter* bacterial strains

Three different marine bacteria strains from the *Roseobacter* clade were used; *Roseobacter denitrificans* OCh114 (Allgaier et al., 2003), *Oceanobulbus indolifex* HEL-45 (Wagner-Dobler et al., 2004) and *Dinoroseobacter shibae* DFL 12 (Shiba, 1991). The strains were obtained from NCIMB Ltd. (Aberdeen, Scotland).

2.1.2 *Roseobacter* culturing conditions

For the culturing of all *Roseobacter* strains, Marine Broth (MB) and MB agar (both Becton, Dickinson & Company, Le Pont de Claix, France) were used for liquid and static cultures respectively, as per supplier instructions. Some protocols required the use of half-concentrated MB agar. This was prepared by dissolving 18.7g of MB powder (instead of the standard 37.4g) in one litre H₂O, adding agar (Sigma-Aldrich, Munich, Germany) to a concentration of 1.5% (w/v), and sterilising the solution by autoclaving (Piekarski et al., 2009).

If required, the solutions were supplemented with antibiotic as detailed below. Shaker cultures were incubated in a Kuhner ISF-1-V Climo-Shaker Incubator (Adolf Kuhner AG, Basel, Switzerland) at 30-37°C at a speed of 200rpm, while static cultures were incubated at 30-37°C in a Memmert High Precision Incubator (Memmert GmbH, Schwabach, Germany) with the petri dish upside down to avoid condensation droplets compromising the cells.

2.1.3 Glycerol stock preparation for *Roseobacter* strains

6ml of sterile MB was inoculated with 6 μ l of the *Roseobacter* strain of interest from a previous glycerol stock or with a colony picked from an MB agar plate on which the strain had been spread

and grown for 12-16 hours at 37°C in an incubator. Following inoculation, the solution was incubated for 12-16 hours in a shaker-incubator at 37°C and rotation of 200rpm, after which it was supplemented with sterile 80% glycerol (Fisher Scientific, MA, U.S.A) at a ratio of 1:4 for glycerol:culture. This was then divided into 380 μ l aliquots and stored at -80°C.

2.1.4 Antibiotics used for culturing of *Roseobacter* strains

Throughout all *Roseobacter* investigations, the following selective agents were added to the media at the appropriate working solutions. 100mg/ml Ampicillin (Fisher Scientific, MA, U.S.A) was prepared by dissolving 1g Ampicillin in 10ml 0.33M NaOH, filter-sterilising the solution and storing in 500 μ l aliquots at -20°C. It was used at a working concentration of 100 μ g/ml. 34mg/ml Chloramphenicol (VWR International Ltd., Leicestershire, U.K) was prepared by dissolving 0.34g of Chloramphenicol in 10ml 100% ethanol and storing in 500 μ l aliquots at -20°C. It was used at a working concentration of 17 μ g/ml. 50mg/ml Kanamycin (Sigma Aldrich, Munich, Germany) was prepared by dissolving 0.5g of Kanamycin in 10ml dH₂O, filter-sterilising the solution and storing in 500 μ l aliquots at -20°C. It was used at a working concentration of 25 μ g/ml.

2.2 Recombinant DNA techniques in *E. coli*

The following section lists *E. coli* strains, recipes and cloning techniques used in the construction of the plasmid encoding the Goodwin oscillator and all studies related to the investigation of *Roseobacter* as a synthetic biology chassis which also required the use of *E. coli* strains.

Throughout all of the cloning processes, plasmid sequences were analysed using ApE - A plasmid Editor(version 2.0.38) (ApE, 2012), which can be used to create graphic maps, analyse restriction digest sites and align sequences, amongst others. All recombinant DNA techniques were carried out at the Department of Biochemical Engineering, UCL.

2.2.1 Plasmids used for construction of the oscillator networks

The pHD1313 and pRP_{eGFP}SIR2rp3 plasmids were provided by Dr. Sam Alsford and Dr. David Horn from the London School of Hygiene and Tropical Medicine (LSHTM). The pHD1313 sequence data was provided via private correspondence by Prof. Christine Clayton from the Center for Molecular Biology, Heidelberg. Commercially prepared plasmids were ordered from Eurogentec Ltd. (Hampshire, U.K).

2.2.2 *E. coli* strains used in *Roseobacter* investigations and the construction of the oscillator networks

Three different strains of *E. coli* were used throughout cloning. The first is the K-12 W3110 non-commercial strain (Hayashi et al., 2006). The second and third strains were two commercial

brands of highly efficient chemically competent *E. coli* cells. These were the *E. coli* One Shot[®] TOP10 cells (Life Technologies Ltd., Paisley, U.K), with a transformation efficiency of 1×10^9 cfu/ μ g plasmid DNA (Life Technologies Ltd., n.d) and the *E. coli* XL10-Gold[®] Ultracompetent cells (Agilent Technologies U.K Ltd., Chesire, U.K) which have a transformation efficiency of over 5×10^9 transformants/ μ g pUC18 DNA (Agilent Technologies UK Ltd., n.d).

2.2.3 *E. coli* culturing conditions

For liquid cultures, *E. coli* cells were grown in LB medium in tubes or flasks which had at least five times the capacity of the culture volume, due to aeration requirements. One litre of LB was prepared using 10g tryptone, 5g yeast and 5g NaCl (all Sigma-Aldrich, Munich, Germany). The solution was then autoclaved. Cultures were grown at 37°C in a shaker-incubator at an agitation of 200rpm.

For static cultures in petri dishes, cells were streaked or spread on solidified LB agar and incubated at 37°C, with the petri dish upside down to avoid condensation droplets affecting the cells. One litre of LB agar was prepared by using 10g tryptone, 5g yeast, 5g NaCl and 15g of agar. The solution was then autoclaved.

If required, the solutions were supplemented with the correct antibiotic, detailed below.

2.2.4 Glycerol stock preparation for *E. coli* strains

E. coli glycerol stocks were prepared as per Section 2.1.3, replacing MB with LB.

2.2.5 Antibiotics used for culturing of *E. coli* strains

Ampicillin was used as a selective marker for all *E. coli* strains. Ampicillin was prepared as per Section 2.1.4.

2.2.6 Generating chemically-competent *E. coli* W3110 cells

Non-competent W3110 cells were streaked from a glycerol stock onto a minimal agar plate and incubated for 16 hours at 37°C. 50ml of minimal agar was made using 1ml filter sterilised 20% (w/v) glucose (Sigma Aldrich, Munich, Germany), 50 μ l filter sterilised 2mg/ml thiamine (Sigma Aldrich, Munich, Germany), 5 μ l filter sterilised 1M CaCl₂ (Sigma Aldrich, Munich, Germany), 100 μ l 1M MgSO₄ (VWR International Ltd., Leicestershire, U.K), 10ml sterile 5X M9 salts and 39ml sterile 1.4% (w/w) agar. 500ml of 5X M9 salts was prepared by using 32g of Na₂HPO₄ (Sigma Aldrich, Munich, Germany), 7.5g of KH₂PO₄ (Sigma Aldrich, Munich, Germany), 1.25g of NaCl and 2.5g of NH₄Cl (Sigma Aldrich, Munich, Germany).

Following incubation, one colony was picked from the minimal agar plate into 5ml LB media. This was supplemented with 100 μ l 1M MgSO₄ and then incubated for 12-16 hours at 37°C in a shaker-incubator and an agitation of 200rpm. From this solution, 1ml was used to inoculate 100ml LB, which was then placed in a 37°C shaker-incubator and shaken at 200rpm until cells were at early log phase of growth and the optical density at 600nm measured 0.3. At this point, the solution was placed on ice for a further 10 minutes and then centrifuged at 4000rpm for 5 minutes at 4°C. As with all medium-volume (larger than 2ml) centrifugation steps, the Avanti J-E Centrifuge (Beckman Coulter, High Wycombe, U.K) was used. The supernatant was discarded and the pellet was resuspended with 10ml ice-cold 0.1M filter sterilised CaCl₂/15% sterile glycerol and placed on ice for 30 minutes, after which the centrifugation was repeated, the supernatant discarded and the pellet resuspended in 1ml 0.1M CaCl₂/15% sterile glycerol. The volume was divided into ten 100 μ l aliquots which were stored at -80°C.

2.2.7 Chemical transformation of chemically-competent *E. coli* cells

In all transformations detailed below, following the final incubation step, the number of colonies on each plate were counted, and individual, clear, round colonies were picked into separate 2ml Ampicillin-selective LB and shaken for 12-16 hours at 200rpm and a 37°C temperature. From each of these, glycerol stocks were prepared as per Section 2.2.4, while a purification and analytical digest was carried out on the remaining 1.5ml culture as per sections 2.2.8 and 2.2.10.1 respectively in order to analyse plasmid identity.

Two controls were used to ensure a correct transformation procedure was followed. To check whether cells are viable, a positive control where no selective marker is added to the LB agar was used. To check whether cells were contaminated with antibiotic resistant strains, a negative control where no plasmid DNA is added to the competent cells was used.

2.2.7.1 Transformation of chemically competent *E. coli* W3110 cells

A vial of chemically competent W3110 cells was taken from storage at -80°C and thawed on ice. 1-8 μ l of plasmid DNA/ligation mix was pipetted onto the surface of the frozen cells and the mixture was incubated on ice for 45 minutes. This was followed by a heat-shock for 10 minutes in a 37°C water bath. After this, 1.3ml LB was added to the cell solution, the volume was transferred to a 15ml tube and placed in a 37°C shaker with an agitation of 200rpm for one hour to allow for cell membrane repair and antibiotic gene expression. The solution was then centrifuged for two minutes at maximum speed at room temperature using a Microcentrifuge 22R Centrifuge (Beckman Coulter, High Wycombe, U.K). After, the supernatant was removed and the pellet re-suspended with 100 μ l LB and spread on two Ampicillin-selective LB agar plates at a 1:9 ratio. The plates were inverted and incubated for 12-16 hours at 37°C.

2.2.7.2 Transformation of chemically competent *E. coli* One Shot® TOP10 cells

A 50 μ l vial of TOP10 cells stored at -80°C was thawed on ice and 1-5 μ l of plasmid DNA/ligation mix pipetted onto the cells. The solution was mixed gently and incubated on ice for 30 minutes.

The heat-shock was carried out by placing the vial in a 42°C water-bath for exactly 30 seconds. Following this, the vial was placed on ice, without mixing or shaking. 250 μ l pre-warmed S.O.C media supplied with the cells was added to the mixture. The vial was then placed on its side in a shaker-incubator and agitated for one hour at 200rpm. Finally, 10 μ l and 90 μ l of this mixture were spread onto two selective LB agar plates. The plates were inverted and incubated at 37°C for 12-16 hours.

2.2.7.3 Transformation of chemically competent *E. coli* XL10-Gold® cells

One litre of NZY+ broth was made using 10g of casein hydrolysate (Sigma Aldrich, Munich, Germany), 5g of yeast extract and 5g of NaCl. The solution was then autoclaved. Prior to use, the solution was supplemented with 12.5ml filter sterilised 1M MgCl₂ (Sigma Aldrich, Munich, Germany), 12.5ml filter sterilised MgSO₄ and 20ml of 20% (w/v) filter sterilised glucose.

A vial of XL10-Gold® cells was taken from -80°C and thawed on ice. A 100 μ l aliquot was pipetted into a pre-chilled 15ml Falcon tube (Fisher Scientific, MA, U.S.A) to which 4 μ l β -mercaptoethanol supplied with the cells (concentration not specified) was added. The tubes were swirled and placed on ice for 10 minutes with further gentle swirling every two minutes. Afterwards, 2 μ l of plasmid DNA/ligation mix was added, the solution mixed and the tubes incubated on ice for 30 minutes. The heat shock was applied by placing the vial in a 42°C water-bath for exactly 30 seconds. The tube was then placed on ice for two minutes and supplemented with 0.9ml of preheated NZY+ broth, after which it was incubated for one hour at 37°C, shaken at 200rpm. A 10 μ l and 90 μ l aliquot was then spread onto two selective LB agar plates. These were inverted and placed in a 37°C incubator for 12-16 hours.

2.2.8 Plasmid DNA elution and purification from *E. coli* cultures, plasmid DNA digests, PCR reactions and gel electrophoresis runs using commercial kits

This section lists all the different commercials kits used to purify and elute plasmid DNA.

The KeyPrep Spin Plasmid DNA Mini Kit Pk100 (Anachem Ltd., Luton, U.K) was used to elute less than 20 μ g of plasmid DNA (miniprep) from *E. coli* cultures, following the miniprep protocol as supplied with the kit. In order to elute 20 μ g-1mg plasmid DNA from *E. coli* cultures, the Qiagen Plasmid Maxi Kit or HiSpeed Plasmid Maxi Kit (both Qiagen, Crawley, U.K) was used as per supplier instructions.

Following DNA preparative digests, dephosphorylation reactions and PCR runs, DNA was purified from enzymes and reagents using the QIAquick PCR Purification Kit (Qiagen, Crawley, U.K), which can purify DNA strands between 100bp and 10kb in size, as per supplier instructions. If DNA was to be purified from a gel, the fragment was first extracted using the White/UV Transilluminator (UVP LLC, CA., U.S.A) and then purified from the encasing gel using the QIAquick Gel Extraction Kit (Qiagen, Crawley, U.K). The protocol supplied with the kit was followed without

modifications.

Following all purification protocols, plasmid concentration and purity were determined through spectrophotometry, using the Thermo Scientific NanoDrop 1000 Spectrophotometer (Fisher Scientific, MA, U.S.A).

2.2.9 Plasmid DNA purification from plasmid DNA digests using phenol-chloroform extraction in preparation for *T. brucei* transfection

Prior to transfection of the oscillator plasmid into *T. brucei*, the (digested) plasmid was purified using phenol-chloroform methods to ensure sterility and prevent the accumulation of salts which ensue from the use of commercial kits. This in turn served to avoid arcing during the transfection process.

The first step was to add an equal volume of phenol-chloroform (Phenol:Chloroform:Isoamyl Alcohol 25:24:1 Saturated with 10mM Tris, pH 8.0, 1mM EDTA; Sigma Aldrich, Munich, Germany) to the volume of digested DNA. This was vortexed for five seconds using the Vortex-Genie® 2 (Scientific Industries Inc., NY, U.S.A). The solution was then spun for one minute at 10,000rpm using a table-top Eppendorf Centrifuge 5415D (Eppendorf, Stevenage, U.K), the top aqueous layer was pipetted to a fresh tube and 400 μ l ice-cold 100% ethanol (Fisher Scientific, MA, U.S.A) was added to it. A five minute centrifugation at 10,000rpm was then carried out, after which the supernatant was discarded and replaced with 1ml ice-cold 70% ethanol. The centrifugation was then repeated, after which the supernatant was discarded and a final one minute centrifugation at 10,000rpm was carried out to remove residual ethanol. After removing the supernatant and air-drying the pellet for 10 minutes, 10 μ l of dH₂O was added and the pellet stored at -20°C.

2.2.10 Plasmid DNA digests

In both digests detailed below, in order to determine whether the digest had been successful, a control was used during the gel electrophoresis stage, wherein 5 μ l of the uncut plasmid being analysed was run alongside the digested samples in order to assess whether a difference can be noted between the two. All enzymes, enzyme buffers and Bovine Serum Albumin (BSA) were supplied by NEB.

2.2.10.1 Analytical DNA digests of less than 1 μ g of plasmid DNA

Analytical digests using up to 1 μ g of DNA (volume of 10 μ l) were run to confirm plasmid identity. A 10 μ l analytical digest was set up using 5 μ l plasmid solution, 10% (v/v) restriction enzyme solution, 10% (v/v) enzyme buffer, 0.5 μ l BSA and, if required, 10% (v/v) of the second restriction enzyme. This was topped up to 10 μ l with dH₂O. The solution was incubated at the enzyme's specified incubation temperature for 3 hours in an Eppendorf Thermomixer® comfort (Eppendorf, Stevenage, U.K). If necessary, this was followed by a 25 minute heat-inactivation period, also using the Eppendorf Thermomixer® comfort. The heat-inactivation temperature used in this stage depended

on the enzyme being applied. Following this, the digest was analysed via gel electrophoresis (Section 2.2.12).

2.2.10.2 Preparative DNA digests of 1-10 μ g of plasmid DNA

Preparative digests using 1-10 μ g of DNA (volume of 50-200 μ l) were run to prepare cassettes for subsequent cloning steps, such as ligation. A preparative digest was set up using 50-150 μ l of DNA, 10% (v/v) restriction enzyme, 10% (v/v) enzyme buffer, 10% (v/v) of the second restriction enzyme (optional) and 1-3 μ l BSA, rounded to a specific volume with dH₂O. Following preparation, the same incubation and heat-inactivation procedure as an analytical digest was followed. Afterwards, if the digest was to be gel-purified, the whole volume was run on a gel (Section 2.2.12) and purified (Section 2.2.8). Otherwise, a 5 μ l sample was analysed via gel electrophoresis in order to confirm success of digestion and purified as per Section 2.2.8.

2.2.11 Dephosphorylation of 5'-ends of plasmid DNA

In all dephosphorylation procedures detailed below, following treatment, gel electrophoresis was carried out as per Section 2.2.12 to confirm DNA was not degraded during the reaction. This was followed by a plasmid purification as per Section 2.2.8.

2.2.11.1 Dephosphorylation using Calf Intestinal Alkaline Phosphatase enzyme (NEB)

The NEB CIP kit consists of the CIP enzyme at a concentration of 10,000u/ μ l and 1X NEBuffer 3. Two different methods were used to carry out this reaction. In the first protocol, DNA was mixed in an eppendorf tube with 20 μ l of 1X NEBuffer 3 and 0.5 units of CIP per μ g of DNA. This was then briefly centrifuged, incubated at 37°C for one hour and heat-inactivated for 25 minutes at a 65°C temperature. As an alternative more time-efficient procedure, and given that incubation temperatures and use of buffers permitted, a preparative digest and CIP-treatment were carried out simultaneously by loading digestion and CIP reagents into the same eppendorf tube. The solution was then incubated at 37°C for three hours and heat-inactivated for 25 minutes at 65°C.

2.2.11.2 Dephosphorylation using Calf Intestinal Alkaline Phosphatase enzyme (Promega)

For CIAP, a 50 μ l dephosphorylation reaction was prepared. The 1u/ μ l phosphorylation enzyme was diluted to a concentration of 0.01u/ μ l and added to 10 μ g digested, purified DNA at a rate of 0.01u CIAP per pmol of DNA. Following this, 10X Alkaline Phosphatase Reaction Buffer supplied with the enzyme was added to a final concentration of 1X and if required, the volume was rounded to 50 μ l with dH₂O. Once prepared, the mixture was incubated for 30 minutes at 37°C, after which the volume was topped up again with the same amount of CIAP as added previously, and incubated for a further 30 minutes at 37°C.

2.2.11.3 Dephosphorylation using Shrimp Alkaline Phosphatase (NEB)

For dephosphorylation using the Shrimp Alkaline Phosphatase (SAP) enzyme, digest and dephosphorylation reactions were carried out simultaneously. The SAP kit was supplied with 10X Antarctic Phosphatase Reaction Buffer. $2\mu\text{l}$ of this was mixed with $10\mu\text{g}$ DNA, 10%(v/v) restriction enzyme, the appropriate NEBuffer for digestion to a final concentration of 1X and $5\mu\text{l}$ of Antarctic Phosphatase. The solution was topped up to $20\mu\text{l}$ with dH₂O if necessary. The mixture was incubated for 15 minutes at 37°C and heat-inactivated for five minutes at 65°C.

2.2.12 Agarose gel electrophoresis

2.2.12.1 Preparation and electrophoresis of agarose gel

One litre of 10X TAE buffer was prepared from 48.5g of Tris base (Sigma Aldrich, Munich, Germany), 11.4ml of glacial acetic acid (VWR International Ltd., Leicestershire, U.K) and 3.7g of EDTA, disodium salt (Sigma Aldrich, Munich, Germany). The gel was prepared by adding 1.5g agarose (Sigma-Aldrich, Munich, Germany) to 150ml 1X TAE in a conical flask. The solution was heated to dissolve the agarose. Once cooled to c. 40°C, the solution was stained with $20\mu\text{l}$ Ethidium Bromide and poured into the gel cast to solidify.

HyperLadder I (Bioline, London, U.K) was used as a DNA size marker, while the 5X Sample Loading Buffer supplied with it was used as a tracking dye as per supplier instructions. The Power Pac Basic (Bio-Rad, CA, U.S.A) equipment was used to run the gel at 100V for 70 minutes until the tracker dye line was assessed visually to have migrated 90% across the length of the gel.

2.2.12.2 Electrophoresis analysis

Gels were imaged using the Gel Doc 2000 camera (Bio-Rad, CA, U.S.A). The Quantity One Basic software (Bio-Rad, CA, U.S.A) was used to manipulate the exposure and light settings when imaging the gel.

2.2.13 Polymerase Chain Reaction

To amplify DNA fragments using either complementary or mutagenic primers, $50\mu\text{l}$ Polymerase Chain Reaction (PCR) were carried out using Phusion® High-Fidelity DNA Polymerase (NEB, MA, U.S.A). The polymerase was part of a kit which also contained 5X HF Phusion Buffer, 5X GC Phusion Buffer, DMSO and 50mM MgCl₂. In addition, 10mM Deoxynucleotide Solution Mix (NEB, MA, U.S.A) was used.

Following recipe preparation as per Table 2.1 and a brief centrifugation to collect all volumes at the bottom of the tube, reactions were run in a Veriti 96-Well Fast Thermal Cycler (AB Applied Biosystems, CA, U.S.A), as per Phusion® kit recommendations and as outlined in Table 2.1.

Ingredient	Amount
1ng/ μ l DNA template	1-3 μ l
10pmol/ μ l Forward primer	2 μ l
10pmol/ μ l Reverse primer	2 μ l
Phusion® Polymerase	0.5 μ l
5X HF Phusion buffer	10 μ l
10mM Deoxynucleotide Mix	1 μ l
DMSO	1.5 μ l
dH ₂ O	to 50 μ l
Total	50 μ l

(a) Recipe used for PCR

Step	Description	Temperature	Duration
1	Initiation	98 °C	30s
2	30 cycles of: Denaturation Annealing Extension	98 °C 55-72 °C 72 °C	10s 30s 30s/kb
3	Final extension	72 °C	10 mins
4	Hold	4 °C	∞

(b) Thermocycling conditions

Table 2.1: Set-up and thermocycling conditions used for the PCR-amplification of DNA template using either complementary or mutagenic primers.

A list of all primers used in PCR reactions can be found in Appendix B.

2.2.13.1 Controls used in Polymerase Chain Reaction

As a positive control, a plasmid having a SEAP insert and PCRBlunt2TOPO backbone was amplified using the forward primer (5'-3')CAGGAATAGGCTGGCGAAGCGC and the reverse primer (5'-3')CAGTCAGTGCCCCGGTAAGCCC. This served to amplify a segment of the SEAP sequence, which results in a fragment of size 292bp. As a negative control, a reaction was run using a blank instead of DNA template. This served to assess whether the enzymes and buffers being used were contaminated.

2.2.14 Ligation of two or more DNA fragments

2.2.14.1 Directional, non-directional & three-fragment ligations using Quick T4 DNA ligase

Directional and non-directional two-way ligations were carried out using both a 3:1 and a 1:1 insert-to-backbone ratio. Three-fragment ligations were similarly prepared using 3:3:1 and 1:1:1 insert-to-insert-to-backbone molar ratios. Molar values for DNA were estimated using the formula

$$mols = \frac{\text{weight in ng}}{\text{length in kilo base pairs}}$$

The Quick Ligation™ Kit (NEB, MA, U.S.A), which consists of Quick T4 DNA ligase enzyme and 2X

	Control A	Control B	Control C	Control D	Control E
10mol dephos., cut, pur. 8kb backbone at a concentration of 20ng/ μ l	-	-	-	4 μ l	4 μ l
10mol cut, pur. 8kb backbone at a concentration of 20ng/ μ l	4 μ l	4 μ l	-	-	-
10mol uncut 10kb backbone at a concentration of 20ng/ μ l	-	-	5 μ l	-	-
Quick T4 DNA Ligase	1 μ l	-	-	1 μ l	-
2X Buffer ligase	10 μ l				
dH ₂ O (to 21 μ l)	6 μ l	7 μ l	6 μ l	6 μ l	7 μ l

Table 2.2: **Controls used in standard, non-directional and three-fragment ligations.** Each control is prepared using a combination of dephosphorylated cut purified backbone OR cut purified backbone OR uncut backbone, Quick T4 DNA Ligase (optional), 2X Buffer Ligase and dH₂O to add up to 21 μ l volume. In this example, all backbone aliquots are assumed to have a concentration of 20ng/ μ l. The cut backbone is assumed to be 8kb in length, while the uncut backbone is 10kb long. 10 mols of backbone is used in each case. 'dephos.' is short for dephosphorylated while 'pur.' is short for purified.

Quick Ligation Reaction Buffer, was used to carry out ligations of volume 21 μ l, using at least 50ng of purified backbone. Recipe variations of all ligation reactions are listed in Chapter 4. Following preparation all mixtures were briefly centrifuged, incubated at room temperature for five minutes and then either placed directly on ice prior to transformation as per Section 2.2.7 or stored at -20°C.

2.2.14.1.1 Controls used in ligation. Several controls were carried out in order to validate ligation results. An example of the controls carried out is found in Table 2.2. In all other cases a similar set-up was followed with the volumes of the backbone and dH₂O aliquots differing depending on the mass, length and concentration of plasmid used. The same molar value of backbone plasmid needs to be used throughout the ligation and control reactions to ensure comparability of results.

If the backbone was dephosphorylated, as when carrying out a non-directional ligation, all five controls A-E were carried out. Otherwise, as when carrying out a standard directional and three-fragment ligation, only controls A-C were used. The colony counts of controls A and B (which use digested purified backbone with/without ligase respectively) can be compared to those of the ligation reaction to quantify the amount of re-ligated plus undigested backbone and undigested backbone respectively. Control C, using uncut backbone vector, was used to assess the competency and viability of cells. Controls D and E which used dephosphorylated digested purified backbone vector with/without ligase were compared to the ligation reaction plates to assess the efficiency of the dephosphorylation reaction.

2.2.14.2 Gibson assembly®

The Gibson assembly method (Figure 2.1) developed by Gibson et al. (2009) improves on the standard method in that ligations do not require ends of the insert and backbone DNA to be restriction sites. In addition, due to the nature of the primers used throughout, the assembled plasmid does not run the risk of having the insert in reverse direction, as can happen in non-directional ligations. The assembly method involves less steps than the standard assembly method. The Gibson Assembly® Master Mix (NEB, MA, U.S.A) was used.

The insert was prepared via PCR amplification using 20-40 base pair long primers, which also serve to attach overhangs (Figure 2.1a). The forward primer attaches overhangs which are complementary to the 15-25 nucleotides upstream of the insertion site *on the backbone* to be used for the construct. Similarly, the reverse primer attached overhangs which are complementary to 15-25 nucleotides downstream of the insertion site *on the backbone* to be used for the construct. In addition, primers can have mutagenic or additional inserts such as restriction sites placed between the backbone-complementary and insert-complementary sequences. The backbone itself was prepared either via digest (Section 2.2.10.2) or via amplification (Section 2.2.13) using complementary primers (Figure 2.1b). Both insert and backbone were purified as per Section 2.2.8.

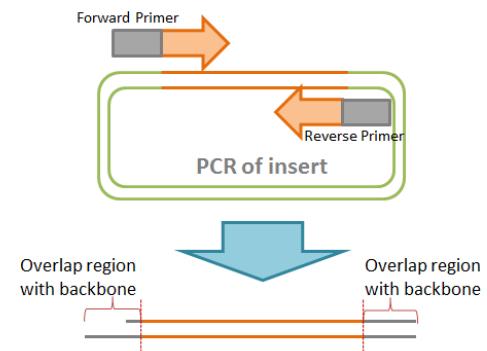
The ligation was carried out by mixing the insert and backbone DNA with the Gibson Assembly® Mix, which has three main components; 5' exonuclease, DNA polymerase enzymes and DNA ligase (Figure 2.1c). The exonuclease serves to chew back the 5' ends of both the backbone and insert, leaving 3' overhangs and allowing for complementary overhangs to anneal. The polymerase then fills in nucleotide gaps. Finally, the ligase creates a phosphodiester bond between the annealed fragments, removing any discontinuities and resulting in a continuous DNA strand. In order to effect the ligation, 0.02-0.5 pmol of backbone and a 2-3-fold molar excess of insert were mixed with the Gibson Assembly® Master Mix for a 20 μ l reaction as per supplier instructions as discussed in Chapter 4. The DNA in pmol was calculated using the formula

$$pmol = \frac{(weight\ in\ ng) \times 1000}{base\ pairs \times 650Da}$$

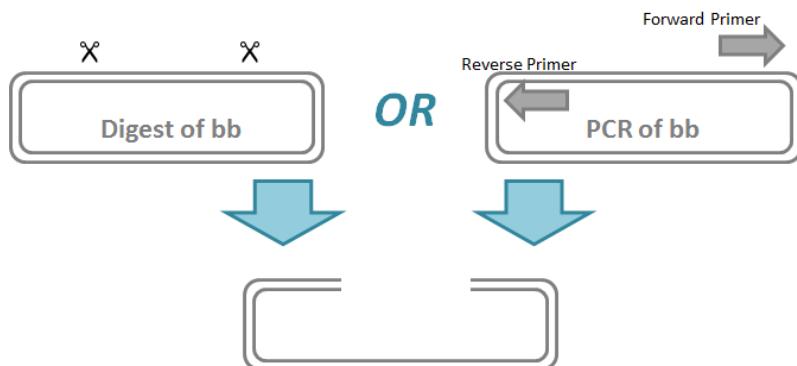
Following a brief centrifugation, the recipe was left to incubate at 50 °C in a thermocycler for 60 minutes. The samples were then either placed on ice in preparation for a transformation as per Section 2.2.7 or else stored at -20 °C.

2.2.14.2.1 Controls used in Gibson assembly®. Several ligation controls were used, an example of which is listed in Table 2.3. In all other cases a similar set-up was followed, although the volumes used differed depending on the length of the backbone, the concentration of the aliquot and the pmol used. Throughout, the same amount of pmol of backbone was used in both ligation and control reactions in order to ensure comparability of results.

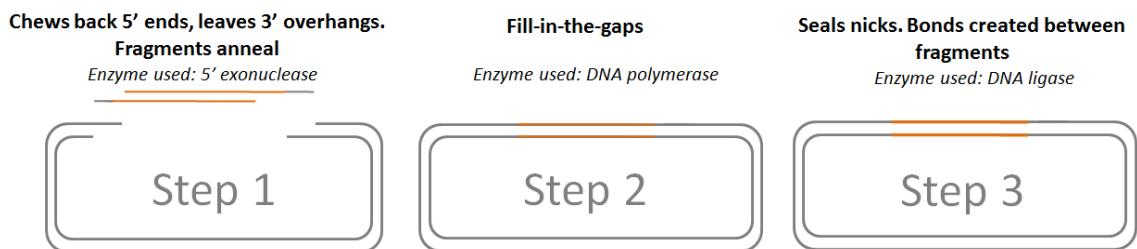
Colony counts from the transformation plates corresponding to the ligation reaction were compared to colony counts from control reactions A and B (which use purified backbone with/without the Master Mix respectively) to assess the amount of uncut plus re-ligated vector and uncut vector respectively. Colony counts for control C, which consists of uncut backbone taken through the ligation and transformation processes, were compared to the ligation reaction colony counts to



(a) The insert is prepared via PCR amplification which also serves to add overhangs which are compatible to the backbone in which the insert will be assembled.



(b) The backbone is prepared for Gibson assembly via either a preparative digest which serves to retain the required DNA backbone fragment or via PCR which serves to amplify the required DNA backbone fragment.



(c) The three main steps and enzymes used in the Gibson assembly method to ligate the prepared insert into the prepared backbone. The insert and backbone are mixed together with the Gibson Assembly mix and then consecutively go through a cleaving reaction which uses 5' exonuclease to chew back the 5' overhangs, allowing for the fragments to anneal (Step 1), followed by an extension reaction using DNA polymerase to fill in any nucleotide gaps (Step 2) and a final reaction using ligase wherein the nicks between the fragments are sealed by the formation of a phosphodiester bond, resulting in a continuous DNA fragment (Step 3).

Figure 2.1: Gibson Assembly Method. Diagram explaining the main steps involved in the preparation and Gibson ligation of two or more DNA fragments: a) Preparation of insert via PCR which adds backbone-compatible overhangs to the insert, b) Preparation of backbone via digest or PCR and c) Steps and enzymes used in the Gibson assembly method wherein the insert and backbone are mixed together with the Gibson Assembly mix and then go through a cleaving, annealing, extension and ligation reaction to assemble the final construct. 'bb' stands for backbone.

	Control A	Control B	Control C
0.02pmol amp./cut, pur. 8kb backbone at a concentration of 25ng/ μ l	4.2 μ l	4.2 μ l	-
0.02pmol uncut 10kb backbone at a concentration of 25ng/ μ l	-	-	5.2 μ l
2X Gibson Assembly Master Mix	10 μ l	-	10 μ l
dH ₂ O (to 20 μ l)	5.8 μ l	15.8 μ l	4.8 μ l

Table 2.3: **Controls used in a Gibson Assembly.** Each control is prepared using a combination of PCR-amplified/cut, purified backbone OR uncut backbone, 2X Gibson Assembly Master Mix (optional) and dH₂O to add up to 20 μ l volume. In this example, the cut backbone is taken to be 8kb long, while the uncut backbone is taken to be 10kb long. Both aliquots have a concentration of 25ng/ μ l. 0.02pmol of DNA is used in all three controls. 'amp.' is short for PCR-amplified, while 'pur.' is short for purified.

assess cell competency and viability following ligation and transformation.

2.2.15 Sequencing Plasmid DNA

As a means of plasmid identity confirmation, sequencing of plasmid DNA was carried out by the Scientific Support Services at the Wolfson Institute for Biomedical Research and the UCL Cancer Institute, UCL.

As per service requirements, 10 μ l of plasmid at a minimum of 100ng/ μ l or PCR product at a minimum of 1ng/ μ l per 100bp were submitted per reaction. Primers were submitted at a volume of 6 μ l at a concentration of 5pmol/ μ l per reaction. Results, consisting of a base sequence and an electropherogram, were given in .fsta and .ab1 format respectively. All primers used for sequencing can be found in Appendix C.

2.3 *T. brucei* strains and culturing conditions

All work concerning *T. brucei* was carried out at the London School of Hygiene and Tropical Medicine (LSHTM) at the laboratory of Dr. Sam Alsford.

2.3.1 *T. brucei* strains

The bloodstream form (BSF) *T. brucei brucei* Lister 427 MITat 1.2 (Clone 221a) (Carrington et al., 1991) strain and procyclic form (PCF) *T. brucei brucei* Lister 427 monomorphic (Cross, 1975) strain were used throughout *T. brucei* investigations. The BSF 2T1 cell line (VSG221 expressing Tagged cell line, clone 1) was used as a control strain (Alsford et al., 2005). All strains were provided by Dr. Sam Alsford at the LSHTM. Unless differentiated, the strains retain their respective form, either bloodstream or procyclic.

2.3.2 BSF *T. brucei* culturing conditions

BSF cells were grown in HMI-9 media (Hirumi & Hirumi, 1984). 4.5 litres of media was prepared using HMI-9 powder (Life Technologies Ltd., Paisley, U.K) supplemented with 15g of sodium bicarbonate (Sigma Aldrich, Munich, Germany), 50ml of Pen-Strep (Life Technologies Ltd., Paisley, U.K) and 70 μ l of β -mercaptoethanol (Sigma Aldrich, Munich, Germany). The solution was filter-sterilised and stored in 450ml volumes at 4°C. Prior to use, 50ml FBS (Sigma Aldrich, Munich, Germany) was added to each 450ml volume for a final concentration of 10% (v/v). The media was then stored at 4°C and pre-warmed to 37°C prior to use.

Cells were incubated in Nunc™ Cell Culture Treated Flasks with Filter Caps (Fisher Scientific, MA, U.S.A) in a Revco Ultima II series CO₂ incubator (Fisher Scientific, MA, U.S.A) at a temperature of 37°C and 0.5 CO₂ level. Cells were maintained at a density below 2.5 \times 10⁶ cells/ml.

2.3.3 Cryopreservation and revival of BSF *T. brucei*

In order to cryopreserve BSF *T. brucei* cells, 100 μ l of 100% glycerol was added in a cryovial to 900 μ l of culture grown to a density of over 1 \times 10⁶ cells/ml. This was then stored at -80°C for 24 hours and then transferred to liquid nitrogen for a staggered freezing profile so as not to damage cells.

In order to thaw cryopreserved BSF cells, a 1ml vial was taken from storage in liquid nitrogen, allowed to thaw and added to 5-10ml of pre-warmed media. The volume was mixed gently and incubated in 37°C in a CO₂ incubator. Following a 24 hour incubation period, in which cells are given time to recover, selective antibiotic, if appropriate, was added to cells.

2.3.4 PCF *T. brucei* culturing conditions

In order to culture procyclic cells, semi-defined SDM-79 media (Schonenberger, 1979) was used. 4.5 litres of media was prepared using SDM-79 powder (Life Technologies Ltd., Paisley, U.K), supplemented with 10g sodium bicarbonate and 50ml Pen-Strep (Life Technologies Ltd., Paisley, U.K). This was then filter-sterilised and stored at 4°C in 450ml volumes. Prior to use, 50ml of 56°C heat-inactivated FBS and 1.5ml 2.5mg/ml hemin (Sigma Aldrich, Munich, Germany) were added to each 450ml volume. The media was then stored at 4°C and pre-warmed to 28°C prior to use.

Cells were kept at a density between 1 \times 10⁶ and 2 \times 10⁷ cells/ml. PCF cells kept at lower densities do not survive. The cells did not require flasks with filter caps to be cultured in and were grown in non-ventilated flasks (Fisher Scientific, MA, U.S.A) at an incubation temperature of 28°C using an LMS Series Two Cooled Incubator (LMS Ltd., Kent, U.K).

2.3.5 Cryopreservation and revival of PCF *T. brucei*

In order to cryopreserve PCF cells, $100\mu\text{l}$ of 100% glycerol was added in a cryovial to $900\mu\text{l}$ of culture at a density of over 1×10^7 cells/ml. Storage was the same as for BSF cryopreserved cells.

In order to thaw cryopreserved PCF cells, a 1ml vial was taken from storage in liquid nitrogen, allowed to thaw and added to 10ml of pre-warmed media. This was then centrifuged for 10 minutes at 2000rpm using a table-top Eppendorf Centrifuge 5804 (Eppendorf, Stevenage, U.K) in order to remove traces of glycerol. The cells were then resuspended in 2ml media to ensure high density and incubated at 28°C . Following a 24 hour incubation period, in which cells are given time to recover, selective antibiotic, if appropriate, was added to cells.

2.3.6 Antibiotics used for the culturing of *T. brucei* strains

In order to select for BSF and PCF trypanosomes transfected with the oscillator plasmid, Hygromycin (Sigma Aldrich, Munich, Germany) was used at a concentration of $2.5\mu\text{g}/\text{ml}$. It was prepared at a stock solution of 5mg/ml in dH₂O, filter-sterilised and stored at -20°C .

2.3.7 Tetracycline induction of *T. brucei* cells transfected with the oscillator plasmid

In order to induce *T. brucei* cells transfected with the oscillator plasmid with tetracycline (Sigma Aldrich, Munich, Germany), a 1mg/ml tetracycline stock solution was prepared by dissolving 10mg of tetracycline in 10ml of dH₂O, filter-sterilising the solution and storing at -20°C . The antibiotic was used at a working concentration of 10pg/ml-1 $\mu\text{g}/\text{ml}$, depending on the study as discussed in Chapter 7. Unless otherwise specified, the culture was then incubated in the appropriate conditions for 24 hours prior to harvesting or analysis, in order to allow for the tetracycline to become effective.

2.3.8 Blocking of *T. brucei* protein synthesis using cycloheximide

In order to determine the half-life of proteins within the oscillator plasmid, cycloheximide was used as per Leung et al. (2011) in order to block translation processes. Cycloheximide (Calbiochem, Nottinghamshire, U.K) was prepared as a 50mg/ml stock solution in ethanol, to enable solubility, and stored for up to three months at -20°C .

Two induced BSF pUnstable clone cultures were grown to 1×10^6 cells/ml. Cycloheximide was added to a concentration of $100\mu\text{g}/\text{ml}$ to one of the cultures. An equal volume of ethanol was added to the other culture to act as a control. Both flasks were then swirled gently and incubated. For both cultures, a protein sample was taken and protein prepared (Section 2.5.1) at the required time-points post-incubation. Following this, protein was analysed via a western blot (Section 2.5.3) and coomassie stain (Section 2.5.4).

2.4 Recombinant DNA techniques in *T. brucei*

2.4.1 Stable DNA transfection of BSF *T. brucei*

Prior to transfection, 10 μ g of insert DNA was linearised using the NotI (NEB, MA, U.S.A) enzyme as per Section 2.2.10.2 and purified using phenol-chloroform techniques as per Section 2.2.9.

2 \times 10⁷ BSF *T. brucei* cells were spun down at 2000rpm for 10 minutes (Alsfeld & Horn, 2012). The supernatant was then removed, cells were resuspended in 0.5ml PBS (Sigma-Aldrich, Munich, Germany) and centrifuged for one minute at 10,000rpm. The supernatant was again removed and replaced with 10 μ l of the phenol-chloroform purified, digested DNA and 100 μ l of Cytomix. Cytomix was prepared as per the following recipe: 120mM KCl, 25mM HEPES at pH7.6, 0.5% glucose, 100 μ g/ml Albumin bovine serum, 150 μ M CaCl₂, 2mM EGTA at pH7.6, 1mM Hypoxanthine, 5mM MgCl₂.6H₂O and 10mM K₂HPO₄/KH₂PO₄ at pH7.6. Note that 100 X Hypoxanthine was prepared using 0.4g NaOH, 100ml H₂O and 1.36g Hypoxanthine and stored at -20 °C. 10X K₂HPO₄/KH₂PO₄ at pH7.6 was prepared using 8.66ml 1 M K₂HPO₄, 1.34 ml 1M KH₂PO₄ and 90ml H₂O and stored at room temperature. All chemicals were supplied by Sigma Aldrich (Munich, Germany). The pH was adjusted to 7.6, the Cytomix filter sterilised and then stored at 4 °C.

The cell, DNA and Cytomix mixture was then transferred to an 0.2cm cuvette and the transfection was carried out using the Amaxa™ Nucleofector™ II (Lonza Group Ltd., Basel, Switzerland) transfection device, using nucleofection programme X-001. Following transfection, 1ml of pre-warmed media was added to the cuvette and this was transferred into 60ml pre-warmed media. The flask was then incubated for six hours to allow for expression of Hygromycin resistant protein, after which a further 140ml pre-warmed media was added to dilute cells and Hygromycin selective antibiotic was applied. 1ml volumes were then aliquoted into two 48-well plates at a neat and 5-fold dilution or into four plates using a neat dilution. The plates were incubated for five days, after which positive clones were identified, transferred to 10ml pre-warmed media with Hygromycin and incubated for 24 hours prior to analysis.

2.4.2 Stable DNA transfection of PCF *T. brucei*

Transfection of PCF cells follows the same protocol except for the required use of conditioned media, which is media in which PCF *T. brucei* have been previously cultured (Archer, 2009b). Conditioned medium was freshly prepared prior to each transfection carried out. 200ml of PCF cells were grown to a density of 1 \times 10⁷cells/ml. This volume was then divided into 50ml aliquots and centrifuged at 2000rpm for 10 minutes to pellet and separate the cells from the medium, excreted proteins and chemicals. After this, the supernatant was retained, transferred to a new tube and the centrifugation repeated. Again, the supernatant was retained, after which it was passed through an 0.2 μ m filter (Acrodisc® Syringe Filters with 0.2 μ m Supor® Membrane, Pall Life Science, Portsmouth, U.K) and stored at 4 °C prior to the transfection process.

Following this, the same transfection process as that used for BSF cells was carried out (Section 2.4.1), with one main difference. Following transfection, the use of pre-warmed media for BSF cells was replaced with pre-warmed conditioned media supplemented with 10% (v/v) 56 °C heat-inactivated FBS. This allows for PCF cell growth even though cell density is below 10⁶cells/ml,

which usually results in cell death (Archer, 2009a).

2.4.3 *T. brucei* immobilisation using chemical agents and microfluidic devices

2.4.3.1 Microfluidic laser cage devices

This work was carried in collaboration with Dr. Brian O'Sullivan at UCL. For the development of microfluidic cage devices, the devices were initially designed in Adobe Illustrator software (v. CS6)(Adobe Systems Inc., CA, U.S.A), after which they were imported into WinMark Pro Laser Marking Software (v. 6.0)(Synrad Inc., WA, U.S.A)

In order to manufacture the devices, slides measuring 24×60mm were cut from 1.5mm thick polymethylmethacrylate (PMMA; RS Components, Northants, U.K) using a CO₂ Synrad 48-Series laser marking head (Laserlines, Banbury, U.K) with a maximum power of 25W at 50% power (12.5W) and a mark speed (laser tracking speed) of 10mm/s. Following this, square chambers with a width of 50-150 μ m were ablated on the slide using a power of 12.5W and a mark speed of 200-300mm/s.

2.4.3.2 Agarose block trap devices

Live cells were immobilised by placing an agarose block on them (Figure 2.2) (Huang et al., 2014). PBSG was prepared by supplementing PBS with 6mM D-glucose. Following this, agarose was added to 20ml of PBSG for a final concentration of 1% (w/v) and dissolved by heating. After allowing to cool to a tepid temperature, 12ml was poured into a 9.2cm diameter petri dish to a thickness of 2cm and allowed to solidify for at least 30 minutes. 1cm-by-1cm agarose blocks were then cut (but left in the petri dish, to avoid drying out). If necessary, the dish was sealed and stored at 4 °C. Prior to use, the agarose blocks were left to set at room temperature.

Prior to cell preparation and analysis, a glass slide and cover slip set-up was prepared by cleaning the microscope glass slide (Fisher Scientific, MA, U.S.A) with 70% ethanol and lining a 24mm×60mm cover slip (Fisher Scientific, MA, U.S.A) with blu tack (Figure 2.2). 1ml of cells at a density of 1×10⁶cells/ml were then centrifuged at 1000rpm for one minute. The supernatant was discarded and the cells were resuspended with an equal volume of pre-warmed PBSG. The centrifugation process was repeated, the supernatant discarded and the cells resuspended in 50 μ l pre-warmed media.

Immediately after this, 5 μ l of cells was pipetted onto the glass slide, covered gently with a 1cm-by-1cm agarose block and the cover slip which is kept in place by the blu tack which attaches to the glass slide. The set-up was then inverted, so that the cells are on top the agarose block.

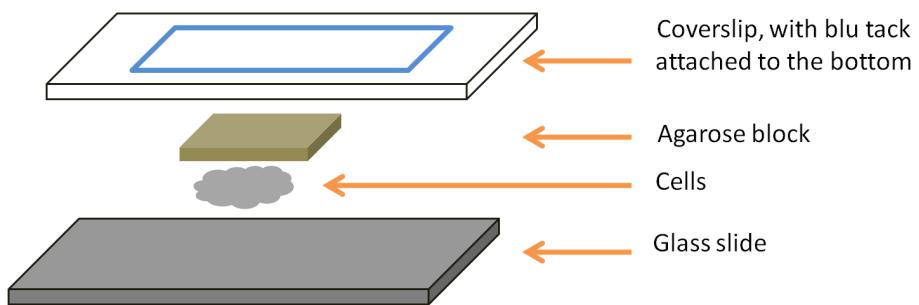


Figure 2.2: **Agarose block trap device:** Main components of the agarose block trap device used to immobilise viable and motile *T. brucei* cells. The device consists of a glass slide, 1cm-by-1cm 1% agarose block and cover slip with blu tack attached to the bottom. 5 μ l of cells are pipetted onto the glass slide, on top of which an agarose block and the cover slip with blu tack are placed. The set-up is then inverted.

2.4.3.3 Agar and agarose solutions

Agar and agarose were both prepared by adding the required % (w/v) of powder to HMI-9 media, dissolving the powder by mixing and then autoclaving the solutions. Prior to use, both solutions were heated in a microwave and if required, aliquots were diluted with HMI-9 media.

2.4.3.4 Xanthan gum

Xanthan gum solutions were prepared by adding the required % (w/v) of xanthan powder (Sigma-Aldrich, Munich, Germany) to HMI-9 media and mixing the solution at a heated temperature of c. 30 °C using the Bibby Scientific HB502 heated magnetic stirrer (Bibby Scientific Ltd., Staffordshire, U.K). Prior to use, the solution was mixed to ensure homogeneity. If required, the solution was diluted using HMI-9 media.

2.4.3.5 Cygel™ thermoreversible gel

Cygel™ (BioStatus Ltd., Leicestershire, U.K) is a thermoreversible product which is liquid below room temperature and solid at temperatures above. It was stored at 4 °C, as per supplier instructions. If required, Cygel™ was diluted using dH₂O prior to use. 40X PBS supplied with the gel was then added to a final concentration of 2.5% (v/v).

2.5 *T. brucei* protein analysis

2.5.1 *T. brucei* protein isolation

To prepare protein from *T. brucei* cells, 10ml of 1 \times 10⁶ cells/ml culture was pelleted at 2000rpm for 10 minutes. The supernatant was removed, the pellet resuspended in 10ml PBS and the centrifugation repeated. Following this, the supernatant was removed and the pellet was resuspended in 500 μ l PBS. This was followed by a two minute centrifugation at 5000rpm, after which the pellet was resuspended in 1X SDS sample buffer to a final concentration of 1 \times 10⁵ cells/ μ l.

The SDS sample buffer was prepared by dissolving 62mM Tris at a pH6.8, 10% glycerol, 2.3% SDS (Sigma Aldrich, Munich, Germany), 5% β -mercaptoethanol and 0.002% bromophenol blue (Sigma Aldrich, Munich, Germany) in rH₂O. A SIGMAFAST™ Protease Inhibitor tablet (Sigma Aldrich, Munich, Germany) was added and dissolved via vortexing, after which the sample was stored at -20°C.

Following resuspension in the SDS sample buffer, the cell solution was vortexed for two seconds in order to ensure homogeneous breakdown of the cell wall and inhibition of protease activity. The solution was stored at -20°C.

2.5.2 SDS-PAGE gel analysis of *T. brucei* proteins

2.5.2.1 Preparation of resolving and stacking gels

Following protein isolation (Section 2.5.1), an SDS-PAGE gel (Laemmli, 1970) was run in order to separate proteins. The Bio Rad Mini-PROTEAN® Tetra System (Bio-Rad, CA, U.S.A) set-up was used as per kit instructions to prepare and run gels.

Once the casting frame was assembled using a glass plate and 0.75mm spacer plate, the 12.5% resolving gel was prepared using 2.08ml 30% Acrylamide (Sigma Aldrich, Munich, Germany), 1.38ml Lower Gel Buffer (1.5M Tris at pH8.8 and 0.4% SDS, 1.75ml dH₂O, 500 μ l 10% Ammonium Persulfate (APS) (Sigma Aldrich, Munich, Germany) and 2 μ l N,N,N',N'-Tetramethylethylenediamine (Temed) (Sigma Aldrich, Munich, Germany). The Temed and APS were added in the end to prevent premature gel polymerisation. 3.5ml of this solution was set to solidify in the casting frame and on top of it, 0.5ml dH₂O was pipetted to help remove bubbles along the gel's surface. The gel was allowed to polymerise for one hour, after which the top layer of water was removed by tipping over the casting frame and using blue paper towel to absorb water.

The 5% stacker gel was prepared using 250 μ l 30% Acrylamide, 375 μ l Upper Gel Buffer (0.5M Tris at pH6.8 and 0.4% SDS), 875 μ l dH₂O, 12.5 μ l 10% APS and 1.25 μ l Temed. The Temed and 10% APS were added in the end to prevent premature polymerisation of the gel. After the addition of the stacker gel to the casting frame on top of the resolving gel, a 10 well comb was inserted on top to form wells. This was followed by a further one hour polymerisation period.

2.5.2.2 Running the SDS-PAGE gel

Prior to running the gel, 1X SDS-PAGE running buffer was prepared. A one litre 10X running buffer solution was prepared using 144g glycine (Sigma Aldrich, Munich, Germany), 1% SDS and 30g Tris. When running one or two gels 700ml of 1X running buffer was used, while when running three or four gels one litre was used. The gels were then slotted into place in the tank, the inner buffer chamber filled with 1X running buffer, the combs removed and the wells cleaned using a 1ml pipette. The samples and ladder were then thawed on ice and 10 μ l of each sample was loaded into the wells. The PageRuler Prestained Protein Ladder (Fisher Scientific, MA, U.S.A) or the Precision Plus Protein™ WesternC™ Standard (Bio-Rad, CA, U.S.A) ladder was loaded into one

of the wells. After this, the outer buffer chamber was filled with the remainder of the 1X running buffer, the lid placed and connected to the mains and a voltage of 200V was run for approximately one hour through the gels using an Electrophoresis Power Supply (Fisher Scientific, MA, U.S.A). The tank was placed in a tray filled with ice to avoid over-heating due to the high voltage being used.

When the run was completed, the voltage was stopped and the mains switched off. The gel was then removed from the glass plates holding it in place, and immediately placed into the appropriate solution, depending on whether the gel was to be taken forward to a western blot (Section 2.5.3) or a Coomassie stain (Section 2.5.4).

2.5.3 Western blotting

2.5.3.1 Semi-dry protein transfer

In order to detect proteins of interest, a western blot was carried out following an SDS-PAGE run using the semi-dry transfer method. The Trans-Blot® Turbo™ Blotting System (Bio-Rad, CA, U.S.A) was used to transfer protein from the gel to a nitrocellulose membrane. Assembly instructions were followed throughout.

One litre of transfer buffer was prepared using 5.83g Tris, 2.93g glycine, 200ml methanol (VWR International Ltd., Leicestershire, U.K) and 0.375g SDS. Following an SDS-PAGE run, the gel was placed in Western Transfer buffer for 15 minutes while being rotated at 20rpm using the Stuart Mini See-Saw Rocker SSM4 (Bibby Scientific Ltd., Staffordshire, U.K).

During this equilibration period, the materials to be used for the semi-dry transfer sandwich were assembled. This consisted of wetting two extra thick filter papers (Extra Thick Blot Paper; Bio Rad, CA. U.S.A) in transfer buffer for at least 10 minutes, with no agitation. In addition, a nitrocellulose membrane (Amersham Hybond ECL Nitrocellulose Membrane; GE Healthcare Life Sciences, Buckinghamshire, U.K) was placed in transfer buffer for 30 seconds, transferred to rH₂O for a further one minute and then placed again into transfer buffer for 10 minutes, this time with 20rpm agitation.

Once the membranes and filter paper were equilibrated, the transfer sandwich was prepared by placing an extra thick filter paper at the bottom, on top of which the nitrocellulose membrane was placed, then the SDS-PAGE gel and finally the second extra thick filter paper. At each addition, any bubbles and excess liquid were removed using a roller. The transfer sandwich was assembled on the base anode cassette, with the cathode cassette placed on top, locked into place and slotted into the apparatus bay, as per manual instructions. The Bio Rad pre-defined Standard SD program was used, where a voltage of 25V and a current of 1A were run through the transfer sandwich for 30 minutes.

2.5.3.2 Blocking of non-specific binding at the membrane

Following transfer, the membrane was placed in blocking solution for a minimum of 30 minutes in agitation in order to prevent non-specific binding. The blocking solution was prepared using 5% milk powder (Marvel, Premier International Food (U.K) Ltd., Lincolnshire, U.K) and 0.05% TWEEN®-20 (Tween) (Sigma Aldrich, Munich, Germany) in TBS or PBS buffer, depending on the antibody being used. One litre of TBS was prepared using 50ml of 1M Tris at pH8.0, 36.4ml 4M NaCl (Fisher Scientific, MA, U.S.A) and topped up with rH₂O.

2.5.3.3 Addition of antibodies for protein detection in western blots

After blocking, the membrane was placed in the primary antibody solution for one hour in 20rpm agitation. The primary antibody was prepared in the blocking solution, and diluted as per antibody supplier instructions. For the detection of the eGFP gene, the GFP Rabbit IgG Polyclonal Antibody Fraction (Life Technologies Ltd., Paisley, U.K) was used at a dilution of 1:5000. PBS was used as a buffer in blocking and antibody solutions. For the detection of the tetR gene, the Anti-Tet Repressor Antibody (EMD Millipore Co., MA, U.S.A) and the Anti-Tet Repressor Antibody (Sigma Aldrich, Munich, Germany), both produced in rabbit, were used at a dilution of 1:1000. TBS was used as a buffer. For the detection of the Myc tag protein, the Mouse IgG1 Monoclonal Antibody to c-Myc (Source BioScience, Nottingham U.K) was used at a dilution of 1:2000. TBS was used as a buffer.

Following incubation, a rinsing process was carried out to remove excess antibody. This consisted of placing the membrane in the same buffer as the one used to prepare blocking and antibody solution but supplemented with 0.05% Tween and agitated for 10 minutes at 20rpm. The solution was discarded and the rinse process was repeated.

After discarding the solution, the membrane was placed in a TBS-T (TBS + 0.05% Tween) solution for a final 5-minute rinse, agitated at a speed of 20rpm. The membrane was then placed in TBS-T plus 0.5% milk powder containing the secondary antibody diluted as per supplier instructions. Following use of the anti-GFP and anti-TetR antibodies, the Goat Anti-Rabbit IgG (H+L)-HRP Conjugate (Bio-Rad, CA, U.S.A) was used at a dilution of 1:12000. Following the use of the c-Myc antibody, which had a Mouse host, the Goat anti-mouse IgG(H+L) Peroxidase Conjugated Antibody (Fisher Scientific, MA, U.S.A) was used at a dilution of 1:12000.

This was then incubated at 20rpm for a further one hour. Following this, the membrane was rinsed in TBS-T for five minutes and agitated at 20rpm, after which the solution was discarded. The rinse process repeated for another two times.

Following the final rinsing process, the membrane was prepared for visualisation and analysis using the Amersham™ ECL™ Western Blotting Analysis system kit (GE Healthcare Life Sciences, Buckinghamshire, U.K) which is used to detect tagged proteins. The detection reagent preparation and applications instructions plus post-handling of the membrane were carried out as per manual instructions. Following preparation, the set-up was transferred to a dark room and an autoradiography film (Amersham™ Hyperfilm ECL, GE Healthcare Life Sciences, Buckinghamshire, U.K) was pressed onto the membrane for an exposure of 15-60 seconds. The Compact X4 (Xograph Healthcare, Gloucestershire, U.K) X-Ray film processor was used to visualise and print

results.

2.5.4 Coomassie stain for protein visualisation following SDS-PAGE

Following an SDS-PAGE (Section 2.5.2), a Coomassie stain was carried out in order to obtain a protein profile of the sample being analysed. The gel was placed into Isopropanol Fixing Agent solution and rotated at 20rpm for 15 minutes. The Fixing Agent solution was prepared using 25% (v/v) isopropanol (VWR International Ltd., Leicestershire, U.K), 10% (v/v) acetic acid (Fisher Scientific, MA, U.S.A) and 65% rH₂O. This was followed by a one hour staining procedure using a Coomassie G-250 stain (Sigma Aldrich, Munich, Germany). The staining solution was prepared using 10% (v/v) acetic acid, 0.006% (w/v) Coomassie brilliant blue G-250 and 90% rH₂O. Again, the gel was rotated at 20rpm to ensure homogeneous application of the stain. Finally, the gel was placed in a destaining solution consisting of 10% acetic acid and rotated at 20rpm for at least 30 minutes.

In order to image the stained gel, the GeneGenius Bio-Imaging System (Syngene, Cambridge, U.K) apparatus was used along with the image acquisition software GeneSnap (v. 6.08, Syngene, Cambridge, U.K.) which was used to control exposure and light conditions.

2.5.5 Densitometry plot

Densitometry graphs based on western blot signal intensities were plotted using ImageJ v1.48 (Schneider et al., 2012). Given a ‘.jpg’ image of the western blot output, the density profile of each band in the blot was calculated via ImageJ. After controlling for background noise manually, the area under each density profile was calculated again using ImageJ. Following this, the relative density readings x_i were extracted for each band i by using either the data point of a positive control or the data point of the band intensity at time zero x_r as a reference point and calculating x_i/x_r .

2.5.6 Immunofluorescence assay

2.5.6.1 Cell fixation process

In order to carry out an immunofluorescence assay to detect fluorescence levels in fixed cells, cells were grown to a density of 1×10^6 cells/ml. 0.5ml 4% (v/v) paraformaldehyde (Sigma-Aldrich, Munich, Germany) was added to 0.5ml of culture. The mixture was inverted gently and placed on ice for 15 minutes. It was then centrifuged for one minute at 5000rpm, 950 μ l of the supernatant removed and the cells resuspended in the remaining solution. 1ml pre-chilled PBS was added and the process of centrifuging, discarding the supernatant and resuspending in 1ml PBS was repeated twice, although in the resuspension step of the last run, 0.5ml of pre-chilled PBS was added instead of 1ml. A final centrifugation step was carried out, and the supernatant removed except for a residual 20 μ l, in which the cells were resuspended. 2 μ l of cells was pipetted onto a 5mm 12-well slide (Fisher Scientific, MA, U.S.A), and the density of the cells was checked using a 25X or 40X objective on a SP-95-1 Inverted Microscope (Brunel Microscopes Ltd., Wiltshire,

U.K) having 10X widefield eyepieces for a total magnification of 250X or 400X. If too sparse, an additional $2\mu\text{l}$ of cells was pipetted onto the cells.

2.5.6.2 Addition of antibodies for protein detection in immunofluorescence assays

If antibody was to be added in order to detect specific proteins, following the air-drying process, instead of staining the cells, $40\mu\text{l}$ of 50% (v/v) FBS in PBS was added and left to act as a blocking solution for 10 minutes. The blocking mixture was removed by softly dabbing a paper towel on the well. $40\mu\text{l}$ of primary antibody diluted 1:200 in 3% (v/v) FBS in PBS were then pipetted onto the well and left to incubate for 45 minutes in a humidity box. The excess solution was then removed by tipping over the slide, which was thrice washed in 50ml PBS for five minutes per wash (with no agitation). $40\mu\text{l}$ of the secondary antibody at a dilution of 1:50 in 3% (v/v) FBS in PBS was then pipetted onto the well and the slide was left to incubate for 45 minutes in a humidity box. The excess solution was removed by tipping the slide, which was then thrice washed in 50ml PBS for five minutes per wash. The slide was then allowed to air dry for at least 30 minutes in the hood. Following this, $1\mu\text{l}$ of DAPI in Vectashield (Vector Labs, Peterborough, U.K) was pipetted onto the cells, a cover slip was placed on top and the edges were sealed with nail varnish.

2.5.6.3 Visualisation

Images of fixed and stained cells were captured using the AxioPlan2 microscope set up with the QImaging® Retiga-2000R Fast 1394 digital camera. Prior to imaging, the surface of the cover slide over each sample was covered with a drop of immersion oil (Sigma Aldrich, Munich, Germany). In order to acquire both brightfield and fluorescent images of the cells, the Velocity® 3D Image Analysis Software, Version 6.3.0 (Perkin Elmer Inc., MA, U.S.A) was used.

2.6 *T. brucei* live cell imaging via fluorescent and brightfield microscopy

2.6.1 DEFRA licensing for *T. brucei* live cell imaging

In order to be able to carry out authorised work with *T. brucei* organisms and carry out live-cell imaging, a Specified Animal Pathogens Order 2008 (SAPO) License, number PATH/151/2010/1 was issued by the Department for Environment, Food and Rural Affairs (DEFRA).

2.6.2 Imaging set-up using live *T. brucei* cells

In order to be able to image viable and motile *T. brucei* cells via fluorescent and brightfield microscopy, cells were embedded in 40% Cygel™ in order to restrict motility. Cells were initially grown to over 1×10^6 cells/ml. An 0.5ml aliquot was taken and centrifuged for one minute at 5000rpm, after which the supernatant was removed and the cells were resuspended in media

Channel	Brightfield	Fluorescent
Exposure	2ms	2.0s
Colour Format	Monochrome	Monochrome
Zeiss Fluo Turret	Filter 1	Screen
Zeiss TL Voltage	5.4V	0.0V
Gain	1.8x	0.0x
Offset	46 levels	0 levels

Table 2.4: Camera settings used in brightfield and fluorescent microscopy to acquire consecutive images of viable cells.

to a concentration of 5×10^7 cells/ml. $5\mu\text{l}$ of this solution was then taken and mixed with an equal volume of 40% Cygel™. $4\mu\text{l}$ of this mixture was pipetted onto a 5mm 12-well slide, on top of which a cover slide was placed and sealed with nail varnish in order to comply with Health and Safety regulations.

2.6.3 Imaging conditions

The AxioPlan2 microscope, set up with the QImaging® Retiga-2000R Fast 1394 digital camera was used to acquire consecutive fluorescent and brightfield images over time. Prior to imaging, the surface of the cover slide over each sample was covered with a drop of immersion oil. In order to acquire both brightfield and fluorescent images of the cells, the Volocity® 3D Image Analysis Software, v. 6.3.0 was used in order to set up automated image capturing protocols. Camera settings such as exposure times and colour format were set up as per Table 2.4. Images were captured every two minutes for a duration of two to four hours.

2.6.4 Background correction

Following image capturing, images were analysed using the Volocity 3D Image Analysis. Background correction was carried out on fluorescent microscopy images in order to better distinguish between fluorescent cells and background noise. This was done by using a blank well as the control. A cover slip was put on top of the blank, a drop of immersion oil was placed on the cover and a fluorescent image acquired using the same settings as for the other wells. This blank image was then selected as the Dark Reference Image when setting up background correction to manipulate the images.

2.6.5 Quantification of fluorescence intensity profiles

Following image acquisition and background correction, the next step was to extract fluorescent trajectories at the single cell level. In order to quantify fluorescence intensities of individual cells, Volocity Quantitation was used to detect and measure fluorescent spots at each time-point via the Measurement view. Objects were excluded by size, with only fluorescent spots of size 20–10,000 μm^2 being measured.

Following this, results were confirmed visually by tracking a cell of interest manually over each image acquired to ensure that the software had measured the correct object in each case.

2.7 Mathematical Modelling of the Goodwin oscillator GRN in a *T. brucei* host

2.7.1 The Gillespie Direct Method stochastic simulation algorithm

The Gillespie Direct Method stochastic simulation algorithm (SSA) as developed by Gillespie (1977) was used to generate stochastic time series simulations of the quantity of molecular components/species within the oscillator GRN over a span of 500-1000 minutes. The algorithm works by selecting an initial time-point t_0 and simulating when a biochemical reaction next takes places and also which reaction takes place. Species' molecule numbers and timings are then updated accordingly and the process is repeated. The decision of when the next reaction takes place and also which type of reaction takes places is based on the generation of two random numbers from the uniform distribution with parameters 0 and 1. This generates a trajectory of the molecular components. The algorithm scheme runs as follows:

Step 1) Given thermal and spatial homogeneity

- Set t_0 and t_{final} . Set $t = t_0$.
- Store initial molecule numbers for the N species, $X_i(t_0)$ for $i \in 1, \dots, N$
- Store reaction rate values of the M reactions, c_j for $j \in 1, \dots, M$

Step 2) Calculate the propensity values $a_j(t) = h_j \cdot c_j$ for $j \in 1, \dots, M$ where h_j is the number of available molecules of the reactant species in reaction j

Step 3) Calculate $a_0 = \sum_{j=1}^M a_j$

Step 4) Generate a random pair (r_1, r_2) from the standard uniform distribution $U(0, 1)$

Step 5) Calculate $\tau = \frac{1}{a_0} \cdot \ln \frac{1}{r_1}$

Step 6) Calculate k such that $\sum_{j=1}^{k-1} a_j < r_2 \cdot a_0 \leq \sum_{j=1}^k a_j$

Step 7) Update $t = t + \tau$ and $X_i(t)$ for $i \in 1, \dots, N$ to reflect the changes in the population from the execution of reaction k .

Step 8) Go to step 2 and repeat until $t > t_{final}$

The Gillespie algorithm was implemented via the ABC-Sysbio algorithm, as discussed in Section 2.7.1.2.

2.7.1.1 Copasi 4.12.65 for coding the oscillator GRN in SBML format

To carry out time series simulations, the network was represented as a set of molecular species and biochemical reactions. This information was coded in Copasi 4.12.65 (Hoops et al., 2006)

which was used for debugging the mathematical model and for translating the data into Systems Biology Markup Language (SBML) (Keating & Le Novere, 2013)

2.7.1.2 ABC-Sysbio 2.05 for running the Gillespie stochastic simulation algorithm

In order to simulate dynamics of the GRN components using the Gillespie SSA, the ABC-Sysbio 2.05 (Liepe et al., 2010) Python-based program was used as per developers' instructions. In order to run the program, the Ubuntu 12.04 (www.ubuntu.com) Linux platform was used to access a Dell PowerEdge C6100 server plus C410X GPU chassis. The system uses the centOS 5.8 operating system using a Tesla M2050, M2090 and K20 GPU. This remote system was used with the permission of Dr. Chris Barnes at the UCL Research Department of Cell and Developmental Biology.

Two scripts; *abc-sysbio-sbml-sum* and *run-abc-sysbio*, were used. The former was used by running the command *abc-sysbio-sbml-sum -files file.xml* in a terminal window, where *file.xml* is the SBML model of the Goodwin oscillator. This command serves to parse and translate the GRN model's SBML file into an ABC-Sysbio compatible input '.xml' file (*input_file_template.xml*). Additionally, it also inserts default settings in the file, to be used for the simulation of the species' quantity profiles.

The file was then manually edited in order to specify the following algorithm and simulation conditions:

1. **<particles>** This represents the number of time series simulations to carry out.
2. **<data> <times>** The time points at which to print molecule numbers of the species which have been simulated.
3. **<models> <model1> <type>** The type of simulations to carry out. This was consistently set to Gillespie.
4. **<models> <model1> <parameters>** The distribution of the range of values which can be used for each parameter during simulations. The parameter values are randomly picked from these distributions at the beginning of each simulation. No simulation is carried out using the same set of parameters.
5. **<models> <model1> <initial>** The initial number of molecules for each of the species within the model.

In order to run Gillespie stochastic simulations, the *run-abc-sysbio* script is run as per the following command: *run-abc-sysbio -infile input_file_template.xml -f -of=res -timing -cuda -simulate* where *input_file_template.xml* is the file generated by *abc-sysbio-sbml-sum* and then edited manually, **-f** asks the program to indicate when each iteration has finished **-of** specifies the name of the folder in which to place simulation results (in this case, *res*), **-timing** asks the program to print timing information, **-cuda** asks the program to use CUDA platform implementation and **-simulate** asks the program to generate time series for the specified model. This script generates two text files. *particles.txt* contains information about the parameter values which were selected from the user defined range of values and used in each simulation. *trajectories.txt* lists the quantities of the

different species at each of the specified time points in `(data) (times)`.

2.7.2 MATLAB R2012a for analysis of Gillespie simulations

Data generated by ABC-Sysbio was analysed using MATLAB (2012) R2012a. The software was also used to plot time series, matrix plots and reconstructed phase space trajectories during quantitative and qualitative mathematical analysis.

2.7.3 Data smoothing by application of imputation or a moving average filter

In order to impute or smooth data during the qualitative analysis process, a moving average filter was applied to the data. Given a set of N data points x_1, \dots, x_N , the moving average of data point k , y_k , using window size n is based on the value of data point k , the previous n raw data points x_i and the subsequent n raw data points x_i , as per the following:

$$y_k = \frac{\sum_{i=k-n}^{k+n} x_i}{2k + 1}$$

given that $k - n > 0$ and $k + n \leq N$. Else, each data point with the index out of bounds is ignored, and the denominator subtracted by 1. This gives a set of N filtered data points y_1, \dots, y_N . By weighting each data point equally, the dominant trend is preserved while eliminating noise signals. This technique was implemented in MATLAB using a custom written algorithm, as seen in Appendix D.

2.7.4 Cubic spline data interpolation

In order to fit a (smooth) curve to a set of data points, a spline was fitted. Given a dataset y with values defined at independent points t_1, \dots, t_N , a cubic spline serves to generate a stable piecewise function, $f : [a, b] \rightarrow \mathbb{R}$ where $[a, b] = [a = t_1 < t_2 < \dots < t_{N-1} < t_N = b]$, consisting of cubic polynomials between each consecutive pair of data points. The polynomials are continuous and continuously differentiable up to the second derivative, even at the interior boundary points (De Boor, 2001).

A cubic spline was fitted to the dataset of interest using the MATLAB function `yy = spline(t, y)` wherein yy is the output representing the coefficients of the spline's polynomials, t is the data array of independent variables taken to be time and y is the set of data points to which the spline must be fitted. Following this, `v = ppval(pp, yy)` was used to generate values of the piecewise polynomial function yy at time-points pp . The `plot` function was used to plot the fitted spline.

2.7.5 Fitting of linear growth trends

In order to fit a linear trend $y = m \cdot t + c$ to a dataset, the *polyfit* function in MATLAB was used. The function works by running the command $yy = \text{polyfit}(t, x, d)$ which fits a polynomial of degree d to the vector x using the vector t as an independent variable, generally taken to be time. In this case $d = 1$. The output yy generates a vector of coefficients representing m and c from the linear trend.

In order to generate data points for the linear trend, $y = \text{polyval}(pp, yy)$ was used wherein the coefficients of the polynomial in yy are used to generate a vector of values y at the vector of time-points pp .

2.7.6 Phase space reconstruction

In order to reconstruct the phase space, a custom written program in MATLAB was used. Given a delay of k , the program used the *plot*(x, y) function to map the vector x against the vector y where y equals the vector x offset by k data points. Let N be the number of data points and $x(i, 1)$ for $i = 1, \dots, N$ be equal to the data points on which the reconstruction will be based, i.e. the variable output. Then the function $\text{plot}(x(1 : (N - k), 1), y(1 : (N - k), 1))$ is used where $y(i, 1) = x(i + k, 1)$ for $i = 1, \dots, N - k$. The program is reproduced in Appendix E.

Chapter 3

Establishing *Roseobacter* marine bacteria as a synthetic biology chassis

In order to establish *Roseobacter* marine bacteria as a synthetic biology chassis by identifying reliable molecular biology procedures which enable the construction of synthetic GRNs, three strains from the clade were used throughout the investigation: *Roseobacter denitrificans* (*R. denitrificans*), *Oceanobulbus indolifex* (*O. indolifex*) and *Dinoroseobacter shibae* (*D. shibae*). The strains were selected firstly because they are dispersed along the *Roseobacter* phylogenetic tree, thus making it possible to check whether the procedures of interest can be applied across the whole *Roseobacter* spectrum. Secondly, the strains' genomes have been sequenced (Brinkhoff et al., 2008) and were used by Piekarski et al. (2009) to develop a fluorescence reporter system, thus providing a good foundation to build on. The strains were also put forward as a potential synthetic biology chassis by the UCL team, Plastic Republic, participating in the 2012 iGEM competition, in order to tackle oceanic bioremediation challenges (http://2012.igem.org/Team:University_College_London).

3.1 Long-term viability of *Roseobacter* glycerol stocks

In order to establish *Roseobacter* strains as synthetic biology chassis, the first step was to test the period of time over which glycerol stocks of the strains *R. denitrificans* and *O. indolifex* stored at -80 °C could be revived in static and shaken cultures. This was carried out in order to assess whether it is possible to maintain long-term storage of cells over time and build master and working cell banks.

3.1.1 Preparation of glycerol stocks

The first step was to prepare glycerol stocks and store them directly at -80 °C. For each *Roseobacter* strain of interest, 20ml of MB was inoculated with 100 μ l from a fresh glycerol stock. From this solution, two 10ml aliquots were taken and shaken for 12-16 hours at 200rpm at 37 °C. Following this, for each of the 10ml cultures:

- A 6ml aliquot was taken and added to 1.6ml of sterile 80% glycerol. The volume was divided into 380 μ l aliquots and stored at -80 °C. These glycerol stocks were then assessed in subsequent weeks.
- A 50 μ l aliquot was taken and diluted 10,000-fold in MB. This was done by diluting the 50 μ l aliquot in 50ml MB broth. From this, a 5ml aliquot was taken and added to 45ml MB. Following this, 50 μ l was spread onto a non-selective MB agar plate and incubated for 12-16 hours at 37 °C, after which the first colony count readings were taken and the CFU/ml calculated, taking into account the dilutions made. This was taken to be the initial reading for the growth of revived *Roseobacter* strains within static cultures.
- A 1ml aliquot was taken and the optical density (OD) at 750nm was measured using the Ultraspec 500 pro (GE Healthcare Life Sciences, Buckinghamshire, U.K) machine as an initial reading of the OD's of shaken revived cultures.

3.1.2 Measurement of long-term viability of glycerol stocks using optical density and colony count measurements

After one week, a glycerol stock of the strain of interest was taken and allowed to thaw, after which 50 μ l was taken and diluted 10,000-fold in MB. From this solution, a 50 μ l aliquot was spread on a non-selective MB agar plate, incubated for 12-16 hours at 37 °C and colonies were then counted, accounting for the dilutions made in the process. In parallel, a second 50 μ l aliquot of the solution was inoculated into 10ml MB and grown for 12-16 hours in a shaker-incubator at 37 °C at 200rpm. Following this, the OD₇₅₀ was measured. This procedure was repeated for 12 weeks in duplicates.

3.1.3 The viability of *Roseobacter* glycerol stocks in liquid and static cultures

Figure 3.1 shows that both the culturing of revived wildtype *Roseobacter* strains in liquid medium and on agar plates retains efficiency for at least 12 weeks following preparation of glycerol stocks. This indicates that master and working cell banks can be set up, minimising preparatory work for experiments and allowing for consistency and reproducibility throughout a batch of experiments. Both *R. denitrificans* and *O. indolicum* had an OD₇₅₀ of over 2.51 and a CFU/ml count of over 2.2×10^7 for the duration of the experiment. Cells grown in liquid cultures show more stability than cells grown on agar plates, with smaller variations seen between duplicate readings and over the weeks. *R. denitrificans* (Figure 3.1a), in particular, exhibit a more stable profile than *O. indolicum* (Figure 3.1b).

Results show that similarly to *E. coli*, *Roseobacter* strains can be grown in a range of temperatures (20 °C-37 °C), with increased growth noted for the higher temperatures (Bruhn et al., 2007; Christie-Oleza et al., 2012). This improves on results from Lafay et al. (1995) who had reported that *R. denitrificans* does not grow at 37 °C, contrary to results in Figure 3.1a. Unless otherwise specified, the culturing of *Roseobacter* strains in a temperature of 37 °C was applied as a measure to simplify the culturing process in subsequent investigations.

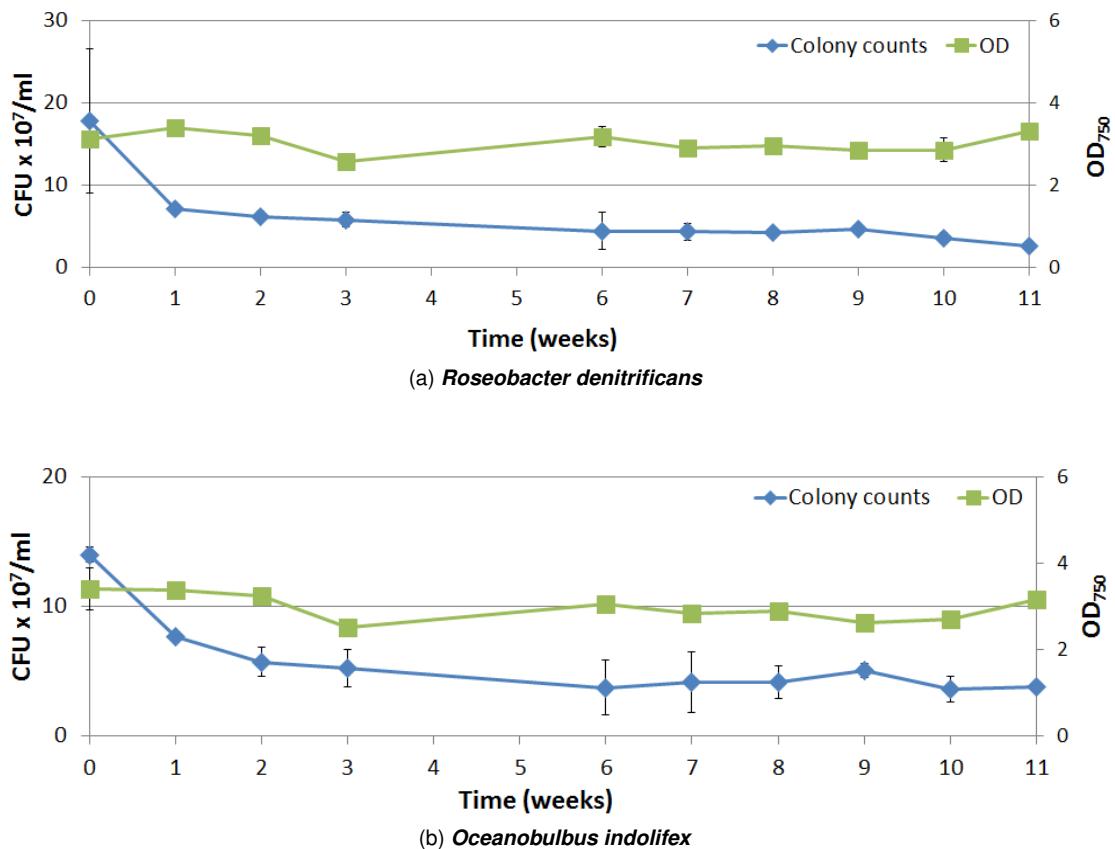


Figure 3.1: Long-term viability of a) *R. denitrificans* and b) *O. indolifex* glycerol stocks stored at -80°C and revived in shaken and static cultures over a period of 12 weeks. Readings for shaken cultures were taken via OD_{750} measurements (right axis of plots) while readings for static cultures were taken via CFU/ml measurements (left axis of plots). Readings were taken after 12-16 hours from initiation of culture. Error bars show standard deviation for duplicate readings.

3.2 Growth of *Roseobacter* strains in response to varying Chloramphenicol concentrations in the culture medium

Antibiotics are essential for the maintenance of plasmids equipped with antibiotic resistance genes (Vandermeulen et al., 2011). The following study was carried out to determine whether non-transformed wildtype *Roseobacter* cells are susceptible to growth inhibition in the presence of antibiotics, and if so, the threshold concentration of antibiotics which enables a distinction between transformed and non-transformed cells. The synthetic BioBrick™ plasmid that was used throughout transformation experiments had a Chloramphenicol resistance gene, hence, this study was carried out with Chloramphenicol antibiotic.

The Chloramphenicol concentrations chosen for testing (Table 3.1) were based on previous work carried out in Piekarski et al. (2009). All three *Roseobacter* strains (*R. denitrificans*, *O. indolifex* and *D. shibae*) were used. The assay was also carried out on *E. coli*, to use as a reference point.

Strain	Chloramphenicol concentrations tested				
	0 μ g/ml	3.4 μ g/ml	7 μ g/ml	17 μ g/ml	34 μ g/ml
<i>R. denitrificans</i>	✓	✓	-	✓	✓
<i>O. indolifex</i>	✓	✓	✓	✓	-
<i>D. shibae</i>	✓	✓	✓	✓	-
<i>E. coli</i>	✓	-	-	-	✓

Table 3.1: The different Chloramphenicol concentrations which each of the three *Roseobacter* strains and *E. coli* were cultured in, in order to assess cell growth inhibition by Chloramphenicol.

3.2.1 Preparation of *Roseobacter* strains in culture medium with varying Chloramphenicol concentrations

To analyse Chloramphenicol sensitivity of *R. denitrificans*, *O. indolifex* and *D. shibae*, the first step was to inoculate 100ml of MB with 100 μ l of cells from a glycerol stock and incubate it for 12-16 hours at 37°C at 200rpm. Four aliquots of 20ml were taken and each was supplemented with 20ml MB plus freshly prepared Chloramphenicol at a different concentration for a total of 40ml. Four Chloramphenicol concentrations were used for each of the different strains, as listed in Table 3.1.

Following this step, each 40ml volume was divided into sixteen 2ml volumes. The remaining volume was discarded. The 2ml aliquots were incubated at 37°C at 200rpm. 2ml volumes were taken out of incubation at hourly intervals at 0-6 hours and 8 hours post-incubation. OD₆₀₀ measurements were made in order to quantify cell growth. Two samples were measured at each time point in order to carry out the experiment in duplicate. Samples were then discarded. *E. coli* was taken through a similar protocol, replacing the use of MB with LB.

3.2.2 The Chloramphenicol concentration threshold for cell growth inhibition of *Roseobacter* strains

A selectivity of $3.4\mu\text{g}/\text{ml}$ Chloramphenicol for *R. denitrificans* (Figure 3.2b) and $7\mu\text{g}/\text{ml}$ for both *D. shibae* (Figure 3.2a) and *O. indolifex* (Figure 3.2c) resulted in minimal growth inhibition over a span of eight hours. Selectivity of $17\mu\text{g}/\text{ml}$ or over resulted in severe growth inhibition over eight hours for all three strains. Cells cultured in medium with Chloramphenicol concentrations higher than $17\mu\text{g}/\text{ml}$ did not recover growth over this span of time. This study confirmed that there exists a Chloramphenicol concentration threshold above which *Roseobacter* cell growth is inhibited. Hence, antibiotic selective screening can be used throughout all processes to promote growth of antibiotic resistant transformed *Roseobacter* strains while inhibiting growth of non-transformed cells.

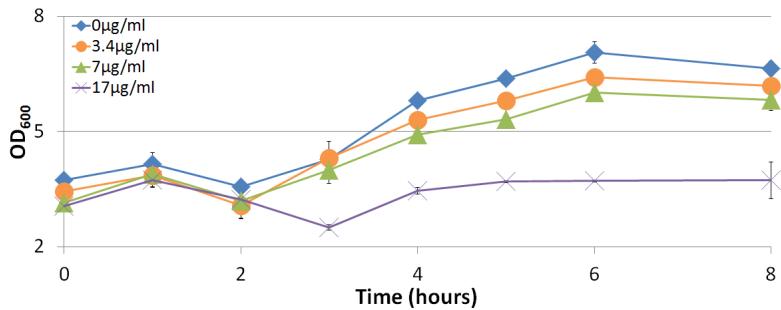
The growth profiles of the three *Roseobacter* strains show a different susceptibility to Chloramphenicol. The growth profiles of *D. shibae* cultured in $3.4\mu\text{g}/\text{ml}$ or $7\mu\text{g}/\text{ml}$ Chloramphenicol are similar to each other. The strain only showed growth inhibition when $17\mu\text{g}/\text{ml}$ Chloramphenicol was applied. *R. denitrificans* and *O. indolifex* showed more distinct profiles when cultured in different antibiotic concentrations. *R. denitrificans*, in particular, started to show distinct weak growth when as little as $3.4\mu\text{g}/\text{ml}$ Chloramphenicol was applied. It showed acute growth inhibition at $17\mu\text{g}/\text{ml}$, which is lower than the minimal inhibitory Chloramphenicol concentration ($30\mu\text{g}/\text{ml}$) reported for this strain in Piekarski et al. (2009).

E. coli was used in order to verify that Chloramphenicol remains effective over the time span over which observations were made. All optical density measurements were made at 600nm , as is standard for *E. coli*, instead of at 750nm as used previously. This modification is based on *Roseobacter* studies (Kalhoefer et al., 2011; Giebel et al., 2011) who similarly take optical density readings at 600nm . As when culturing *Roseobacter* in 37°C instead of 30°C , this was undertaken as a simplification measure in order to facilitate and standardise culturing of *Roseobacter*.

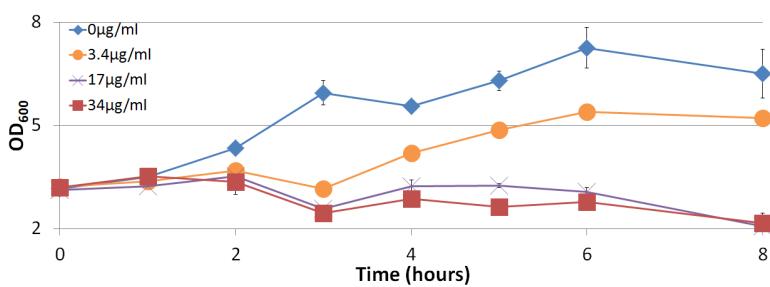
3.3 Transformation of *Roseobacter* clade bacteria using chemical transformation and electroporation methods

A fundamental procedure within a cloning process and within synthetic biology is the introduction of foreign plasmids into cell platforms. Therefore, different competence preparation and transformation protocols were implemented to analyse whether *Roseobacter* strains can be transformed with plasmids, and moreover, whether they can be transformed using a simple and straightforward technique. The selected protocols are based on established *Roseobacter* and *E. coli* methods, with an emphasis made on the development of an electroporation based method, rather than chemical transformation, due to the former method being more efficient in other bacterial organisms (Anthony, 2001).

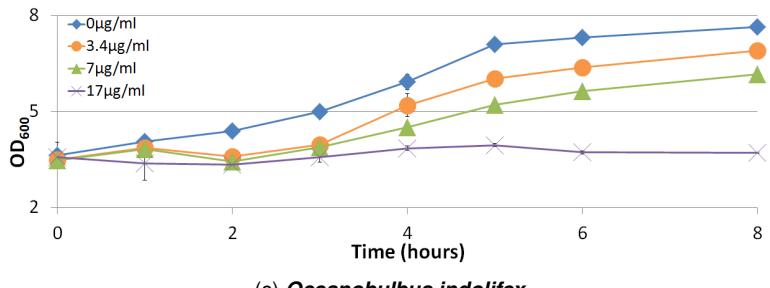
The synthetic plasmids used for the transformation procedures, including BioBricksTM, are listed in Table 3.2. The BioBrickTM part pSB3C5 (EU496103) was supplied by the Registry of Standard Biological Parts (MA, U.S.A). The pA0815 plasmid was purchased from InvitrogenTM (Life Technologies Ltd., Paisley, U.K) The pBBR1MCS1, pBBR1MCS2 and pBBR1MCS4 plasmids



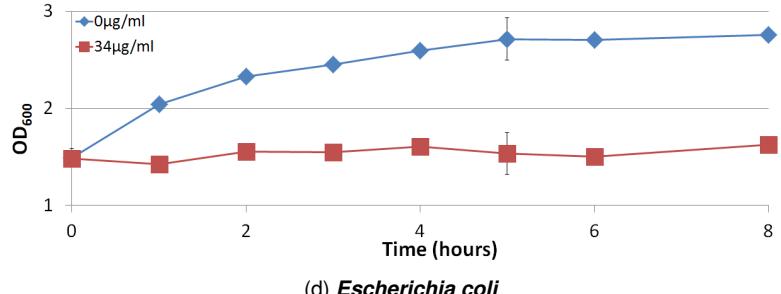
(a) *Dinoroseobacter shibae*



(b) *Roseobacter denitrificans*



(c) *Oceanobulbus indolifex*



(d) *Escherichia coli*

Figure 3.2: Growth of *Roseobacter* strains in liquid culture in response to varying Chloramphenicol concentrations for a) *D. shibae*, b) *R. denitrificans*, c) *O. indolifex* and d) *E. coli*. Each plot shows OD₆₀₀ readings versus Time in hours. Error bars show standard deviation for measurements carried out in duplicate.

	Plasmid	Ori	Selectivity
1	BioBrick TM pSB3C5	p15A	Chloramphenicol
2	pA0815	pBR322	Ampicillin
3	pHD1313	pUC, F1	Ampicillin
4	pRP _{GFP} SIR2rp3	pUC, 2 μ	Ampicillin
5	pUBeK	pUC	Ampicillin
6	pBBR1MCS1	pBBR1	Chloramphenicol
7	pBBR1MCS2	pBBR1	Kanamycin
8	pBBR1MCS4	pBBR1	Ampicillin

Table 3.2: BioBrickTM and engineered plasmids used in the chemical and electro transformation of *Roseobacter* strains.

were purchased from CBS-KNAW Fungal Biodiversity Centre (Utrecht, Netherlands). pH1313, pRP_{GFP}SIR2rp3 and pUBeK were provided as listed in Section 2.2.1. Transformation were also carried out on *E. coli* to help inform results.

3.3.1 Transformation based on variations of electrocompetence and electroporation methods

3.3.1.1 Efficiency of electroporation-based transformation as a function of plasmid mass

Cells were transformed using an electroporation protocol based on Piekarski et al. (2009). The authors present the protocol as ‘a fast and easy method to transfer plasmids into *Roseobacter* cells’ which gives transformation efficiencies up to the 10⁷ CFU/ μ g range. This is in line with the aims of this *Roseobacter* investigation to develop standardised straight-forward methods for the use of *Roseobacter* cells as a synthetic biology platform. Minor modifications were made to the protocol for simplification purposes, such as incubating cells in 37°C rather than 30°C. All three *R. denitrificans*, *O. indolicum* and *D. shibae* strains, along with *E. coli*, were transformed with the pSB3C5 BioBrickTM.

3.3.1.1.1 Preparation and electroporation of electrocompetent cells. *Roseobacter* strains were prepared by growing 50 μ l of non-competent cells from a glycerol stock in 50ml MB medium at 37°C and an agitation of 200rpm for 12-16 hours to an OD₆₀₀ of over 0.5. Samples were then centrifuged for 15 minutes at 4000rpm at 4°C. The supernatant was discarded and cell pellets were resuspended in 10ml pre-chilled 10% (v/v) sterile glycerol after which the centrifugation-and-resuspension process was repeated another four times, with the cell pellet being resuspended in 400 μ l 10% (v/v) glycerol in the last step. The suspension was divided into 50 μ l aliquots using pre-chilled 1.5ml eppendorf tubes. Samples were used straight away, therefore no storage was required. *E. coli* cells were prepared similarly, substituting the use of MB with LB medium.

Following preparation of electrocompetent cells, 5 μ l plasmid-DNA at the required concentration was added to 50 μ l electrocompetent cells in a pre-chilled 0.2cm pulser cuvette (Bio-Rad, CA, U.S.A). For each strain, four electroporation reactions were carried out using pSB3C5 at a mass

of 50-1000ng as listed in Table 3.3. Following preparation, the mixture was then pulsed in a Gene Pulse Xcell™ System (Bio-Rad, CA, U.S.A) using a field strength of 2.5kV, capacitance of $25\mu\text{F}$ and resistance of 200Ω . Afterwards, cells were supplemented with 1ml chilled MB and incubated for 12-16 hours at 37°C with shaking at 200rpm to allow for synthesis of the antibiotic resistant protein. Following this, $100\mu\text{l}$ of the culture was transferred onto MB agar plates with antibiotic selectivity as listed in Table 3.3 and incubated at 37°C . For each strain, two different Chloramphenicol concentrations were used, based on Chloramphenicol resistance results as seen in Figure 3.2; a concentration of $17\mu\text{g/ml}$ which inhibits cell growth of non-transformed cells and a lower concentration, dependant on the strain being used, at which cell growth, albeit not optimal, can still be observed. After plating, *D. shibae* and *R. denitrificans* plates were left to incubate for four days at 37°C while *O. indolifex* plates were left to incubate at 37°C for two days. In parallel, control transformations with *E. coli* were carried out as per above, replacing the use of MB with LB and incubating plates at 37°C for a total of 12-16 hours.

Table 3.3: Electroporation based transformation using different plasmid masses for electroporation. Reactions and controls carried out in electroporation transformation of *Roseobacter* and *E. coli* strains with the BioBrick™ pSB3C5. Each reaction lists the strain which was transformed, the plasmid mass used, whether cells were pulsed, the Chloramphenicol concentrations used to prepare transformation plates and the colony counts following incubation times. 'Lawn' refers to +250 small, non-isolated colonies.

	Strain	Plasmid Mass (ng)	Pulse	Chloramphenicol ($\mu\text{g/ml}$)	Colony Count
1	<i>R. denitrificans</i>	50	Y	7, 17	0, 0
2	<i>R. denitrificans</i>	250	Y	7, 17	0, 0
3	<i>R. denitrificans</i>	750	Y	7, 17	0, 0
4	<i>R. denitrificans</i>	1000	Y	7, 17	0, 0
5	<i>R. denitrificans</i>	500	N	-	250
6	<i>R. denitrificans</i>	1000	N	-	200
7	<i>R. denitrificans</i>	-	N	-	200
8	<i>R. denitrificans</i>	-	Y	-	150
9	<i>R. denitrificans</i>	-	N	7, 17	0, 0
10	<i>O. indolifex</i>	50	Y	10, 17	0, 0
11	<i>O. indolifex</i>	250	Y	10, 17	0, 0
12	<i>O. indolifex</i>	750	Y	10, 17	0, 0
13	<i>O. indolifex</i>	1000	Y	10, 17	0, 0
14	<i>O. indolifex</i>	500	N	-	~50
15	<i>O. indolifex</i>	1000	N	-	~20
16	<i>O. indolifex</i>	-	N	-	~20
17	<i>O. indolifex</i>	-	Y	-	~20
18	<i>O. indolifex</i>	-	N	10, 17	0, 0
19	<i>D. shibae</i>	50	Y	10, 17	0, 0
20	<i>D. shibae</i>	250	Y	10, 17	0, 0
21	<i>D. shibae</i>	750	Y	10, 17	0, 0
22	<i>D. shibae</i>	1000	Y	10, 17	0, 0
23	<i>D. shibae</i>	500	N	-	lawn

Table 3.3: **Continued - Electroporation based transformation using different plasmid masses for electroporation.** Reactions and controls carried out in electroporation test to transform *Roseobacter* and *E. coli* strains with the BioBrick™ pSB3C5. Each reaction lists the strain which was transformed, the plasmid mass used, whether cells were pulsed, the Chloramphenicol concentrations used to prepare transformation plates and the colony counts following incubation times. ‘Lawn’ refers to +250 small, non-isolated colonies.

	Strain	Plasmid Mass (ng)	Pulse	Chloramphenicol (μg/ml)	Colony Count
24	<i>D. shibae</i>	1000	N	-	lawn
25	<i>D. shibae</i>	-	N	-	lawn
26	<i>D. shibae</i>	-	Y	-	lawn
27	<i>D. shibae</i>	-	N	10, 17	0, 0
28	<i>E. coli</i>	50	Y	34	0
29	<i>E. coli</i>	250	Y	34	0
30	<i>E. coli</i>	750	Y	34	0
31	<i>E. coli</i>	1000	Y	34	~20
32	<i>E. coli</i>	500	N	-	150
33	<i>E. coli</i>	1000	N	-	500
34	<i>E. coli</i>	-	N	-	lawn
35	<i>E. coli</i>	-	Y	-	500
36	<i>E. coli</i>	-	N	34	0

3.3.1.1.2 Effectiveness of electroporation-based approach using varying plasmid mass for *Roseobacter* transformation. Colony counts from transformation plates (for example, Reactions 1-4), listed in Table 3.3 showed that the three marine strains did not retain the BioBrick™ plasmid following electroporation. Competent cells which were mixed with the plasmid but not pulsed (for example, Reactions 5-6) showed good growth, indicating that the strain was not damaged by the presence of the plasmid. Similarly, growth on plates of electrocompetent cells which were pulsed/not pulsed in the absence of the plasmid (for example, Reactions 7-8) indicated that both the preparation of electrocompetent cells and the electroporation shock were not damaging to the strains. No growth was observed for non-pulsed cells spread onto selective plates (for example, Reaction 9), indicating that no antibiotic-resistance contamination was present. Similar results were observed for *E. coli* reactions. In this case, the electroporation worked when using a very large DNA mass (1μg) (20 colonies were noted). The pSB3C5 BioBrick™ had a 260/280 purity ratio of 1.9, making it unlikely that plasmid impurities contributed to the observed results.

The high colony counts from *D. shibae* transformation plates indicated that the strain was the more robust of the *Roseobacter* strains tested. Moreover, no differences were observed in colony counts on transformation plates which had low and high antibiotic selectivities, indicating that even the lower concentration effectively inhibits growth for non-resistant cells grown on selective agar plates. However, based on previous findings, the higher concentration was retained to distinguish between plasmid-positive and plasmid-negative *Roseobacter* cells.

	Plasmid Mass (ng)	Voltage (kV)	Chloramphenicol (µg/ml)	Colony Count
1	250	0.5	17	0
2	250	1.0	17	0
3	250	1.5	17	0
4	250	3.0	17	0
5	1000	0.5	17	0
6	1000	1.0	17	0
7	1000	1.5	17	0
8	1000	3.0	17	0
9	250	0.5	-	~10 & lawn
10	250	1.0	-	~20 & lawn
11	250	1.5	-	~10 & lawn
12	250	3.0	-	~10 & lawn
13	1000	0.5	-	~5 & lawn
14	1000	1.0	-	~10 & lawn
15	1000	1.5	-	~20 & lawn
16	1000	3.0	-	~10 & lawn

Table 3.4: **Transformation using different output voltages in electroporation.** Reactions and controls used in electroporation transformation of *D. shibae* with the BioBrick™ pSB3C5. Each reaction lists the BioBrick™ mass used, the voltage set during electroporation, the Chloramphenicol concentration with which plates were prepared and colony counts following incubation times. 'Lawn' refers to +250 small, non-isolated colonies.

3.3.1.2 Efficiency of electroporation-based transformation as a function of voltage

3.3.1.2.1 Preparation and electroporation of electrocompetent cells. As a means of optimisation, cell preparation and electroporation steps were as per the previous method, with one main difference; cells were pulsed using an output voltage between 0.5kV and 3kV, instead of the previously used 2.5kV voltage. The pSB3C5 BioBrick™ plasmid was used at a mass of 250ng or 1µg for insertion into *D. shibae*. If shown to be successful, the modified protocol would then be applied to *R. denitrificans* and *O. indolifex*. A list of the reactions and controls which were prepared can be found in Table 3.4. All reactions and controls were pulsed.

3.3.1.2.2 Effectiveness of electroporation-based approach using varying voltage for *Roseobacter* transformation. Colony counts listed in Table 3.4 show that *D. shibae* cells which were mixed with BioBrick™ plasmid, pulsed and spread onto selective plates (Reactions 1-8) did not retain the plasmid, irrespective of the plasmid mass and voltage used. Colony counts for cells taken through the same process but spread onto non-selective plates (Reactions 9-10) indicated that cells were not compromised by the electroporation. Controls from the previous electroporation process (Table 3.3) had already shown that cells retained viability in the presence of the plasmids and following the electrocompetence preparation procedure, and that no antibiotic resistant contamination was present.

3.3.1.3 Efficiency of electroporation-based transformation based on a modified electroporation protocol and a broader range of plasmid ori sequences

Based on previous results, a different electrocompetence preparation and electroporation protocol, widely used for *E. coli*, was tested on *D. shibae* cells. In addition, four different engineered plasmids were utilised instead of pSB3C5; pA0815, pHD1313, pRP_{GFP}SIR2rp3 and pUBeK. These plasmids have different origins of replication and antibiotic resistance genes, as seen in Table 3.2.

This *E. coli* based protocol (Sambrook et al., 2001) was investigated for applicability with the *D. shibae* strain due to a lack of *Roseobacter*-specific protocols available. In one of the few related studies, a protocol developed by D'Alvise (2013) is used to transform *Ruegeria mobilis*, which is a strain in the *Rhodobacteraceae* family in which the *Roseobacter* clade is also nested. Similarly, Miller and Belas (2006) transform strains from the *Silicibacter* genus which is also nested in *Rhodobacteraceae*. Although the preparation and pulsing conditions are similar to the electroporation runs discussed above, both protocols use different culturing temperatures and more complex media which are not applicable to *Roseobacter* strains. This renders the *Rhodobacteraceae*-optimised methods infeasible for the purpose of the current investigation.

3.3.1.3.1 Preparation and electroporation of electrocompetent cells. 10ml of MB was inoculated with 10 μ l of cells from a *D. shibae* glycerol stock and incubated for 12-16 hours in a shaker-incubator at 37°C and agitation of 200rpm. Following this, 400ml MB were inoculated with 4ml of the culture and incubated at 37°C at 200rpm until OD₆₀₀ was 0.3 or higher. The culture was chilled on ice for 30 minutes, transferred to pre-chilled centrifuge bottles and centrifuged for 15 minutes at 6000rpm at 4°C. The supernatant was discarded by decanting and the pellet was resuspended in one volume of ice-cold sterile distilled H₂O. The centrifugation, supernatant discarding and pellet resuspension steps were repeated another two times, with the supernatant being removed through aspiration instead of decanting. Additionally, in the final step, cells were resuspended in 50ml of ice-cold sterile 10% (v/v) glycerol and transferred to a pre-chilled 50ml centrifuge tube. A final 15 minute centrifugation step was carried out at 6000rpm at 4°C, the supernatant was removed using a pipette and the pellet resuspended in 2ml of ice-cold sterile 10% (v/v) glycerol. This was then divided into 50 μ l aliquots in pre-chilled PCR tubes, which were kept on ice and used immediately.

In summary, the preparation of electrocompetent cells (Sambrook et al., 2001) differs from the previous both in the two-step inoculation process which is used to grow non-competent cells initially and in the harvesting of cells. This includes using a different speed of centrifugation (6000rpm instead of 4000rpm), resuspension volume (one volume instead of a standard 10ml), resuspension solution (ice-cold sterile distilled H₂O instead of pre-chilled 10% (v/v) sterile glycerol) and final resuspension volume (2ml instead of 400 μ l).

Following the preparation of electrocompetent cells, 5 μ l plasmid DNA at the required concentration was added to 50 μ l electrocompetent cells in a pre-chilled 0.2cm pulser cuvette. A list of the reactions and controls prepared can be found in Table 3.5. Only one plasmid mass, of 500ng, was used for electroporation. The mixture was then pulsed in a Gene Pulse Xcell™ System using a field strength of 2.5kV, capacitance of 25 μ F and resistance of 200 Ω . After pulsing, cells were

	Plasmid	Plasmid Mass (ng)	Ampicillin ($\mu\text{g/ml}$)	Colony Count
1	pAO815	500	100	0
2	pHD1313	500	100	0
3	pRP _{GFP} SIR2rp3	500	100	0
4	pUBeK	500	100	0
5	pAO815	500	-	180
6	pHD1313	500	-	190
7	pRP _{GFP} SIR2rp3	500	-	150
8	pUBeK	500	-	140
9	-	-	-	170

Table 3.5: Transformation using electrocompetence protocol by Sambrook et al. (2001). Reactions and controls used to transform *D. shibae* using a modified electroporation protocol and a broader range of plasmid ori sequences. Each reaction lists the plasmid and plasmid mass used to transform *D. shibae* with, the Ampicillin concentration used to prepare plates with and the colony count following the incubation period.

supplemented with 1ml of pre-warmed MB and incubated for one hour at 37°C with shaking at 200rpm to allow expression of the antibiotic resistant proteins. Following incubation, the culture was then centrifuged for three minutes at 8000rpm and the supernatant was removed, except for 100 μl in which the pellet was resuspended in order to condense cell density. The culture was then divided equally between selective and non-selective plates and left to incubate for four days at 37°C.

Thus, it can be seen that this protocol is similar to the previous in the electroporation conditions used, but differs in the treatment of cells following electroporation, allowing cells to recover for only one hour prior to plating, instead of 12-16 hours.

3.3.1.3.2 Effectiveness of modified electrocompetence and broader plasmid range approach to *Roseobacter* transformation. Colony counts from transformation plates wherein electrocompetent *D. shibae* cells were mixed with plasmid, pulsed and spread onto selective plates (Reactions 1-4 in Table 3.5) indicated that cells did not retain the plasmid following electroporation. Colony counts of transformation plates wherein pulsed cells were spread onto non-selective plates (Reactions 5-8) showed good growth, indicating that the plasmids were not toxic to the cells. Colony counts from transformation plates on which *D. shibae* cells which were taken through the process using a blank instead of a plasmid (Reaction 9) indicated that cells taken through the preparation and electroporation process retained viability.

3.3.2 Transformation using the Piekarski et al. electroporation method

In light of the above results, the next step was to apply the preparation and pulsing procedure developed by Piekarski et al. (2009) for the transformation of *D. shibae* and *E. coli* cells. The range of plasmids used for the transformation was widened again. In addition to the ones used above, the

electroporation was carried out using three additional plasmids which were used in Piekarski et al. (2009): pBBR1MCS1, pBBR1MCS2 and pBBR1MCS4 (Kovach et al., 1994). The plasmids have the same backbone but have different antibiotic resistance genes (Table 3.2). The pBBR1MCS1-4 plasmids were prepared as per supplier instructions. Only pBBR1MCS1 and pBBR1MCS2 were used in the electroporation protocol.

3.3.2.1 Preparation and electroporation of electrocompetent cells

Plasmids were used at a mass of 50ng as per Table 3.6. Cells were prepared as per Section 3.3.1.1.1, using incubation temperatures of 30°C instead of 37°C and measuring optical density of cells at 578nm until density is at 0.5 instead of taking optical density measurements at 600nm. Following electroporation and incubation, instead of plating cells which were suspended in an MB solution, cells were serially diluted once in 1.7% sea-salts (Sigma-Aldrich, Munich, Germany) to get concentrations of 1X and 0.5X. 100 μ l of the cells were plated on hMB agar plates (instead of MB agar plates) with selective drug and without, as a control measure. *D. shibae* plates were then left to incubate for four days at 30°C.

When using *E. coli*, the same electroporation process was carried out, replacing use of MB and hMB with LB, using 37°C instead of 30°C incubation temperatures and incubating LB agar plates for 12-16 hours.

3.3.2.2 Effectiveness of the Piekarski et al. approach to *Roseobacter* transformation

Results from colony counts following incubation periods, as seen in Table 3.6, indicated that the electroporation was successful, with colonies noted on all selective plates on which electroporated cells were grown (Reactions 1-6). By comparing colony counts of *D. shibae* plates wherein electroporated cells were spread onto non-selective plates (Reactions 7-12) with those spread onto selective plates (Reactions 1-6), a larger number of colonies was observed on the former, indicating that the antibiotics were being effective and that colony counts on selective plates resulted from plasmid-positive cells.

Based on these counts, two colonies from each of the selective plates were grown for 12-16 hours in selective LB at a temperature of 30°C and agitation of 200rpm, after which the plasmid DNA was extracted, digested and analysed via gel electrophoresis. Gel results (not shown) of the analytical digests showed no clear bands, indicating that the plasmids which were used had been discarded or broken down by the cell. This was corroborated by the low concentration of the DNA sample following elution.

This is in contrast to results reported in Piekarski et al. (2009) who used a similar procedure to isolate plasmid DNA. Irrespective of this, alternative methods can be adopted to confirm the presence of inserted plasmids in *Roseobacter*. One option is to add EDTA to reduce nuclease activity which may be inhibiting the isolation process. Otherwise, a phenol-chloroform extraction method can be tested. In order to bypass DNA extraction, a plasmid encoding a fluorescent gene (Piekarski et al., 2009) or allowing for blue/white screening can be used, although available literature indicated that the latter has not been used with a *Roseobacter* culture beforehand. A third

	Strain	Plasmid	Plasmid Mass (ng)	Antibiotic (µg/ml)	Colony Count
1	<i>D. shibae</i>	pAO815	50	Amp 100	100
2	<i>D. shibae</i>	pHD1313	50	Amp 100	100
3	<i>D. shibae</i>	pRP _{GFP} SIR2rp3	50	Amp 100	100
4	<i>D. shibae</i>	pUBeK	50	Amp 100	20
5	<i>D. shibae</i>	pBBR1MCS1	50	Cm 17	lawn
6	<i>D. shibae</i>	pBBR1MCS2	50	Kan 100	2
7	<i>D. shibae</i>	pAO815	50	-	lawn
8	<i>D. shibae</i>	pHD1313	50	-	lawn
9	<i>D. shibae</i>	pRP _{GFP} SIR2rp3	50	-	lawn
10	<i>D. shibae</i>	pUBeK	50	-	lawn
11	<i>D. shibae</i>	pBBR1MCS1	50	-	lawn
12	<i>D. shibae</i>	pBBR1MCS2	50	-	lawn
13	<i>D. shibae</i>	-	-	-	lawn
14	<i>E. coli</i>	pAO815	50	Amp 100	20
15	<i>E. coli</i>	pHD1313	50	Amp 100	100
16	<i>E. coli</i>	pRP _{GFP} SIR2rp3	50	Amp 100	140
17	<i>E. coli</i>	pUBeK	50	Amp 100	200
18	<i>E. coli</i>	pBBR1MCS1	50	Cm 17	0
19	<i>E. coli</i>	pBBR1MCS2	50	Kan 100	140
20	<i>E. coli</i>	pAO815	50	-	lawn
21	<i>E. coli</i>	pHD1313	50	-	lawn
22	<i>E. coli</i>	pRP _{GFP} SIR2rp3	50	-	lawn
23	<i>E. coli</i>	pUBeK	50	-	lawn
24	<i>E. coli</i>	pBBR1MCS1	50	-	lawn
25	<i>E. coli</i>	pBBR1MCS2	50	-	lawn
26	<i>E. coli</i>	-	-	-	lawn

Table 3.6: **Transformation using the method by Piekarski et al. (2009).** Reactions and controls used in the electroporation of *D. shibae* and *E. coli* cells with plasmids having a range of ori sequences. For each reaction, the table lists the strain which was transformed, the plasmid and plasmid mass used, the antibiotic concentrations which was used to prepare transformation plates and colony counts following incubation times. 'Amp' refers to Ampicillin, 'Cm' to Chloramphenicol and 'Kan' to Kanamycin. 'Lawn' refers to +250 small, non-isolated colonies.

alternative is to colony boil PCR transformed cells (Buchan et al., 2000).

Similar results were observed for *E. coli* reactions (Reactions 14-25), indicating that plasmids were not retained by the host cells. Reactions wherein *D. shibae* and *E. coli* were taken through the electrocompetence and electroporation processes without adding any plasmids and then plated onto non-selective plates (Reactions 13, 26 respectively) showed good growth indicating that the procedure was not compromising the cells.

3.3.3 Transformation using the heat shock method

Given the above results, an alternative strategy was adopted and a heat-shock based chemical transformation was applied instead. A CaCl_2 -based chemical transformation method was used to transform the *D. shibae* strain with the pSB3C5 plasmid.

3.3.3.1 Preparation and transformation using the heat shock method of chemically competent marine bacterial cells

D. shibae cells were prepared and transformed as per Section 2.2.6 and 2.2.7.1 respectively, replacing the use of LB with MB. A list of reactions carried out can be found in Table 3.7.

3.3.3.2 Effectiveness of the heat-shock method to *Roseobacter* transformation

Colony counts from transformation plates (Table 3.7) indicated that whereas the transformation was successful when using *E. coli* hosts (Reactions 8-10), transformation of *D. shibae* (Reactions 1-3) was ineffective. Colony counts for transformation reactions which were plated onto non-selective plates (Reactions 4-5) showed growth, indicating that the plasmid was not toxic to the cells. Colony counts for the reaction wherein no plasmid was added to chemically competent cells taken through transformation and plated on non-selective plates (Reaction 6) indicated that the transformation process did not compromise *D. shibae*. Finally, having eliminated the probability that the electroporation and transformation protocol were toxic, colony counts for reactions wherein no plasmid was added to cells taken through the transformation process and plated onto selective plates (Reaction 7) showed that no antibiotic-resistant contamination was present. Piekarski et al. (2009) also carried out a RbCl_2 chemical transformation in their study of *Roseobacter* cells. Similarly, the transformation was not successful and therefore was not replicated here.

A summary of the electroporation and chemical transformations which were made can be found in Figure 3.3, which shows colony counts from *D. shibae* and *E. coli* selective transformation plates following incubation. The plot highlights the wide range of plasmid mass and voltage conditions examined. Repeatedly, no positive colonies for *D. shibae* were detected on selective transformation plates. Based on this and controls which show that cells were consistently viable and uncontaminated, it was concluded that the transformation of *Roseobacter* cells using either electroporation or heat-shock based methods is a non-trivial process and that further work is required to develop an efficient, reproducible transformation procedure for *Roseobacter* cells.

	Strain	Plasmid	Plasmid Mass (ng)	Chloramphenicol ($\mu\text{g}/\text{ml}$)	Colony Count
1	<i>D. shibae</i>	pSB3C5	100	17	0
2	<i>D. shibae</i>	pSB3C5	250	17	0
3	<i>D. shibae</i>	pSB3C5	500	17	0
4	<i>D. shibae</i>	pSB3C5	100	-	lawn
5	<i>D. shibae</i>	pSB3C5	500	-	lawn
6	<i>D. shibae</i>	-	-	-	lawn
7	<i>D. shibae</i>	-	-	17	0
8	<i>E. coli</i>	pSB3C5	100	17	~ 200
9	<i>E. coli</i>	pSB3C5	250	17	~ 200 & lawn
10	<i>E. coli</i>	pSB3C5	500	17	~ 200 & lawn
11	<i>E. coli</i>	pSB3C5	100	-	lawn
12	<i>E. coli</i>	pSB3C5	500	-	lawn
13	<i>E. coli</i>	-	-	-	lawn
14	<i>E. coli</i>	-	-	17	0

Table 3.7: **Transformation using the heat-shock method.** Reactions and controls used to transform *D. shibae* with the pSB3C5 BioBrick™ using chemical instead of electro transformation. For each reaction, the table shows the strain which was transformed, the plasmid and plasmid mass which were used, the Chloramphenicol concentration used on transformation plates and the colony counts on plates following incubation times.

3.4 Discussion

Three strains drawn from different lineages of the *Roseobacter* phylogenetic tree, *O. indolicum*, *D. shibae*, and *R. denitrificans* were investigated to establish fundamental molecular biology procedures which enable the construction and characterisation of synthetic GRNs within *Roseobacter* platforms. This investigation was the first of its kind to be carried out with the aim of establishing *Roseobacter* as a chassis within synthetic biology. The following findings and conclusions were made:

1. Demonstrated that cell banks of *Roseobacter* marine strains can be developed and maintained for at least 12 weeks, possibly more. Glycerol stocks stored at -80 °C were successfully revived in both shaken and static culture for the duration of this period.
2. Demonstrated that antibiotic selectivity can be applied throughout cloning and transformation procedures to select for plasmid-positive cells. Wildtype *Roseobacter* growth was severely inhibited in the presence of Chloramphenicol concentrations as low as 17 $\mu\text{g}/\text{ml}$. This was lower than the previously reported 30 $\mu\text{g}/\text{ml}$ concentration (Piekarski et al., 2009).
3. *Roseobacter* strains were successfully cultured in temperatures up to 37 °C.
4. Mapped the region of experimental design space within which three *Roseobacter* species are resistant to plasmid uptake with respect to plasmid DNA concentration, plasmid origin of replication sequence, pulsing voltage and heat-shock/electroporation method used. Throughout, *Roseobacter* cells were shown to have retained viability following the addition

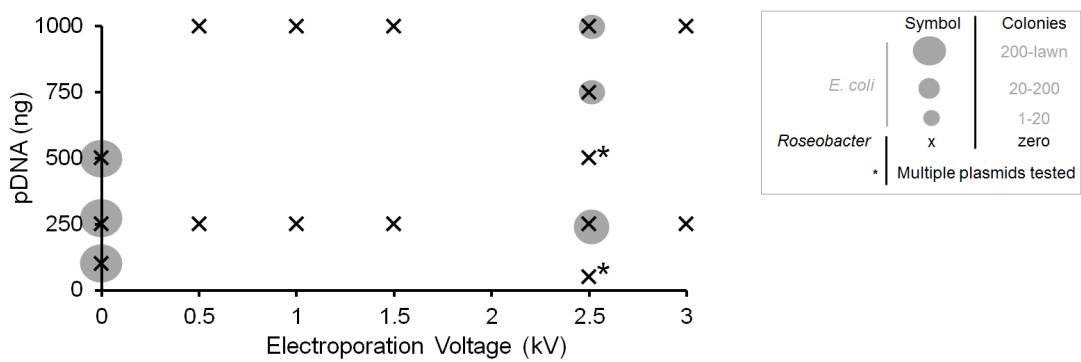


Figure 3.3: Plot mapping the different electroporation and chemical transformations which were run to insert foreign plasmids into *D. shibae* and *E. coli* strains. The colony counts from selective transformation plates for each insertion are shown as a function of the mass of plasmid DNA in ng used in the transformation (y-axis) against the voltage in kV used during the transformation (x-axis). Colony counts are as per the legend on the right. Colony counts marked with an * indicate that multiple transformations using different engineered plasmids were run under the specified conditions.

of synthetic plasmid and the transformation process.

5. The development of a simple and rapid transformation protocol for *Roseobacter* strain is non-trivial and further investigative work is required in order to both enable the development of synthetic GRNs in *Roseobacter* and to develop a de-skilled practice (Tucker, 2012) which allows for the participation of both academics and citizen scientists.

Chapter 4

Design and construction of the stable and unstable trypanosomal Goodwin oscillators

In the scope of establishing *T. brucei* as a synthetic biology chassis by developing a Goodwin oscillator encoded by a novel synthetic GRN, the first steps were to design and assemble the trypanosomal synthetic plasmid. Results of this process to develop both a stable and unstable Goodwin oscillator are presented here.

4.1 The stable and unstable Goodwin oscillators

The Goodwin oscillator was developed as a simplified mathematical model of the circadian rhythm (Goodwin, 1963). It has two main components, as seen in Figure 4.1; a single gene (*gene A*) which expresses a protein A, and an operator to which the protein A binds and inhibits further expression. This inhibitory regulation of expression is represented by the red arrow in Figure 4.1. In order to re-start expression, there are two options; to allow protein A to degrade via natural processes, following which expression re-initiates autonomously or to add an inducer molecule which causes protein A to unbind from the operator and instead bind to the inducer towards which it has a greater affinity.

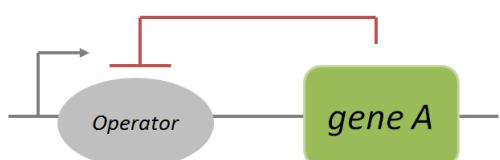


Figure 4.1: Diagram of a theoretical Goodwin oscillator showing the two main components of the GRN; *gene A* and the associated *operator*. *Gene A* produces a protein A which binds to the *operator*, resulting in inhibitory feedback which shuts down protein synthesis, as represented by the red arrow.

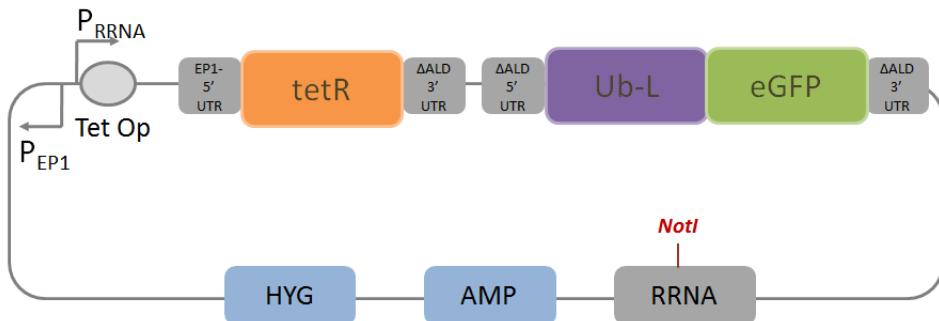


Figure 4.2: Diagram of the plasmid encoding the stable Goodwin oscillator which is to be constructed showing the main molecular species in the network: the P_{RRNA} and P_{EP1} promoters; the tet operator, Tet Op; the *tetR* gene; the Ub-L-eGFP fusion gene; the 5' and 3' UTRs for both sets of genes; the Hygromycin ('Hyg') and Ampicillin ('Amp') selectable markers and the rRNA gene targeting locus having a *NotI* restriction site.

4.1.1 The GRN which encodes the stable Goodwin oscillator

The novel synthetic GRN which encodes the stable Goodwin oscillator, pStable (Figure 4.2), has singular dominant negative feedback. The device has four main sets of components; the regulatory gene and operator equivalent to gene A and the operator in Figure 4.1 respectively plus additional parts which allow for correct recombinant DNA and assembly procedures and facilitate characterisation efforts, as discussed below:

1. **Genes & Proteins.** One of the main features of the network is the *tetR* gene which expresses the Tetracycline Repressor (TetR) protein. This protein binds to the tetracycline operator (Tet Op/tet operator) and inhibits its own expression and that of all genes within the same operon, based on the inhibiting mechanism found on the *E. coli* Tn10 transposon (Hillen & Berens, 1994). The *tetR* gene is equivalent to *gene A* in Figure 4.1 and is responsible for regulation of gene expression. The tetracycline regulated system was chosen not only because of its relative simplicity but also because it is a tightly regulated system across several species (Gossen & Bujard, 1992; Kamionka et al., 2004) and has been shown to work well in trypanosomes in multiple expression regulation studies (Alibu et al., 2005; Peacock et al., 2005).

A modified Green Fluorescent Protein (eGFP) which acts as a reporter protein to indirectly reflect TetR levels in the network (Zhang et al., 1996) is expressed within the same operon and is therefore under the influence of the tet operator. Thus, the GRN is a proof-of-principle construct which uses a fluorescing protein as an output to enable straightforward characterisation.

The N-terminus of the eGFP gene is fused to a mutated Ubiquitin gene, a Ubiquitin degron, wherein the 76th (and final) amino acid is changed to Leucine (Ub-L) as seen in Dantuma et al. (2000). The fusion serves to flag up eGFP for ubiquitin-proteasome dependant degradation. The final amino acid replacement is based on the N-end rule which states that a protein's half-life is determined by the amino acid at its terminus (Bachmair et al., 1986). Leucine is calculated to result in a protein half-life of 5-6 hours in HeLa cells (Dantuma et al., 2000; Gonda et al., 1989). Failing this destabilisation step, regulation through TetR feedback is masked as the wildtype eGFP reporter protein does not have a sufficiently fast turnover to reflect changes from TetR inhibition (Li et al., 1998). This means that the trans-

fected organism would fluoresce irrespective of whether expression is ongoing or not.

The network has an Ampicillin gene used during *E. coli* assembly stages to select for transformants and a Hygromycin gene which serves to select stable *T. brucei* transfectants.

2. **Operators, Promoters & Selectable Markers** The plasmid has one operator, the tet operator, which lies downstream of the trypanosomal, class I, tetracycline-dependant P_{RRNA} promoter (Schimanski et al., 2004) and to which the TetR protein can bind and inhibit expression. The *tetR* and Ub-L-eGFP genes are under the influence of the P_{RRNA} promoter.

Apart from this, the network has a P_{EP1} promoter from the EP procyclin locus (Alibu et al., 2005) which drives expression of Hygromycin and a P_{Amp} promoter (not shown in Figure 4.2) which drives expression of the Ampicillin resistance gene in *E. coli*.

3. **Untranslated regions** The *tetR* gene is flanked by two untranslated region sites (UTRs), EP1-5' and Δ ALD-3', which are regions in mRNA which influence gene expression and stability (Drozdz & Clayton, 1999; Clayton, 1999). The former is an EP1 procylic 5'-UTR, while the latter is a truncated fructose biphosphate aldolase 3'-UTR. The EP1 UTR contains both activating and inhibiting sequences which control stability and translation (Schurch et al., 1997). Δ ALD-3' has been shown to increase expression levels in BSF trypanosomes (Clayton & Shapira, 2007). The aldolase UTR can be truncated since specific sites such as the initial 136 bases are not required for regulation (Hug et al., 1993; Hotz et al., 1995).

The eGFP gene is flanked by a Δ ALD -5' and a Δ ALD 3' region, which serve similar purposes.

4. **rRNA targeting region** The rRNA region contains a NotI restriction site at which the plasmid can be linearised. The plasmid is then introduced into the trypanosomal genome via transfection, wherein the plasmid is inserted via recombination at a non-transcribed rRNA spacer loci (Alsfeld & Horn, 2008, 2005).

4.1.2 Emergence of oscillations in protein production in the stable Goodwin oscillator

Figure 4.3 explains how oscillatory protein production within the stable GRN results from auto-inhibitory transcriptional regulation by the TetR protein. There is no positive feedback within the GRN. In terms of protein expression dynamics, it is assumed that at time zero there is no ongoing protein production (Figure 4.3a). The P_{RRNA} promoter initially drives expression of *tetR* and eGFP, resulting in increasing TetR and fluorescence levels (Figure 4.3b). Once dimerised (not shown), TetR protein can bind to the tet operator and this leads to a halt in expression (Figure 4.3c). TetR and eGFP levels then decline due to natural degradation processes (Figure 4.3d). Once TetR protein bound to the operator is degraded, the pathway for expression is re-activated and the oscillation process repeats itself (Figure 4.3a). This increase, peak and decrease in protein levels is considered to be an oscillation.

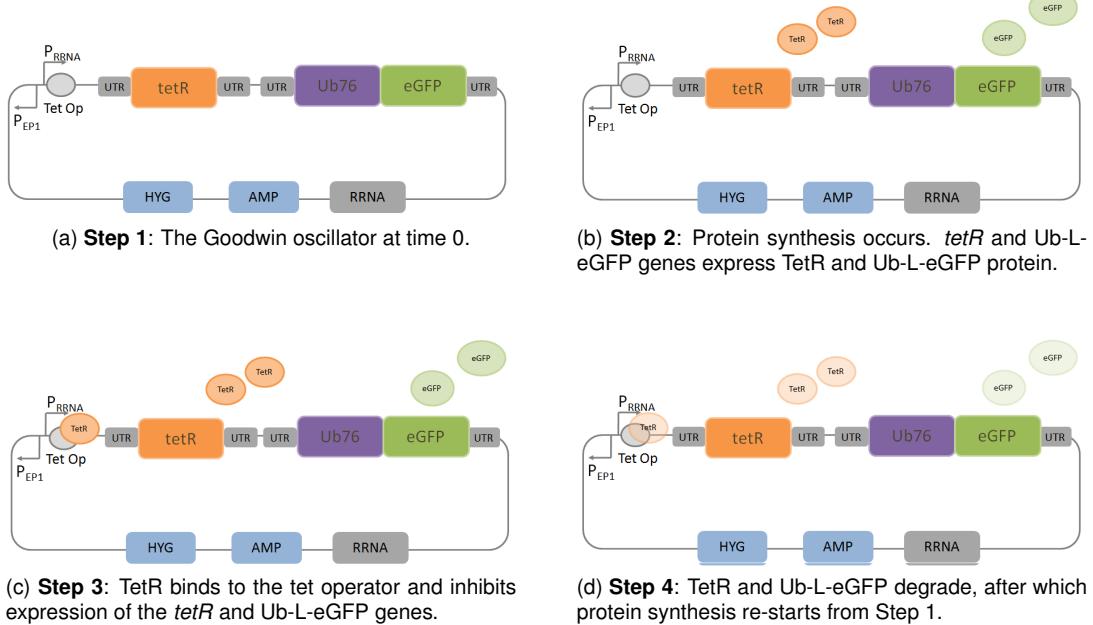


Figure 4.3: Diagram explaining how oscillatory protein production dynamics in the Goodwin oscillators occur as a result of the negative feedback loop within the GRN. The negative feedback loop results from the binding of TetR to the tet operator.

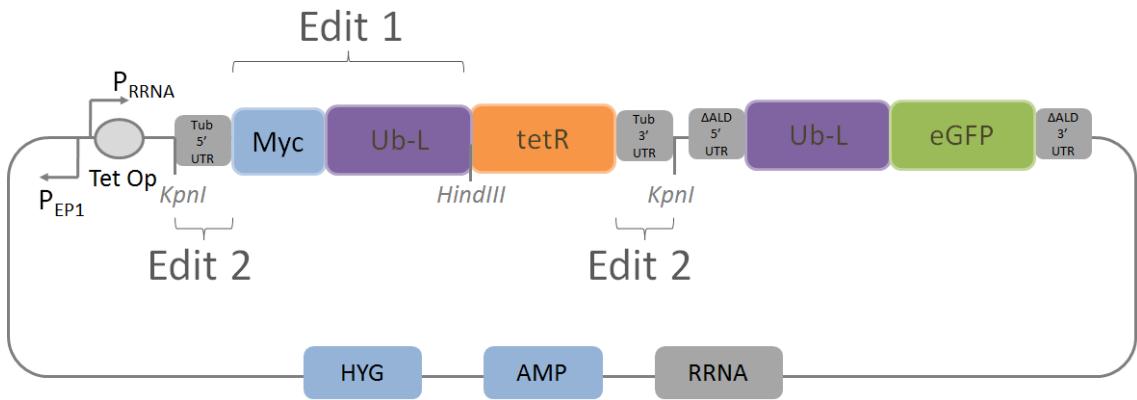
4.1.3 Theoretical design principles implemented in the stable Goodwin oscillator which enable oscillations

The Goodwin oscillator incorporates a set of fundamental design principles, discussed in Section 1.2.1.5 which allow for oscillations to occur: negative feedback, time-delay and non-linearity. A dominant negative feedback loop is implemented through the auto-inhibition of *tetR* synthesis, wherein a TetR dimer binds to the tet operator and represses expression. No positive loops are present in the network. Although this is not expected to have an effect on the nature of the dynamics it is likely to have an effect on the quality of the oscillations, in that oscillations will be less consistent (Tsai et al., 2008; Stricker et al., 2008).

The second necessary design feature, time-delay, is implemented due to natural synthesis processes (transcription, translation, post-translation) within eukaryotes which result in a delay between when expression of a protein is initiated and when the protein becomes functional. A time-delay implies that the system output (TetR and eGFP) reflects the amount of system input (*tetR* and eGFP expression) at a previous time, rather than the input's current state, thus preventing the network system from stabilising. The third design feature, non-linearity, is inherently present in transcription which is dependant on the amount of TetR present, although not linearly so since the relation depends on the number of operator-bound TetR within the cell. In mathematical modelling, this is represented by a Hill function. Similarly to time-delay, this serves to cause instability within the system.

4.1.4 The unstable Goodwin oscillator

The unstable Goodwin oscillator, pUnstable, builds on the stable design and retains all the main characteristics and components. Two main sets of changes were applied, as represented in Fig-



(a) **Unstable Goodwin oscillator**, showing the two main modifications which were applied to the stable oscillator; the fusion of a Myc tag and Ub-L destabilising gene to the *tetR* gene (Edit 1) and the change of the UTRs flanking the Myc-Ub-L-*tetR* moiety to Tubulin 5'UTR and 3'UTR (Edit 2).

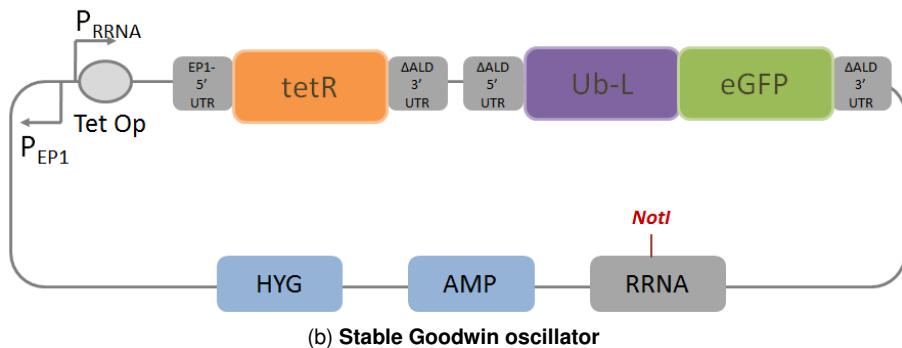


Figure 4.4: Diagram of the plasmids encoding the: a) unstable and b) stable Goodwin oscillator constructs. Edit 1 and Edit 2 highlight the changes applied to the stable device in order to develop the unstable network.

ure 4.4a and as discussed below. These include the insertion of a Myc tag and Ub-L destabilising gene upstream of the *tetR* gene (but downstream of the EP1-5'UTR) to form a moiety, and the replacement of the EP1 5' UTR and Δ ALD 3' UTRs flanking the *tetR* gene with Tubulin UTRs. Oscillations within the unstable Goodwin oscillator occur via the same mechanisms as the stable Goodwin device (Section 4.1.2). The same design principles implemented in the stable Goodwin oscillator (Section 4.1.3) are also implemented here.

4.1.4.1 Modification 1: Addition of a c-Myc tag peptide and modified Ubiquitin gene upstream of the *tetR* gene

In the first modification, a c-Myc tag and Ub-L destabilising gene were fused upstream to the *tetR* gene. The c-Myc tag (Evan et al., 1985) was added in order to allow for an easier characterisation process when quantifying proteins via use of antibodies, to detect for the c-Myc tag instead of the *tetR* protein directly. Although there are other methods for quantifying TetR (Hillen et al., 1982), this was considered to be a more straightforward, tried-and-tested quantification technique.

The Myc epitope was added upstream of the *tetR* gene rather than at the C-terminus since TetR dimerisation occurs at the α 7- α 10 structures (Rossi et al., 1998) which are located in the C-terminus half of a TetR monomer. Fusing the Myc tag at the C-terminus results in the compromising of the TetR structure and a leaky expression as seen in Dragosits et al. (2012). Research has

<i>Ub-L Original</i>	1 CAG ATC TTC GTG AAA ACT CTT ACC GGT AAA ACC ATT GCT CTC
<i>Ub-L Modified</i>	CAG ATT TTT GTG AAG ACT CTT ACG GGT AAG ACG ATC GCC CTC
<i>Ub-L Original</i>	43 GAG GTC GAA GCC AGT GAC ACC ATT GAG AAC GTA AAG GCC
<i>Ub-L Modified</i>	GAG GTA GAG GCT AGT GAC ACG ATT GAA AAT GTA AAA GCT
<i>Ub-L Original</i>	82 AAG ATC CAA GAT AAG GAG GGT ATC CCT CCG GAT CAG CAG
<i>Ub-L Modified</i>	AAA ATA CAG GAC AAA GAA GGC ATC CCT CCC GAC CAG CAG
<i>Ub-L Original</i>	121 CGA CTG ATT TTT GCT GGT AAA CAG TTG GAG GAG GGC CGC
<i>Ub-L Modified</i>	CGC CTG ATT TTC GCC GGA AAA CAG CTC GAG GAA GGC CGT
<i>Ub-L Original</i>	160 ACG CTT GCG GAC TAC AAC ATT CAG AAG GAG TCG ACG CTG
<i>Ub-L Modified</i>	ACT CTC GCC GAT TAT AAC ATA CAG AAA GAA TCT ACA CTG
<i>Ub-L Original</i>	199 CAT CTT GTG CTT CGT CTT CGA GGT GGT TTG
<i>Ub-L Modified</i>	CAT TTG GTA TTA AGA CTT CGC GGG GGT CTT

Figure 4.5: Diagram highlighting the different nucleotides between the original 228bp Ub-L sequence (excluding the start codon) found upstream of eGFP (top row of each line) and the codon optimised Ub-L sequence to be inserted upstream of the *tetR* ORF (bottom row of each line) in the unstable oscillator. The original Ub-L sequence is in black, while the codon optimised Ub-L sequence is in blue, with nucleotides which do not correspond to the original highlighted in red. The numbers on top at the beginning of each line represent the base pair number.

shown that the N-terminus of the TetR (dimer) protein plays a role in the binding of the protein to the tet operator (Berens et al., 1992). Based on this information, it was decided to append the tag at the N-terminus so as not to compromise TetR stability and structure (Berens et al., 1992). The effect of the appendage on the binding properties of TetR will be discussed during characterisation.

The second addendum to the *tetR* gene was the destabilising Ub-L gene in order to destabilise the TetR protein (Bachmair et al., 1986), hence why this GRN is labelled as the unstable oscillator. This was added at the N-terminus of the *tetR* gene, downstream of the Myc tag. The TetR protein was destabilised to increase the frequency of oscillations and thus develop a tuned oscillator plasmid. In order to facilitate identification of the unstable plasmid a HindIII restriction site was inserted between the Ub-L and *tetR* genes.

The Ub-L codon usage was optimised for the *T. brucei* platform. Optimisation was carried out in order to increase translational efficiency of the moiety and reduce the risk of recombination due to the homologous (original) Ub-L sequence which is present upstream of the eGFP gene (Menzella, 2011). Optimisation was carried out using an application by Integrated DNA Technologies, Inc. (Leuven, Belgium, <http://eu.idtdna.com/CodonOpt>). The edited nucleotides were scattered throughout the Ub-L sequence in order to decrease long stretches of homology with the original Ub-L sequence. See Figure 4.5 for a comparison of the original and optimised Ub-L genes. In total, 49 nucleotides were changed, meaning that there is a 21% discrepancy between the original and modified Ub-L sequences. Note that the start codon in the optimised Ub-L sequence was removed, since Ub-L was preceded by the Myc-tag.

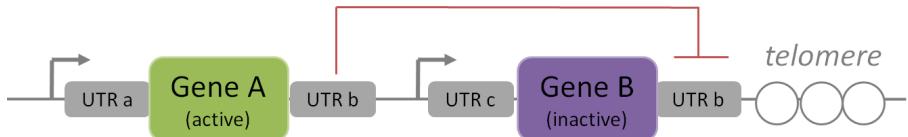


Figure 4.6: Diagram representing the silencing of genes by active genes at sub-telomeric loci via inhibitive cross-talk between homologous 3' UTRs.

4.1.4.2 Modification 2: Change of UTRs flanking *tetR*

In the second main modification applied to the stable oscillator, the UTRs flanking the *tetR* gene were changed from the EP1 5' UTR and Δ ALD 3' UTR to the Tubulin 5' and 3' UTRs, represented as the Tub 5' and 3' UTRs (Figure 4.4a). The Tubulin 5' and 3' UTRs have been shown to allow for higher expression levels than other UTR options (McLatchie et al., 2013). The KpnI sites flanking the previous UTRs were retained, to facilitate plasmid identification during cloning.

The decision to change the UTRs is based on findings on proteins at telomeric or sub-telomeric locations by Hutchinson et al. (2014). A telomere is a repetitive region at the end of a chromatid. The authors found that VSG proteins at telomeric or sub-telomeric loci are inhibited and kept inactive by active proteins using cross-talk between homologous 3' UTRs of the active and inactive proteins, as represented in Figure 4.6. Although the result has not yet been proved for non-telomeric regions, the homologous Δ ALD 3' UTR region downstream of the *tetR* gene in the Goodwin oscillator was edited nonetheless to remove the possibility of inhibiting cross-talk.

4.2 Construction of the stable Goodwin oscillator

Four assembly strategies, each using different recombinant DNA techniques for the construction of the stable oscillator plasmid pStable were developed and taken forward in parallel. A brief outline of the strategies, labelled the Non-directional, Three-fragment, TetR-first and eGFP first strategies, is found in Figure 4.7 which serves to show the main cloning steps and plasmids used in each assembly. This approach was developed as a time-efficient strategy which acknowledges that recombinant DNA assembly strategies are not fail-safe (Perkel, 2014) and with the aim of using the construct first past the post. These assembly strategies will be discussed next.

4.2.1 Outline of the principle cloning steps

The plasmid pRP_{eGFP}SIR2rp3 (Figure 4.8a) was used as a backbone to construct the stable oscillator which has size 8.8kb. All cloning was carried out in *E. coli* cells, so as to take advantage of standard and time-efficient cloning protocols. Each assembly entailed three main steps, as represented in Figure 4.8b and as explained below:

1. Insert the EP1 5' UTR-*tetR*- Δ ALD 3' cassette (1.27kb), extracted from pHD1313, into pRP_{eGFP} SIR2rp3 at the unique KpnI site.

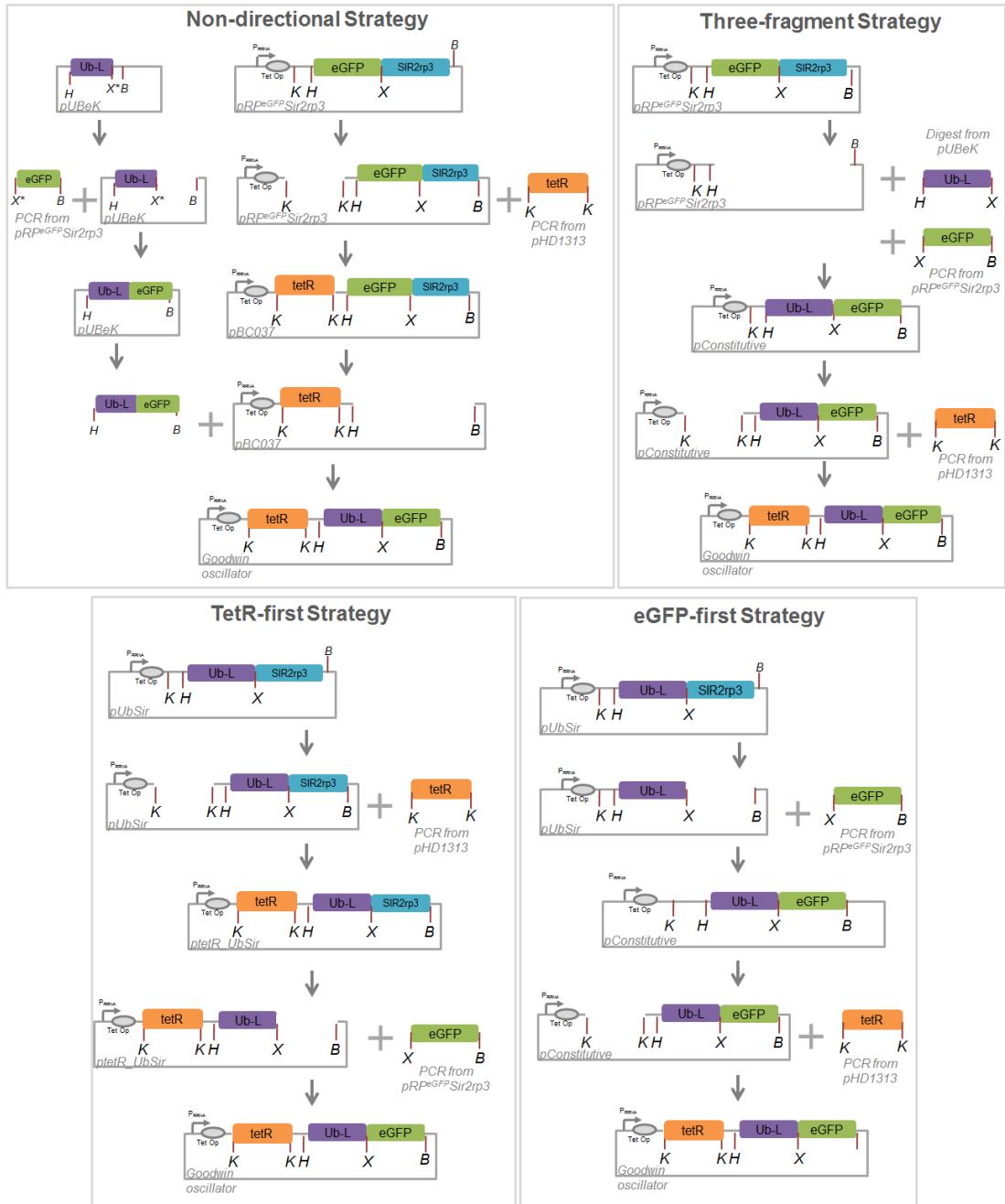
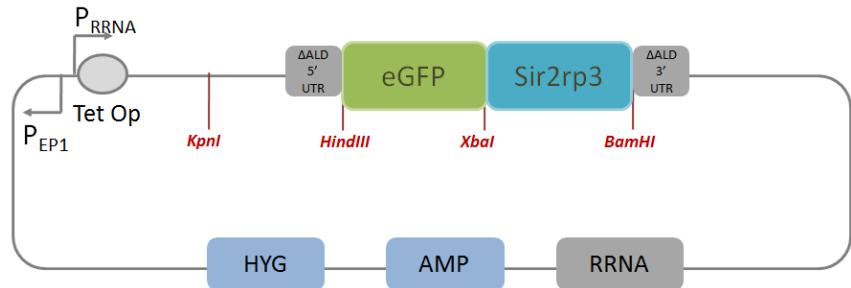
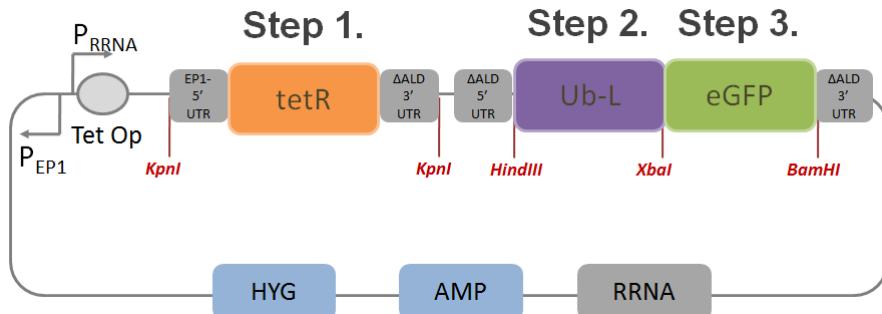


Figure 4.7: Summary of the four Assembly strategies developed to construct the plasmid encoding the Goodwin oscillator. The strategies make use of single and double preparative digests, PCR amplifications, directional, non-directional, three-fragment and Gibson ligation techniques. The strategies make use of the pUBeK, pRP_{eGFP}SIR2rp3, pBC037, pConstitutive, pUbSir and ptetR-UbSir plasmids during the assembly steps. The letters K, H, X, X* and B represents the restriction sites KpnI, HindIII, XbaI and BamHI used in digest and ligation steps.



(a) pRP_{eGFP}SIR2rp3 plasmid having the pRP backbone used as the basis for the assembly of the stable oscillator



(b) The stable oscillator highlighting the three main cloning steps which are required for assembly using pRP as a backbone; the insertion of the UTR-tetR-UTR cassette at the KpnI site (Step 1), the insertion of Ub-L at the HindIII/XbaI sites (Step 2) and the insertion of eGFP at the XbaI/BamHI sites (Step 3).

Figure 4.8: Diagram showing: a) the pRP_{eGFP}SIR2rp3 plasmid having the pRP backbone used as a backbone for the construction of the stable oscillator and b) the stable oscillator device highlighting the three main construction steps required for assembly.

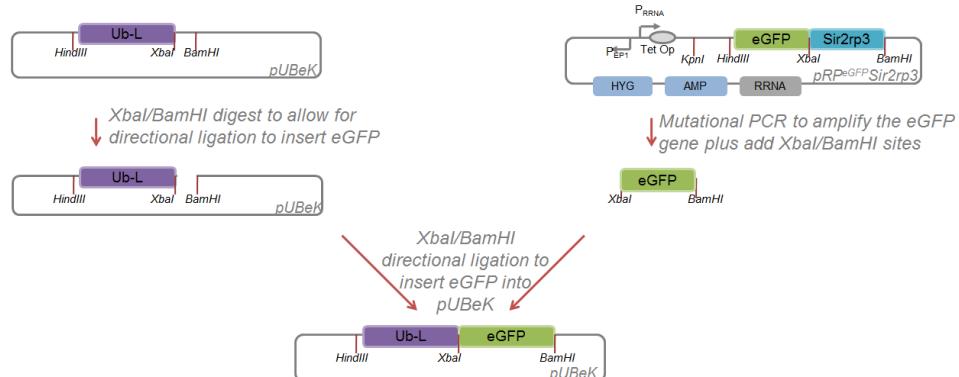
2. Replace eGFP gene site on pRP_{eGFP}SIR2rp3 with the Ub-L gene (0.36kb), extracted from pUBeK.
3. Replace the Sir2rp3 gene on pRP_{eGFP}SIR2rp3 with the eGFP gene (0.72kb) extracted from the same pRP_{eGFP}SIR2rp3 plasmid.

4.2.2 Oscillator assembly strategy using non-directional ligation

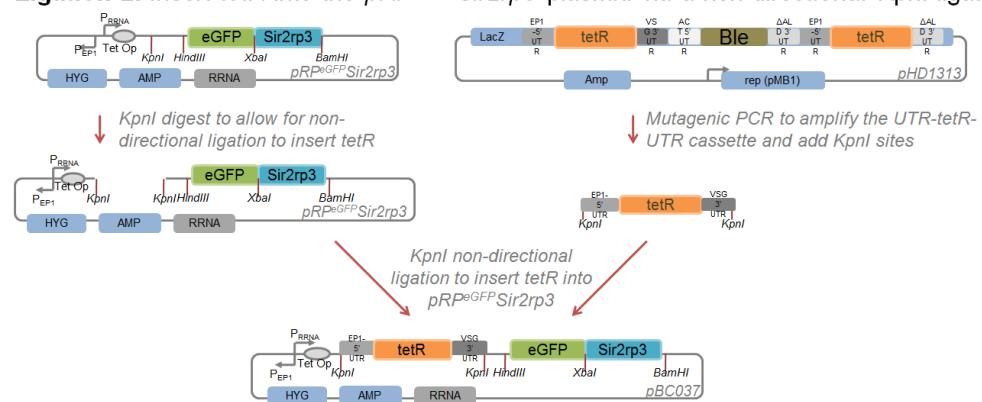
4.2.2.1 Introduction

The first assembly strategy using non-directional ligation is detailed in Figure 4.9. The aim of this strategy was to ligate the eGFP gene into the pUBeK backbone, downstream of the Ub-L gene so as to create the Ub-L-eGFP moiety via an XbaI/BamHI directional ligation (Ligation 1 in Figure 4.9), with the XbaI/BamHI ligation scar designed to be in-frame throughout all ligations. In parallel, the *tetR* gene would be inserted in the pRP backbone of pRP_{eGFP}SIR2rp3 via a KpnI non-directional ligation (Ligation 2). Once completed, the Ub-L-eGFP genes would be ligated into the latter to give the stable oscillator via a HindIII/BamHI ligation (Ligation 3). The diagram omits details such as purification and other treatments which are referred to further on if valid to the discussion.

Ligation 1. XbaI/BamHI ligation of the eGFP insert into the pUBeK backbone to get a Ub-L-eGFP cassette



Ligation 2. Insert *tetR* into the *pRPeGFPSir2rp3* plasmid via a non-directional KpnI ligation



Ligation 3. HindIII/BamHI ligation of the Ub-L-eGFP moiety into pBC037 to get pStable

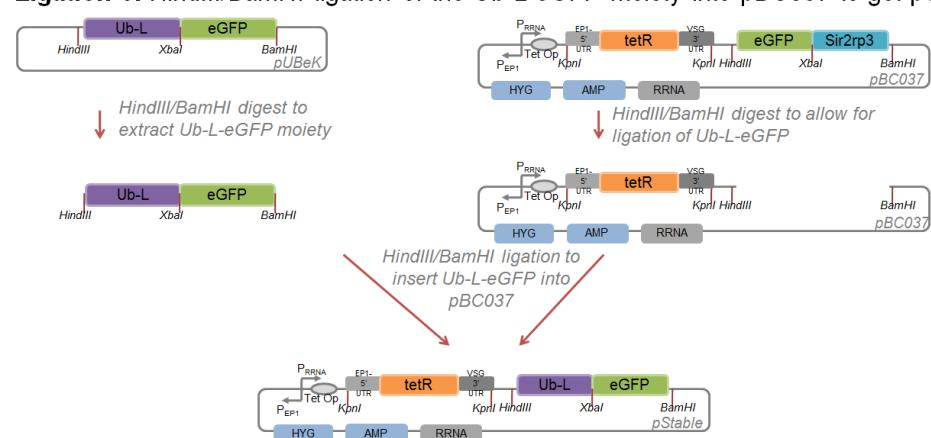


Figure 4.9: Oscillator assembly strategy using non-directional ligation.

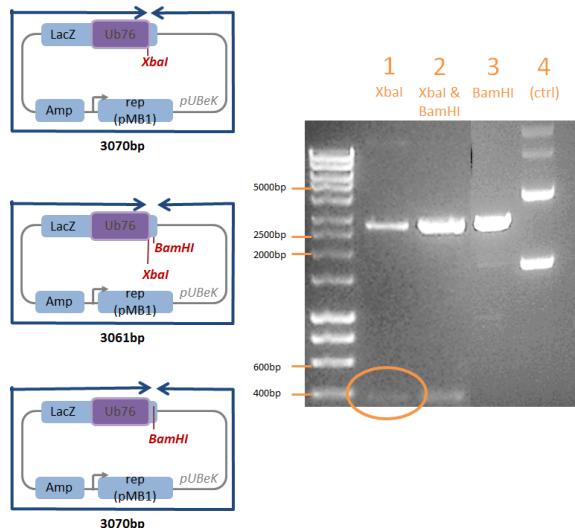


Figure 4.10: Analysis of the XbaI and BamHI restriction sites on pUBeK. Diagrams on the left show the pUBeK plasmid cut with XbaI (top), XbaI & BamHI (middle) and BamHI (bottom). Gel results on the right of XbaI and BamHI single and double digests show:

- 1) 10 μ l pUBeK, XbaI cut,
- 2) 10 μ l pUBeK, XbaI/BamHI cut,
- 3) 10 μ l pUBeK, BamHI cut,
- 4) 5 μ l pUBeK, uncut.

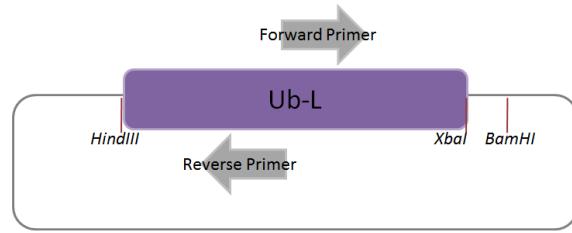
4.2.2.2 Restriction site mapping of the pUBeK backbone

Prior to ligation steps, analytical digests were carried out in order to confirm identity of the pUBeK plasmid. Specifically, digests were carried out to assess whether XbaI and BamHI sites, used during the assembly stages, were unique. Digest results seen in Figure 4.10 indicated that the sites were not unique, as seen in Lane 1 which shows two bands for a single XbaI digest.

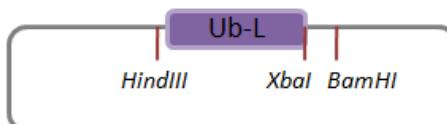
Following this, sequencing was carried out using the forward primer (5'-3') ATAAGGAGGGTATC-CCTCCGG and reverse primer (5'-3') AATCAGTCGCTGCTGATCCGG. The primers annealed to the middle of the Ub-L gene (Figure 4.11a), such that the gene itself and its surrounding region were sequenced. Interest lay in sequencing the nucleotides immediately upstream and downstream of the gene which might have been inserted during construction of pUBeK. As a backbone, the commercially available pUC57 backbone was used and hence, its sequence was easily available. Results in Figures 4.11b and 4.11c show a diagram of the known restriction sites prior to and after sequencing. In addition to the known XbaI site downstream of Ub-L, there is a second XbaI site upstream of the gene which was previously undetected and which is in agreement with analytical digest results in Figure 4.10. This compromised the ligation strategy wherein the eGFP gene was to be inserted into the pUBeK plasmid via an XbaI/BamHI ligation (Ligation 1 in Figure 4.9). This is because if the backbone is cut with XbaI and BamHI in preparation of the ligation, the Ub-L gene is cut out of the backbone in the process. Thus an alternative strategy was adopted.

4.2.2.3 Preparation of eGFP for insertion downstream of the Ubiquitin degron region

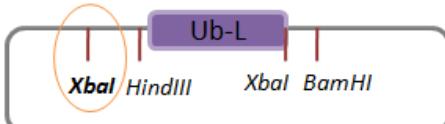
Apart from the digestion of pUBeK, the first step of the assembly strategy required the preparation of the eGFP insert by PCR amplifying the gene from pRP_{eGFP}SIR2rp3 using the forward primer (5'-



(a) Diagram showing the pUBeK plasmid and the primers used to map restriction sites immediately upstream and downstream of the Ub-L gene.



(b) Known restriction sites prior to sequencing.



(c) Known restriction sites post sequencing.

Figure 4.11: Diagram showing: a) Primers used in the sequencing of pUBeK plasmid to map region surrounding the Ub-L gene, b) Known restriction sites surrounding Ub-L gene prior to sequencing, c) Known restriction sites surrounding Ub-L gene post sequencing.

3') TCTAGAGTGAGCAAGGGCGAGGAGCTGTTCACCGGGG to add the XbaI site and remove the start codon and reverse primer (5'-3') GGATCCGCCTTCAAGACTTGTACAGCTCGTCC to add the BamHI restriction site and stop codon. This results in a cassette of size c. 730bp which includes the eGFP gene plus flanking XbaI and BamHI sites to be used in ligation.

Results of PCR reactions showed no amplified DNA regions (not shown). The first step in troubleshooting this issue was to repeat the PCR as per supplier instructions using different recipes wherein four different variables were tested for optimisation:

Addition of dimethyl sulfoxide (DMSO) - As per manual instructions, to a final concentration of 3% which serves to optimise template denaturation by removing DNA supercoils.

Addition of 0.5mM magnesium chloride (MgCl₂) - Two different concentrations were tested, 0.5mM MgCl₂ and 1mM MgCl₂, to act as a cofactor to the magnesium-dependant Phusion polymerase.

Annealing temperature - A 45°C instead of a 55°C annealing temperature was tested, in order to facilitate binding.

Buffer - The use of GC instead of the standard HF buffer. GC buffer can be used in GC-rich templates. Four different experiments were carried out, two at an annealing temperature of 55°C using either no additional MgCl₂ or an additional 0.5mM MgCl₂, and two reactions using an annealing temperature of 45°C following the same recipes.

Throughout, positive controls always showed a clear band, indicating that the PCR reagents were working and the thermocycling process allowed for PCR (images not shown). However, no bands were observed when a sample of the PCR on pRP_{eGFP}SIR2rp3 was run on a gel. Therefore, three complementary primers were designed: a forward primer (5'-3') GAGTGGTACCTGCACGC-CGC, a second forward primer (5'-3') ATGGTGAGCAAGGGCGAGGAG and a reverse primer (5'-3') AATCATACGATTGGGCCGCC.

These primers served to amplify cassettes of size 745bp or 844bp which include the eGFP cas-

sette. Each of the forward complementary primers were tested and both were successful, as seen in Figure A.1a in Appendix A which contains images of gel electrophoresis results run with PCR, ligation and digest DNA samples. Subsequent amplifications of these strands using the original mutagenic primers were then applied to amplify the eGFP gene and add the necessary overhangs (See Figure A.1b). Bands had a size of c. 730bp, as expected.

Thus, amplification and addition of XbaI/BamHI restriction sites to eGFP was achieved in a series of two steps, the first using complementary primers to extract a DNA fragment including the eGFP gene which then serves to minimise probability of non-specific annealing when the second PCR using mutagenic primers to introduce the necessary nucleotides was applied. This ligation step was suspended due to the non-unique restriction sites present in pUBeK as discussed above.

4.2.2.4 KpnI non-directional ligation of *tetR* into pRP_{eGFP}SIR2rp3

In parallel to the above, Ligation 2 of the assembly was taken forward.

Preparation of *tetR* insert. PCR was carried out using the forward mutagenic primer (5'-3') GGTACCGTCTTGGTGTGTCGACCTTGCAGGC and reverse mutagenic primer (5'-3') GGTACCTTGTACATATTGTCGTTAGAACGCGGC to extract the EP1 5'UTR-*tetR*-ΔALD 3'UTR cassette from the pHD1313 plasmid, add KpnI sites upstream and downstream of the cassette and prepare it for ligation into the pRP_{eGFP}SIR2rp3 plasmid backbone. This PCR product will be referred to as the *tetR* cassette and is expected to be 1241bp long. No optimisation measures for PCR were required, with gel electrophoresis results in Figure A.2a. Amplified DNA was purified and identity again confirmed via gel electrophoresis (not shown).

In preparation of the ligation, a preparative KpnI digest was carried out on the *tetR* cassette. Two digests were carried out in parallel, one using the digestive enzyme and another where in addition to it, CIP-treatment was applied to dephosphorylate 5' ends of the insert and minimise chances of re-ligation. Following plasmid identification via gel electrophoresis, the DNA was gel purified and a final sample taken to confirm identity via gel electrophoresis (results not shown).

Preparation of the pRP_{eGFP}SIR2rp3 backbone. The pRP_{eGFP}SIR2rp3 backbone was taken through a KpnI digest and dephosphorylation reactions, as with *tetR* (Figure A.2b). Gel-purification steps were swapped for purification using the QIAquick PCR Purification Kit (Qiagen) for a more efficient process.

Ligation of fragment encoding TetR ORF into pRP_{eGFP}SIR2rp3 backbone. Following preparation of the *tetR* insert and pRP_{eGFP}SIR2rp3 backbone, ligation reactions were carried out (see Table A.1, which also lists colony counts observed on transformation plates).

Transformation of the ligations was carried out in both W3110 and commercial TOP10 cells. In the TOP10 cells growth was noticed on all plates, however, this was limited to small non-isolated

colonies. In the W3110 cells, on the other hand, a small number of isolated large colonies were noticed surrounded by small non-isolated colonies indicating suboptimal growth. When comparing the ligations using CIP-treated inserts to those using non-CIP-treated inserts, no clear distinction was made, indicating that CIP was not having a discernible effect. Controls also showed that the digest was suboptimal, since no difference was observed between reactions using cut plasmid wherein ligase was added or omitted. Two colonies were picked from the ligation plates and a sample taken for a KpnI analytical digest and run on a gel (image not shown). Bands showed that no insert was present.

4.2.2.5 Appraisal of non-directional ligation strategy

Following the implementation of this assembly strategy using non-directional ligation to construct the pStable oscillator, the following findings and observations were made:

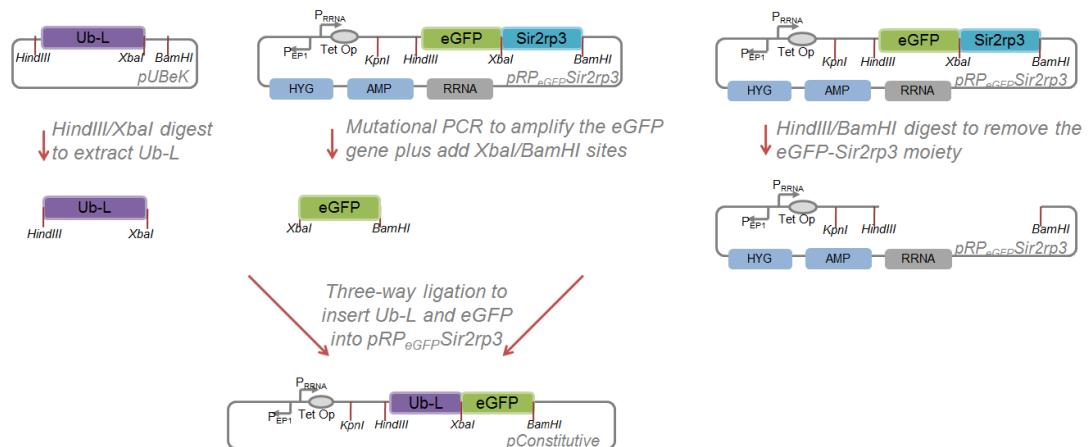
1. The amplification of eGFP from the pRP_{eGFP}SIR2rp3 backbone requires a two-step PCR. The first PCR uses complementary primers to extract a DNA segment including the eGFP gene from the backbone. This reduces the probability of the second set of primers mis-annealing along the backbone. Mutagenic primers are then applied to the segment to extract the eGFP gene and add restriction sites, remove the start codon and add the stop codon.
2. Digests and sequencing to map restriction sites of the pUBeK backbone revealed unknown restriction sites which prevent the use of the pUBeK backbone for cloning purposes to insert the eGFP gene downstream of the Ub-L gene and develop the Ub-L-eGFP moiety.
3. Results of the non-directional KpnI ligation to insert *tetR* into the pRP_{eGFP}SIR2rp3 backbone indicated that the backbone had re-ligated to itself. Additionally, the colonies observed on transformation plates on which chemically transformed W3110 or TOP10 cells were spread, were small and non-isolated. Thus, alternative chemically competent XL10-Gold® cells available commercially will be used.

4.2.3 Oscillator assembly strategy using three-fragment ligation

4.2.3.1 Introduction

The second cloning strategy, as outlined in Figure 4.12, was adopted as an alternative approach of fusing the eGFP gene to Ub-L on the pUBeK backbone via an XbaI/BamHI ligation given that the XbaI site on pUBeK is not unique. The first step (Ligation 1 in Figure 4.12) was to insert the Ub-L and eGFP genes into the pRP backbone via a three-fragment HindIII/XbaI/BamHI ligation to get the pConstitutive plasmid. This avoids the use of pUBeK as a backbone. The pConstitutive plasmid has the Ub-L-eGFP moiety but not the regulating *tetR* gene, resulting in constitutive eGFP expression. Following this, the *tetR* cassette was to be inserted via a non-directional KpnI ligation in pConstitutive to construct pStable (Ligation 2 in Figure 4.12). The diagram leaves out details such as purification and other treatments which are referred to further on if valid to the discussion.

Ligation 1. Insert the Ub-L and eGFP genes into the the *pRP_{eGFP}Sir2rp3* plasmid instead of the eGFP-Sir2rp3 moiety via a three-way HinIII/XbaI/BamHI ligation to get pConstitutive



Ligation 2. Insert *tetR* cassette into *pConstitutive* via a non-directional *KpnI* ligation to get *pStable*

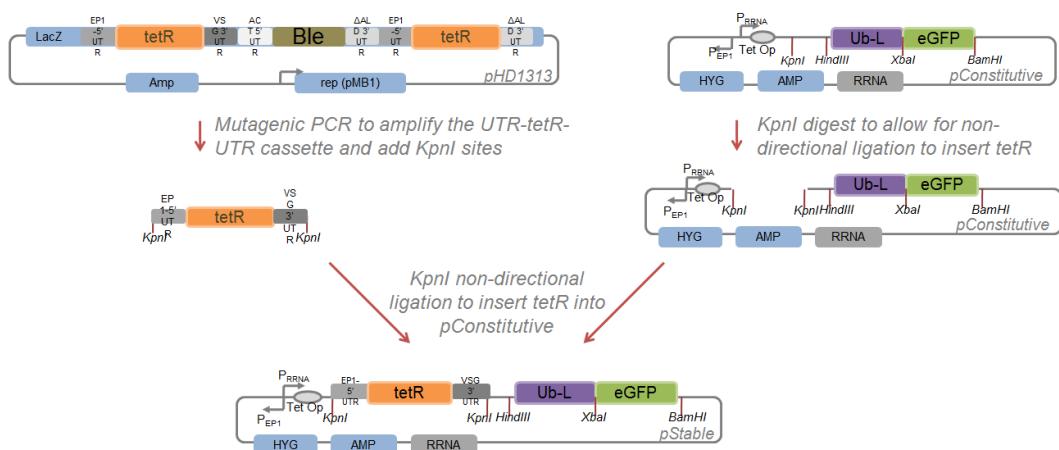


Figure 4.12: Assembly strategy using three-fragment ligation.

4.2.3.2 Ligation 1: Assembly of pConstitutive plasmid by ligation of Ub-L and eGFP fragments into pRP_{eGFP}SIR2rp3 backbone

HindIII/BamHI digest of pRP_{eGFP}SIR2rp3 backbone. The first step was to carry out a HindIII/BamHI preparative digest on pRP_{eGFP}SIR2rp3, after which the reaction was run on gel (Figure A.3a). The backbone was extracted via gel purification and an aliquot run on a gel (not shown) to confirm DNA is still intact.

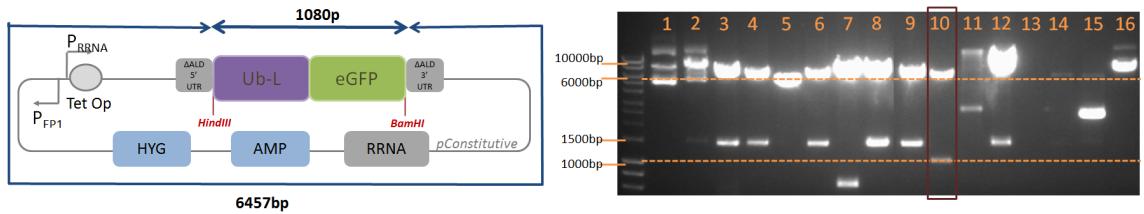
XbaI/BamHI digest of eGFP insert. To amplify the eGFP gene from pRP_{eGFP}SIR2rp3, the two-step PCR process using complementary primers followed by mutagenic primers as outlined in the previous strategy was used (Figure A.3b). Following this, the DNA was purified, digested with XbaI and BamHI and purified. A sample was run on a gel to confirm plasmid identity (not shown).

HindIII/XbaI digest of Ub-L insert. In order to extract the Ub-L gene from pUBeK, an XbaI/HindIII digest was carried out. A sequential digest was carried out instead of a simultaneous digest as an optimisation measure. The digest was run on a gel (Figure A.3c) which showed the required 0.4kb Ub-L band and the discarded 2.7kb backbone band, after which it was gel purified and a small aliquot run on a second gel to confirm plasmid remained intact throughout purification (not shown).

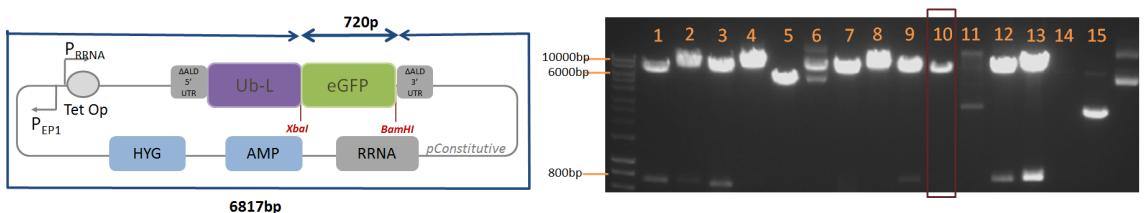
Three-fragment ligation. Following preparation of inserts and backbone, ligations were prepared (data not shown) and transformed using XL10-Gold® cells. Colony counts from control plates were indicative of uncut backbone and a suboptimal purification process which did not eliminate unwanted DNA cassettes. Moreover, colonies which were picked from ligation plates and analysed indicated that only re-ligated backbone plasmid was present.

A second three-fragment ligation was carried out using a larger plasmid mass of better quality, eluted via a maxiprep protocol rather than a miniprep kit. DNA stocks were at least twice as dense as stocks eluted from minipreps. In all subsequent steps, DNA eluted from a maxiprep will be used unless otherwise specified. The inserts and backbone were prepared as per the previous methods (images from gels not shown) and ligations were carried out (Table A.2).

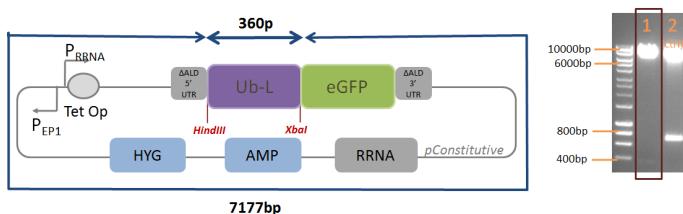
Control reactions were indicative of a successful digest but a suboptimal purification of the backbone following digestion. Colonies were picked from the ligation plates, grown for 12-16 hours in LB medium, purified and an aliquot taken for HindIII/XbaI, HindIII/BamHI and XbaI/BamHI analytical digests to distinguish between the inserts. Results can be seen in Figure 4.13. Of the clones picked and digested using both HindIII and BamHI (Figure 4.13a), only one clone (clone 10) shows the required bands; one at 1.1kb representing the Ub-L-eGFP fused insert and one at 6.5kb representing the pRP_{eGFP}SIR2rp3 backbone. When the same clones were run through an XbaI and BamHI digest (Figure 4.13b), none of the clones show the required bands at 0.72kb for the eGFP insert and 6.8kb for the backbone. The final analytical HindIII and XbaI digest (Figure 4.13c) was only carried out on clone 10, using the pRP_{eGFP}SIR2rp3 backbone as a control. The resultant bands are as expected showing the 0.36kb Ub-L insert and the remaining 7.2kb backbone.



(a) **HindIII/BamHI analytical double digest:** Diagram on the left shows pConstitutive cut with HindIII/BamHI. The right panel shows digest results. Lanes show $10\mu\text{l}$ aliquots of 16 different clones analysed via HindIII/BamHI digest.



(b) **XbaI/BamHI analytical double digest:** Diagram on the left shows pConstitutive cut with XbaI/BamHI. The right panel shows digest results. Lanes show $10\mu\text{l}$ aliquots of 16 different clones analysed via XbaI/BamHI digest.



(c) **HindIII/XbaI analytical double digest:** Diagram on the left shows pConstitutive cut with HindIII/XbaI. The right panel shows digest results.

Lanes:

- 1) $10\mu\text{l}$ from clone 10 seen in above images, HindIII/XbaI cut,
- 2) $10\mu\text{l}$ pRP_{eGFP}SIR2rp3, uncut.

Figure 4.13: Assembly strategy using three-fragment ligation - Assessing success of ligation: Analytical double digests using: a) HindIII/BamHI, b) XbaI/BamHI and c) HindIII/XbaI on clones picked from ligation plates to assess success of the second HindIII/XbaI/BamHI three-fragment ligation attempt to insert Ub-L and eGFP in pRP_{eGFP}SIR2rp3 to get pConstitutive.

Clone 10 was sequenced using primers from Table C.1 in Appendix C. Results showed that whereas Ub-L was successfully inserted into the pRP backbone, the eGFP gene was not inserted and instead the Sir2rp3 gene remained present. The resultant plasmid was named pUbSir and used in subsequent assembly strategies. See Figure 4.7 for reference.

XbaI/BamHI ligation of eGFP into pRP_{eGFP}SIR2rp3. As an alternative to the three-fragment ligation, both inserts could be inserted via a series of standard ligations; the first to insert Ub-L via a HindIII/XbaI ligation and the second using an XbaI/BamHI ligation to insert eGFP, or vice-versa. For the insertion of eGFP into the pRP backbone, DNA was prepared as per above. Gel results of the prepared material are not shown. Once the DNA was prepared, ligations were carried out (Table A.3).

Controls again indicated that the digest worked as expected but showed a suboptimal purification

process which did not successfully remove all unwanted Sir2rp3 DNA. Colonies were picked from the ligation plates and analysed via gel electrophoresis following an Apal/HindIII analytical digest. The Apal enzyme was chosen because the Apal site is present in the Sir2rp3 cassette but not the eGFP one. Thus, different bands will be observed depending on whether the eGFP or Sir2rp3 is present. Bands corresponded to the pRP backbone having the Sir insert rather than eGFP, similarly to previous cases, indicating that the issue lies with the digestion reactions. Thus, the Gibson ligation method was adopted as an alternative assembly method, as will be discussed in subsequent assembly strategies.

4.2.3.3 Ligation 2: KpnI non-directional ligation of the *tetR* cassette into pRP_{eGFP}SIR2rp3

In parallel to the above ligations, a non-directional ligation for the insertion of the *tetR* cassette into pRP_{eGFP}SIR2rp3 was carried out. The backbone and insert were prepared as per Section 4.2.2.4 but using higher quality backbone DNA eluted via maxi-preps. Moreover, the digest and CIP-treatment of the *tetR* gene were carried out in separate reactions so as to maximise efficiency as previous results has shown both reactions to be inefficient when carried out simultaneously. Gel results of the prepared material are not shown.

Following ligation (Table A.4), transformations were carried out using XL10-Gold® cells. Colony counts are indicative of an inefficient digest and CIP-treatment. Colonies were picked for further analysis, grown in selective LB medium for 12-16 hours, purified and a KpnI analytical digest carried out on eluted DNA. The digest was then run on a gel (not shown). All colonies showed re-ligated pRP_{eGFP}SIR2rp3 backbone, once again highlighting the challenge of performing a non-directional ligation.

The first step was to test the effectiveness of different dephosphorylation enzymes on the pUbSir backbone. Thus, instead of the previously used CIP, Calf Intestinal Alkaline Phosphatase (CIAP, Promega) and Shrimp Antarctic Phosphatase (SAP, NEB) were used. Three volumes of 10 μ g of pUbSir backbone were taken through three different preparation steps respectively as listed in Table 4.1. Ligation reactions using only digested pUbSir backbone were then carried out as per Table 4.2. No insert DNA was used, as the aim of this exercise was to assess whether the dephosphorylation enzymes could prevent re-ligation. If successful, colony counts should be 0. Non-dephosphorylated pUbSir backbone was used as a positive control in Ligation 7.

	Dephosphorylation enzyme		
	Reaction 1 CIAP (Promega)	Reaction 2 SAP (NEB)	Reaction 3 -
Steps			
1	KpnI digest	KpnI digest & SAP	KpnI digest
2	PCR purification	PCR purification	PCR purification
3	CIP-treatment (Promega)	-	-
4	PCR purification	-	-

Table 4.1: Assembly strategy using three-fragment ligation - Dephosphorylation: Steps used in the preparation of pUbSir backbone for non-directional ligation by making use of different dephosphorylation enzymes. In the first reaction, Calf Intestinal Alkaline Phosphatase (CIAP) by Promega is used. In the second reaction, Shrimp Antarctic Phosphatase (SAP) by NEB is used, while the third reaction is a control.

		Ligation							
		1	2	3	4	5	6	7	
Ingredient		Amount	Amount	Amount	Amount	Amount	Amount	Amount	Amount
1	pUbSir backbone CIAP (Promega) dephosph., lin., (20ng/ μ l)	2.5 μ l (7mols) (50ng)	-	-	2.5 μ l (7mols) (50ng)	-	-	-	-
2	pUbSir backbone SAP (NEB) dephosph., lin., pur. (25ng/ μ l)	-	2 μ l (7mols) (50ng)	-	-	(7mols) (50ng)	2 μ l	-	-
3	pUbSir backbone Lin., pur. (25ng/ μ l)	-	-	2 μ l (7mols) (50ng)	-	-	(7mols) (50ng)	2 μ l	-
4	pUbSir backbone Uncut (25ng/ μ l)	-	-	-	-	-	-	(7mols) (50ng)	2 μ l
5	Quick T4 DNA Ligase	1 μ l	1 μ l	1 μ l	-	-	-	-	1 μ l
6	2X Ligase Buffer	10 μ l	10 μ l	10 μ l	10 μ l	10 μ l	10 μ l	10 μ l	10 μ l
7	dH ₂ O	to 21 μ l	to 21 μ l	to 21 μ l	to 21 μ l	to 21 μ l	to 21 μ l	to 21 μ l	to 21 μ l
	Total	21 μ l	21 μ l	21 μ l	21 μ l	21 μ l	21 μ l	21 μ l	21 μ l
	Colony Count	30	50	25	0	30	0	0	+250

Table 4.2: Assembly strategy using three-fragment ligation: Recipes for ligation reactions and controls used in study to assess efficiency of different dephosphorylation enzymes to prepare the pUbSir backbone for non-directional ligation. Each ligation lists the cut purified dephosphorylated backbone or uncut backbone in volume, mols and mass, the volume of Quick T4 DNA ligase (optional), 2X Ligase Buffer and dH₂O (optional) used for ligation. Colony counts are from transformation plates following incubation times. 'dephos.' stands for purified plasmids, 'lin.' stands for linearised plasmids.

Colonies were observed on all three transformation plates for Ligations 1-3 in spite of the application of the dephosphorylation enzymes to the digested backbone. This indicates that the tested dephosphorylation enzymes are not removing phosphate groups of DNA ends effectively. Colony counts from transformation plates corresponding to Ligation 6 (ligation carried out using cut backbone, no ligase) showed that the KpnI digest was successful when carried out separately from the dephosphorylation reaction, with no growth observed on transformation plates, but did not work as well when applied in conjunction with the dephosphorylation enzyme, as seen in the other ligations.

4.2.3.4 Appraisal of three-fragment ligation strategy

The second assembly strategy to construct the pStable device featured the use of a three-fragment ligation to insert Ub-L and eGFP into the pRP backbone and the repeat of a non-directional ligation to insert the *tetR* cassette. Following implementation, the following developments and observations were made:

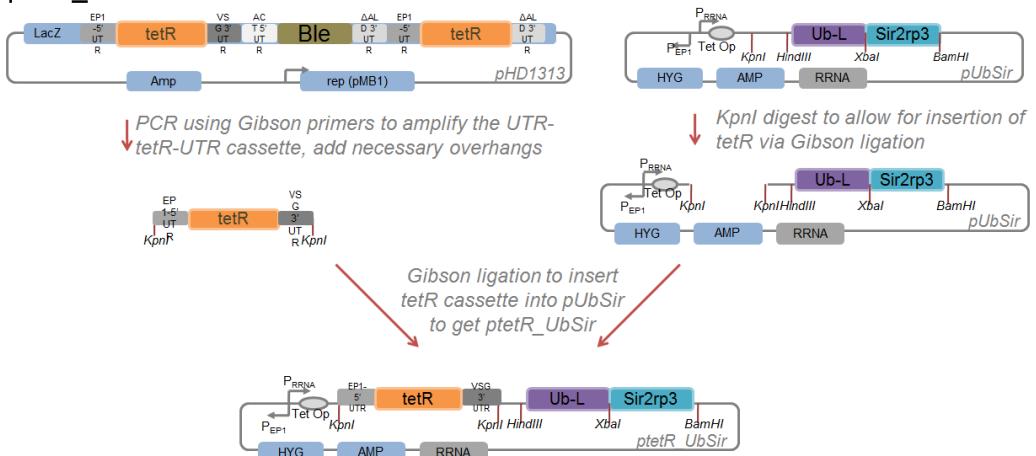
1. The pUbSir plasmid, in which ubiquitin-tagged Sir2rp3 protein is constitutively expressed, was developed as a by-product of a three-fragment ligation implemented to insert the Ub-L and eGFP genes into the pRP backbone. The plasmid will be used as a backbone in subsequent cloning steps. The three-fragment ligation was twice attempted, using a larger DNA mass of better quality the second time. In both cases, the Sir2rp3 gene was re-ligated into the pRP backbone, instead of the eGFP gene. A Gibson assembly will be used in the next strategy to address this.
2. In parallel to the three-fragment ligation, a non-directional KpnI ligation to insert the *tetR* cassette into the pRP backbone was performed. In spite of using DNA material of higher quality, modified protocols and different dephosphorylation enzymes, results repeatedly showed dephosphorylation to be an ineffective process, leading to the re-ligation of the pRP backbone. As with above, a Gibson assembly will be implemented to bypass the use of a non-directional ligation.
3. In both sets of ligations, DNA eluted from a maxiprep kit instead of a miniprep kit was used since it is of better quality. Moreover, preparative double digests were prepared sequentially rather than simultaneously to allow for a more effective digest.

4.2.4 Oscillator assembly strategy inserting the *tetR* cassette followed by the eGFP gene via Gibson assembly

4.2.4.1 Introduction

In the third assembly strategy, the pUbSir plasmid, constructed during the previous assembly strategy, replaced the use of pRP_{eGFP}SIR2rp3 as a backbone. The second main feature of this strategy is the incorporation of the Gibson ligation method (Gibson et al., 2009) as an alternative to standard and non-directional ligation techniques. The Gibson ligation allows for the addition of specific overhangs to the insert, which overhangs correspond to the backbone of interest, automatically making the ligation directional. This bypasses the requirement for CIP-treatment and digestion of the insert, which results in less preparation steps and hence, a smaller loss of DNA

Ligation 1. Insert the *tetR* cassette into pUbSir backbone via a Gibson ligation to get ptetR_UbSir



Ligation 2. Insert the eGFP gene into ptetR_UbSir via a Gibson ligation to get pStable

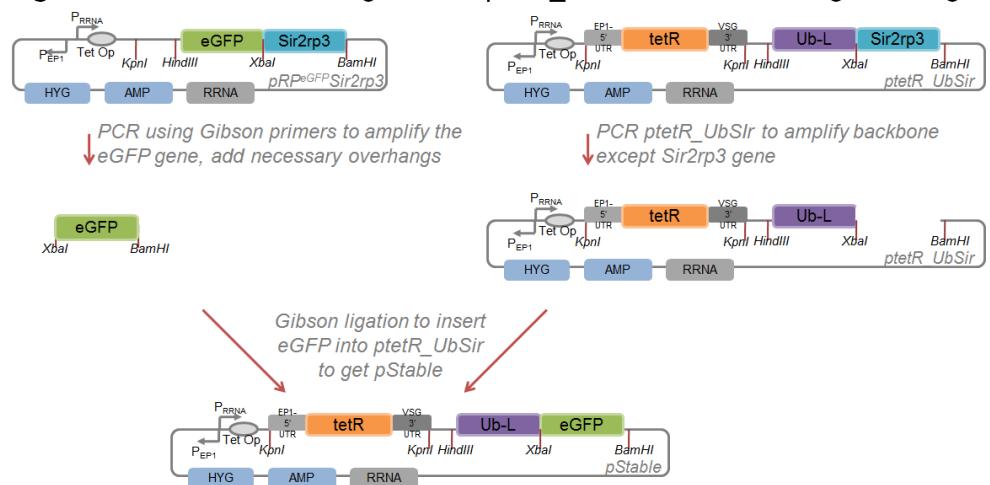


Figure 4.14: Assembly strategy using Gibson assembly to insert the *tetR* cassette followed by the eGFP gene into pUbSir.

material during preparative and purification steps. The backbone, on the other hand, can be prepared via either PCR amplification or preparative digest. In addition, the Gibson cloning technique requires fewer ligation controls. For example, there is no need to control for the CIP-treatment. Once again, this makes for a simpler cloning process.

The main cloning steps of the strategy are represented in Figure 4.14. The first step was to insert the *tetR* cassette into the pUbSir backbone via a Gibson ligation in order to get the ptetR_UbSir plasmid. The insert was amplified using Gibson primers while the backbone was digested at the unique *Kpn*I site. Note that the insertion site was still to be made at the *Kpn*I site on the backbone, similarly to the previous strategies. The second step was to amplify the eGFP insert and ptetR_UbSir backbone, wherein the insert was amplified using Gibson primers which attached complementary overhangs to ptetR_UbSir while the backbone was amplified minus the *Sir2rp3* gene. Following this, the eGFP gene was to be inserted into the ptetR_UbSir plasmid between the *Xba*I and *Bam*HI sites, again via Gibson ligation, in order to get the final pStable construct. The diagram leaves out details such as purification and other treatments which are referred to further on if valid to the discussion.

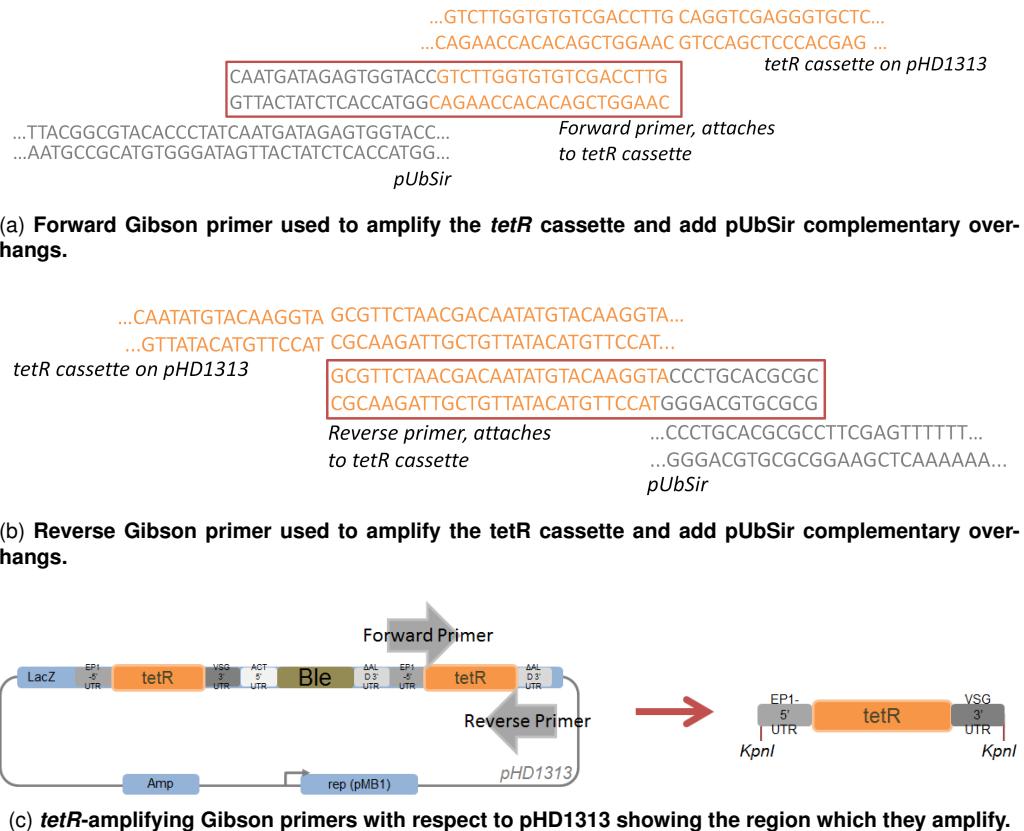


Figure 4.15: TetR-first assembly strategy: Representation of Gibson primers used for the amplification of the *tetR* cassette to be inserted into the pUbSir backbone: a) The upstream Gibson primer (middle) in relation to the *tetR* cassette on pHd1313 (top) and pUbSir backbone (bottom), b) The downstream Gibson primer (middle) in relation to the *tetR* cassette on pHd1313 (top) and pUbSir backbone (bottom). The primers are outlined in red and attach to and amplify the *tetR* cassette (orange) while adding overhangs corresponding to the pUbSir backbone (grey) where the cassette will be inserted via Gibson assembly. c) shows the Gibson primers with respect to pHd1313 indicating the region of amplification.

4.2.4.2 Ligation 1: Gibson assembly of ptetR_UbSir by insertion of the *tetR* cassette into pUbSir

Design of Gibson primers. The main preparation step in the Gibson assembly technique was the design of plasmids which attach overhangs to the *tetR* insert, which overhangs correspond to DNA sequences upstream and downstream of the insertion site on the pUbSir backbone as represented in Figure 4.15.

Preparation of DNA and ligation process. The *tetR* insert was prepared via PCR amplification using the forward Gibson primer (5'-3') CAATGATAGAGTGGTACCGTCTTGGTGTGTC-GACCTTG and reverse Gibson primer (5'-3') GCGCGTGCAGGGTACCTTGACATATTGTCGT-TAGAACGCC. The pUbSir backbone, on the other hand, was prepared by a KpnI digest. Once the preparative digest and amplification were completed, aliquots were run on a gel to confirm DNA identity (Figure A.4a shows the 7.5kb pUbSir backbone linearised at the KpnI site and Figure A.4b shows the amplified 1.2kb *tetR* cassette). Afterwards, PCR purification was carried out and quality was confirmed via gel electrophoresis (not shown).

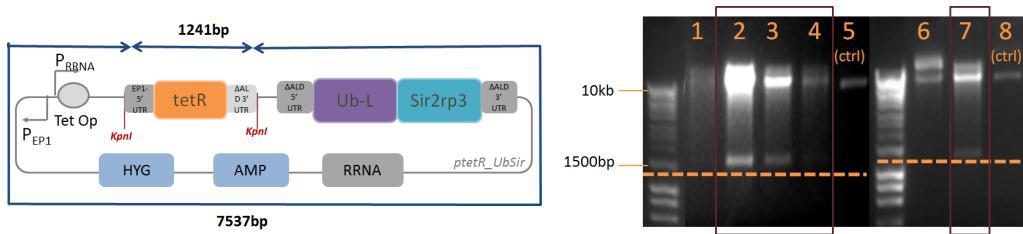


Figure 4.16: TetR-first assembly strategy - Assessing success of ligation: Diagram on the left shows p_{tetR}_UbSir cut with KpnI. The right panel shows gel results for digest of DNA eluted from clones picked from ligation plates following the Gibson ligation to insert *tetR* into pUbSir. Lanes: 1-4) 10 μ l clones, KpnI cut, 5) 5 μ l pUbSir, uncut, 6-7) 10 μ l clones, KpnI cut, 8) 5 μ l pUbSir, uncut.

The Gibson assembly and ligation controls were carried out (Table A.5), after which reactions were transformed into XL10-Gold® cells. Colony counts showed that the positive control supplied by the NEB supplier did not work, although this was not considered to be an issue. Colony counts for negative controls indicated that while the Master Mix was working, the digest had not been efficient. Six colonies were picked from the transformation plates for further analysis, grown in selective LB medium for 12-16 hours, the DNA purified and aliquots taken for a KpnI analytical digest. Each of the digests was then run on a gel, with results as observed in Figure 4.16. Of the six colonies analysed, those in wells 2, 3, 4 and 7 showed the required bands which represent the 1.2kb *tetR* insert and 7.5kb backbone. Glycerol stocks were then prepared and samples sent for sequencing. The presence of the *tetR* gene was confirmed in all the clones, although only one had no mutations present. This was labelled the p_{tetR}_UbSir plasmid and taken forward for DNA purification using a maxiprep.

4.2.4.3 Ligation 2: Gibson assembly of pStable plasmid by insertion of eGFP into p_{tetR}-UbSir backbone

In order to construct the Goodwin oscillator, the insertion of the eGFP gene into the p_{tetR}_UbSir backbone was required as a final step. The Gibson assembly technique was used as an alternative to standard ligation since this had previously led to the re-ligation of the Sir gene. In this Gibson assembly, both the backbone and insert were amplified. During the amplification process of the insert, overhangs corresponding to the p_{tetR}_UbSir backbone were added as per Gibson assembly requirements. The forward Gibson primer (5'-3') CTAGACAAGTTCTAGAG-GTGAGCAAGGGCGACGAG and reverse Gibson primer (5'-3') GCCAACTAAATGGGCAGGATC-CGCCTTCAAGACTTG were used.

The backbone was amplified rather than XbaI/BamHI digested, as the restriction sites were not unique due to the presence of the *tetR* gene. A digest using XbaI and BamHI enzymes would have led to the fragmentation of the p_{tetR}_UbSir plasmid rather than the required linearisation. The forward primer (5'-3') TGCCCATTAGTTGGCTT and reverse primer (5'-3') CTCTAGAAACTTGTC-TAGCCAATTG were used.

Whereas amplification of the eGFP insert did not require optimisation (Figure A.5a), PCR conditions for p_{tetR}_UbSir required optimisation as per the following:

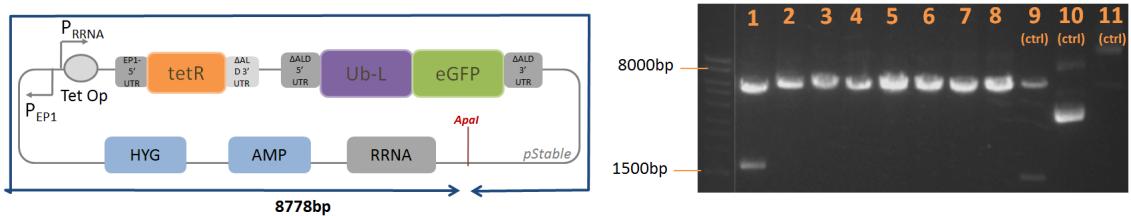


Figure 4.17: TetR-first assembly strategy - Assessing success of ligation: Diagram on the left shows pStable cut with Apal to assess success of Gibson ligation to insert eGFP into ptetR-UbSir. The right panel shows gel results of the digest. Lanes:

- 1-8) 10 μ l clone, Apal cut,
- 9) 10 μ l ptetR_UbSir, Apal cut,
- 10) 5 μ l clone, uncut,
- 11) 5 μ l ptetR_UbSir, uncut.

Addition of 0.5mM MgCl₂ to act as a cofactor to the magnesium-dependant Phusion polymerase.

Addition of DMSO to a final concentration of 3% which serves to optimise template denaturation by removing DNA supercoils.

Using lower annealing temperatures as low as 45°C instead of 50-65°C, in order to facilitate binding.

Using higher annealing temperatures as high as 73°C, contrary to the above approach, in order to minimise chances of mis-annealing.

Using longer extension times up to 2min/kb instead of the protocol's suggested 30s/kb so as to ensure completion of synthesis of DNA strands.

Using GC buffer instead of the standard HF buffer. GC buffer can be used in GC-rich templates.

Using a combination of the addition of 1mM MgCl₂, 2mM DMSO and a 2min/kb extension time, the PCR was partially successful when using an annealing temperature of 58°C or 65°C, in that the correct band appeared at c. 8kb along with an incorrect band at c. 0.8kb indicating non-specific annealing of primers (Figure A.5b). Gel purification was carried out on the 8kb backbone band. Ligations were then carried out (Table A.6).

Ligations were transformed using XL10-Gold® cells. No non-amplified, circular backbone DNA was detected in reactions run without ligase using cut backbone indicating successful amplification and purification procedures. Colonies from the ligation plate were picked and grown in 2ml selective LB medium for 12-16 hours, after which the DNA was purified, a 5 μ l aliquot taken for an Apal digest and run on a gel. The Apal enzyme was used so as to distinguish between plasmids which have the Sir2rp3 present which will show two bands, and those where the gene has been replaced by eGFP which show one band only. Results in Figure 4.17 indicate that while clones 2-8 have only one band, it is smaller than the expected 8.7kb. The clones were taken forward for sequencing using primers from Table C.1 in Appendix C. Sequencing showed that the inserted DNA did not match the required eGFP DNA.

4.2.4.4 Appraisal of *tetR*-first ligation strategy

The following findings and conclusions were made regarding the assembly strategy which uses Gibson ligation as an alternative to standard ligation methods, in order to insert the *tetR* cassette and eGFP gene into the pRP backbone to get pStable:

1. Standard and non-directional ligation methods were replaced with the Gibson assembly method. This requires extensive primer design but results in a directional ligation and a more streamlined process which requires fewer insert and backbone preparation steps.
2. The Gibson ligation to insert the *tetR* cassette into pUbSir was successful where a non-directional ligation had not been on multiple modified previous trials. The backbone was prepared via digest, while the *tetR* cassette insert was prepared via amplification, which also served to add backbone-complementary overhangs.
3. A new construct, ptetR_UbSir, was developed via the above Gibson assembly. The plasmid, having the pRP backbone, has a regulatory *tetR* gene and a Ub-Sir2rp3 moiety which results in the TetR-regulated expression of ubiquitin-tagged Sir2rp3.
4. In the second Gibson ligation to insert eGFP into the ptetR_UbSir, results showed that the insert was not present in the backbone. Both backbone and insert were prepared via amplification, due to the presence of non-unique restriction sites in the ptetR_UbSir backbone which did not allow for backbone linearisation via digest. This required extensive optimisation which resulted in non-specific bands being amplified alongside the backbone. In order to address this, the same assembly strategy will be carried out, but the two ligation steps will be switched, as discussed next.

4.2.5 Oscillator assembly strategy inserting eGFP gene followed by the *tetR* cassette via Gibson assembly

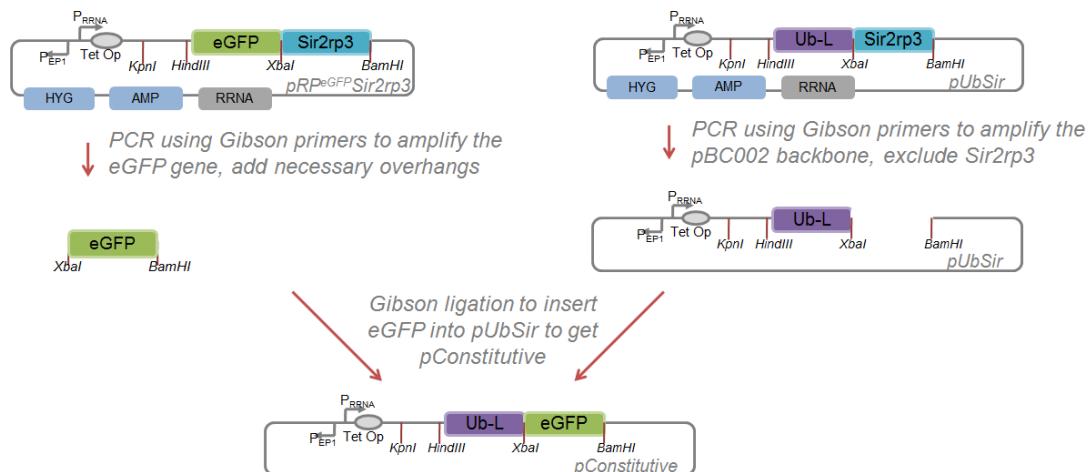
4.2.5.1 Introduction

The strategy uses Gibson assembly methods to construct the pStable oscillator in two main steps as seen in Figure 4.18. In Ligation 1, the eGFP gene is inserted into the pUbSir backbone at the unique XbaI and BamHI sites to get plasmid pConstitutive. In Ligation 2, the *tetR* cassette is inserted into pConstitutive at the KpnI sites to get the final pStable construct. The strategy is similar to the previous strategy but reduces the probability of primer mis-annealing by ligating eGFP first and the *tetR* cassette secondly since it has XbaI and BamHI sites which can affect amplification and ligation steps. The diagram leaves out details such as purification and other treatments which are referred to further on if valid to the discussion.

4.2.5.2 Ligation 1: Gibson assembly of pConstitutive by inserting eGFP gene into pUbSir

The pUbSir backbone could be prepared via an XbaI/BamHI sequential preparative digest but BamHI digestion has been shown to be suboptimal in previous strategies (Section 4.2.3.2). Hence, the backbone was prepared through PCR-amplification using forward primer (5'-3') GGATCC-TGCCCATTTAGTTGGC and reverse primer (5'-3') TCTAGAACTTGTCTAGCC. Similarly, the eGFP insert was prepared via PCR-amplification using forward Gibson primer (5'-3') CTAGA-CAAGTTCTAGAGGTGAGCAAGGGCGAGGAG and reverse Gibson primer (5'-3') GCCAACT-

Ligation 1. Insert eGFP into pUbSir backbone via a Gibson ligation to get pConstitutive



Ligation 2. Insert the tetR cassette into pConstitutive via a Gibson ligation to get pStable

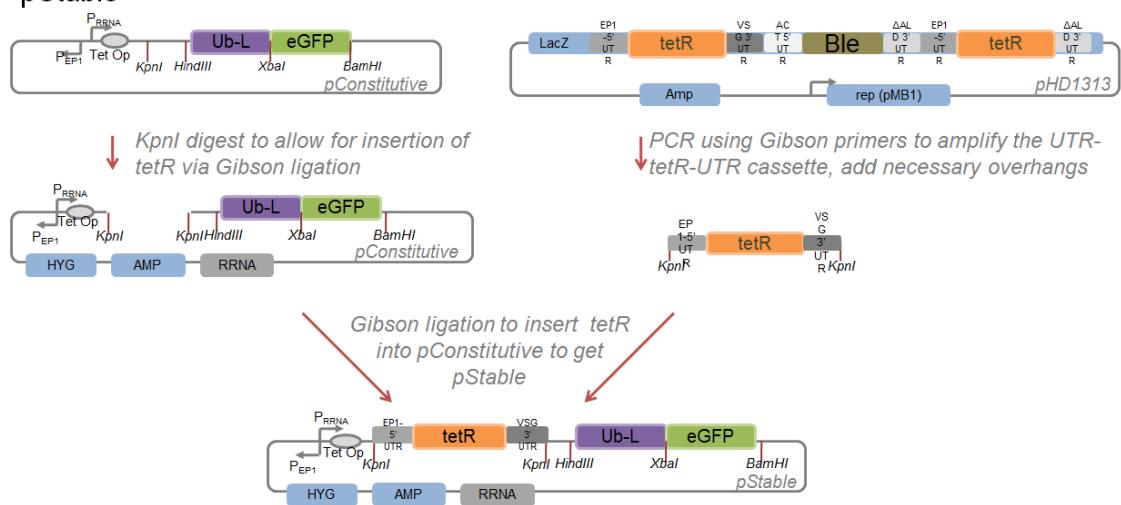


Figure 4.18: Assembly strategy to insert eGFP followed by the *tetR* cassette via Gibson ligation.

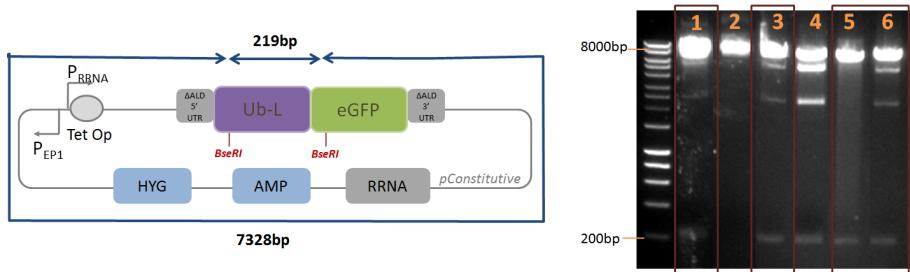


Figure 4.19: eGFP-first assembly strategy, Assessing success of ligation: The left panel shows pConstitutive digested with BseRI. The right panel shows gel results of digests of clones picked following Gibson ligation to insert eGFP into pUbSir to get pConstitutive. Lanes show 10 μ l of BseRI analytical digest of different clones.

AAATGGGCAGGATCCGCCTTCAAGACTTG which also served to add pUbSir complementary nucleotides.

Although the primers used for the amplification of eGFP in the previous strategy could have been used, new primers which added XbaI and BamHI restriction sites upstream and downstream of eGFP were designed so as to increase the size of the complementary regions between the amplified insert and backbone. Amplification of both DNA sequences was straightforward and required no optimisation. Once insert and backbone were successfully prepared (Figure A.6), ligations were set-up (Table A.7), after which transformations were carried out using XL10-Gold[®] cells.

Since the insertion of eGFP has been a repeated challenge in the construction process, a 10:1 insert-to-backbone ratio was used to maximise chances of ligation. However, this resulted in only one colony. It is possible that the excess of insert DNA interfered with the enzymes and thus decreased ligation efficiency. Growth was noted on the negative transformation plate corresponding to the ligation wherein no ligase was added, indicating the presence of non-amplified, circular backbone DNA.

Nonetheless, six colonies were picked, grown in selective LB medium for 12-16 hours and the DNA purified, after which a BseRI digest was carried out to distinguish between clones which have the original Sir2rp3 insert which should show one band, and those which have successfully ligated eGFP which should show two bands at c. 7.3kb and 0.22kb. This was adopted as an alternative to the use of Apal (used in the previous strategy) which only informed whether Sir2rp3 is present or not. BseRI, on the other hand, indicated whether eGFP is present or not. As seen in Figure 4.19, clones 1, 3, 5 and 6 appear to have the eGFP insert, showing both required bands although there are also unexplained bands present which can be due to uncut plasmid. Each clone was sequenced using primers from Table C.1 in Appendix C. Results confirmed the presence of the eGFP insert and showed no mutations present within the plasmid, confirming the successful construction of pConstitutive. Clone 1 was taken forward for use in the next step of the strategy.

4.2.5.3 Ligation 2: Gibson assembly of pStable oscillator by inserting the *tetR* cassette into pConstitutive

The next step was to carry out a second Gibson ligation in order to insert the *tetR* gene into the pConstitutive backbone to get the Goodwin oscillator construct. The *tetR* insert was prepared via

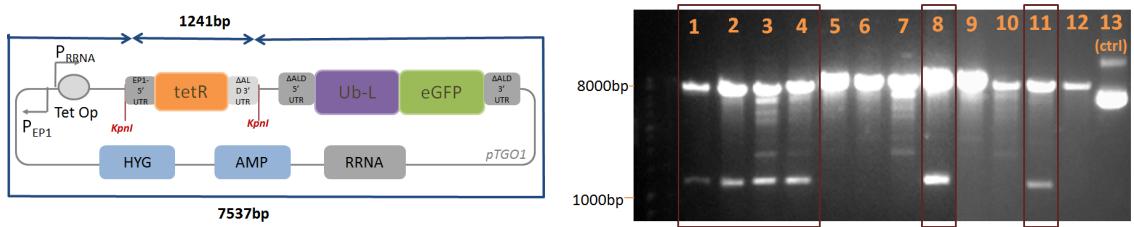


Figure 4.20: eGFP-first assembly strategy - Assessing success of ligation: Diagram on the left shows pStable cut with KpnI to assess success of Gibson ligation to insert *tetR* into pConstitutive to get pStable. The right panel shows gel results of digest. Lanes:
 1-12) 10 μ l clones, KpnI cut,
 13) 5 μ l clone, uncut.

amplification, using the primers (5'-3') CAATGATAGAGTGGTACCGTCTGGTGTGACCTTG and (5'-3') GCGCGTGCAGGGTACCTTGTACATATTGTCGTTAGAACGCC as used in the previous strategy when the *tetR* cassette was amplified for insertion into pUbSir (instead of pConstitutive). This was viable since both pUbSir and pConstitutive have the same pRP backbone. Following amplification, a sample was run on a gel after which DNA was purified and a second aliquot run on a gel to confirm DNA is not compromised (Figure A.7a).

The pConstitutive backbone was linearised at the unique KpnI restriction site at which the *tetR* cassette was to be inserted, bypassing the need for PCR amplification. In addition, in order to remove unwanted circular non-ligated backbone, a DpnI reaction was added in order to break down circular DNA. Following this, a sample of the reaction was run on a gel, the DNA purified and a second sample run on a gel to confirm DNA identity (Figure A.7b). Once the insert and backbone were prepared, ligations were carried out as per standard Gibson assembly protocols (Table A.8). Reactions were transformed using XL10-Gold[®].

Colony counts showed that reactions using DpnI had fewer colony counts than those which did not, indicating that the enzyme was being effective in breaking up non-digested DNA. However, for the corresponding negative control wherein a ligation was run using cut and DpnI treated backbone but no ligase, one colony was observed on the transformation plate, indicating uncut plasmid. 12 colonies were picked for further analysis from the transformation plates, grown in selective LB Medium for 12-16 hours and purified, after which an aliquot was taken for a KpnI digest to assess whether the *tetR* insert was present. As seen in Figure 4.20, clones 1, 2, 3, 4, 8 and 11 appear to have the insert, showing the correct bands at 1.2kb for the *tetR* cassette and 7.5kb for the pConstitutive backbone. Hence, the clones were sequenced using primers from Table C.1 in Appendix C, at which stage it was confirmed that the *tetR* cassette was present and hence, the plasmid encoding the Goodwin oscillator, pStable, was successfully constructed.

4.2.5.4 Appraisal of eGFP-first ligation strategy

Following the use of Gibson ligation to insert the eGFP and *tetR* cassette into pUbSir to get pStable, the following observations can be made:

1. The Gibson assembly was used to successfully construct pStable.
2. The strategy had the same steps as those used in the previous strategy but in reverse order,

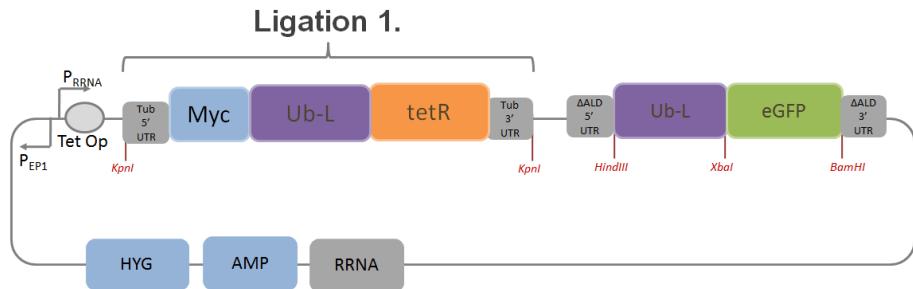


Figure 4.21: Diagram showing main cloning step required to construct the unstable oscillator plasmid, pUnstable, by inserting the UTR-Myc-Ub-L-*tetR*-UTR cassette at the *KpnI* site on the pConstitutive GRN in one step.

i.e. inserting the eGFP gene into pUbSir and then inserting the *tetR* cassette to get pStable. This allowed for the preparation of the DNA and inserts using a series of amplification and digestion steps which bypassed reactions which had previously been inefficient.

3. In the first step to ligate the eGFP gene into pUbSir, both insert and backbone were prepared via PCR. The backbone could have been prepared via *Xba*I/*Bam*HI digest but since digest of the *Bam*HI site had proved temperamental in previous attempts, amplification was used instead. In the final step to ligate *tetR* into pConstitutive, the insert was prepared via PCR while the backbone was prepared via *Kpn*I digest, since the *Kpn*I restriction site was unique.
4. The construction process also resulted in the development of the pConstitutive plasmid, which has a pRP backbone and a Ub-eGFP moiety which results in the constitutive expression of ubiquitin-tagged eGFP.

4.3 Construction of the unstable Goodwin oscillator

In order to assemble the plasmid encoding the unstable Goodwin oscillator, pUnstable, a DNA sequence was designed to encode the unstable oscillator features as depicted in Figure 4.21. Eurogentec Ltd. (Hampshire, U.K) successfully synthesized the DNA fragment, highlighted as 'Ligation 1' in Figure 4.21, and subcloned it into pConstitutive.

4.4 Transfection of *T. brucei* cells with DNA encoding the stable and unstable Goodwin oscillators

Following the successful assembly of both the pStable and pUnstable plasmids, the next step was to insert the two oscillator networks into *T. brucei* via transfection, as discussed below. The networks were transfected into both BSF and PCF forms of *T. brucei* in order to profile the network in both platforms and so as to analyse whether it performs better in one of the forms. PCF strains are more robust than their BSF counterparts and do not require a CO_2 incubator for maintenance. However, they need to be cultured at densities of over 1×10^6 cells/ml in order to survive, which results in longer transfection and thawing processes (see Sections 2.3 and 2.4).

4.4.1 Transfection of BSF and PCF *T. brucei* cells with DNA encoding the stable Goodwin oscillator

BSF trypanosomes transfected with pStable had a transfection efficiency of 7.5×10^{-7} cells, which is in line with expectations, while PCF trypanosomes had a transfection efficiency of 3.5×10^{-7} cells, which is considered to be low for PCF strains (MacGregor et al., 2013). Following transfection and incubation, all positive clones (15 for BSF cells, 4 for PCF transfection) were picked, grown to sufficient concentrations (1×10^6 cells/ml for BSF and 1×10^7 cells/ml for PCF) and glycerol stocks prepared.

The next step was to analyse which of the clones expressed protein the strongest. This is because transfection of exogenous (linearised) DNA in *T. brucei* is based on integration by homologous recombination at a ribosomal spacer which serves to separate foreign DNA from native sequences (Alsfeld et al., 2005). However, there are multiple copies of ribosomal spacers in the genome, therefore, it follows that the linearised pStable plasmid was integrated randomly in one of the 18 non-transcribed rRNA loci. Strength of protein expression depended on which loci integration occurred at. It was important to select the clone with the strongest signal in order to allow for clear-cut fluorescent observations of the plasmid's phenotype.

In order to assess which of the clones gave the strongest signal and also to ensure there was no leaky expression, protein samples were prepared for each clone. In addition, each clone was induced with $1 \mu\text{g}/\text{ml}$ tetracycline for 24 hours, after which protein was also prepared and uninduced/induced samples were run via a western blot to detect for the eGFP reporter protein. eGFP was selected for detection over the TetR protein due to the more established and reliable antibody products which are available.

Results from the western blot and the corresponding coomassie stain of eight induced/uninduced BSF clones can be found in Figure 4.22. At a protein band size of 27kDa, which corresponds to the eGFP protein, clones 1 and 5 have the strongest signals when induced. In addition, no leaky expression was observed, in that none of the uninduced clones had detectable eGFP signals. The coomassie stain showed that differences in observed signal strengths were due to protein expression rather than incorrect sample preparation. Based on this data, subsequent analysis was carried out on clone 1. The clone will be referred to as the BSF pStable clone.

For the transfection of PCF clones with pStable, the signal strength evaluation procedure as used for BSF pStable clones was implemented. Results from the western blot and corresponding coomassie stain can be seen in Figure 4.23, which shows tetracycline-induced protein samples alongside uninduced samples of four PCF *T. brucei* clones bearing the pStable GRN. The 27kDa eGFP signal was strongest for tetracycline induced clone 2. In addition, no leaky expression was noted and the coomassie showed correct sample preparation. Hence, clone 2 was used in subsequent analysis. It will be referred to as the PCF pStable clone.

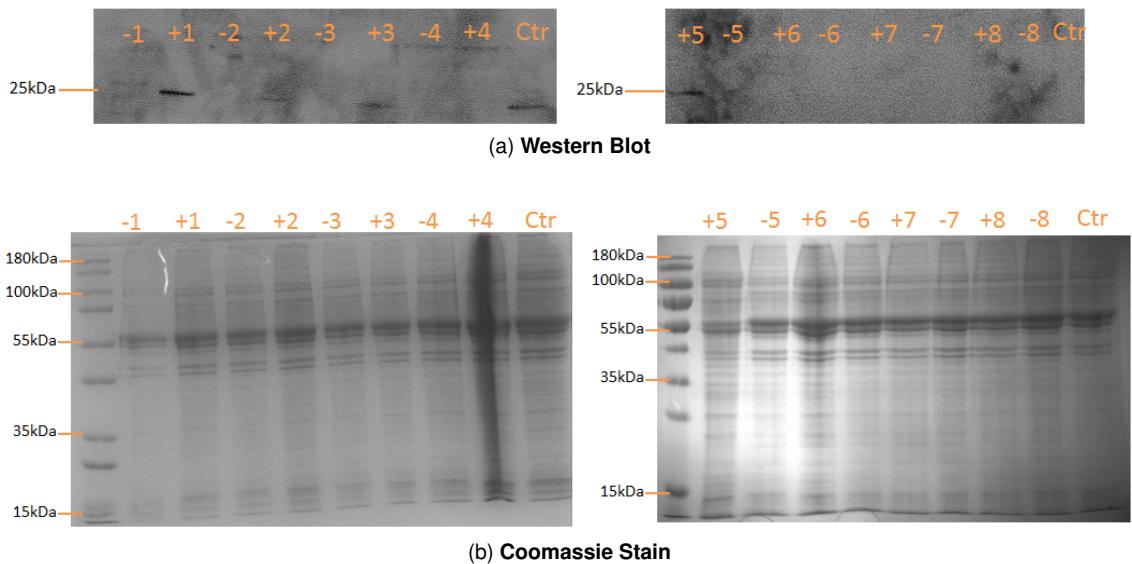


Figure 4.22: Assessing the strength of protein expression of eight different BSF *T. brucei* clones bearing the pStable plasmid and uninduced (-) or induced (+) with 1 μ g/ml tetracycline. Following protein preparation, analysis is carried out via a) western blot detecting for the eGFP signal and b) corresponding coomassie stain. Clones labelled as '-1', '-2', etc. are uninduced. Clones labelled as '+1', '+2', etc. are induced. The control is a protein sample of the constitutively expressed pConstitutive plasmid.

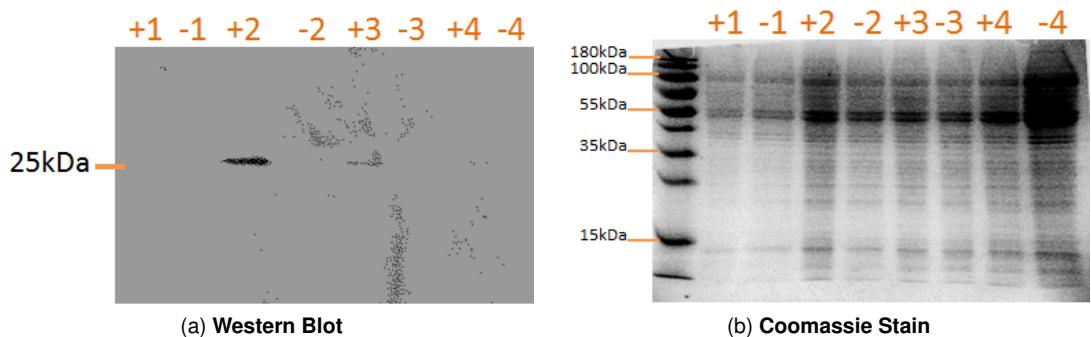


Figure 4.23: Assessing the strength of protein expression of four different PCF *T. brucei* clones bearing the pStable plasmid and uninduced (-) or induced (+) with 1 μ g/ml tetracycline. Following protein preparation, analysis is carried out via a) western blot detecting for the eGFP signal and b) corresponding coomassie stain. Clones labelled as '-1', '-2', etc. are uninduced. Clones labelled as '+1', '+2', etc. are induced.

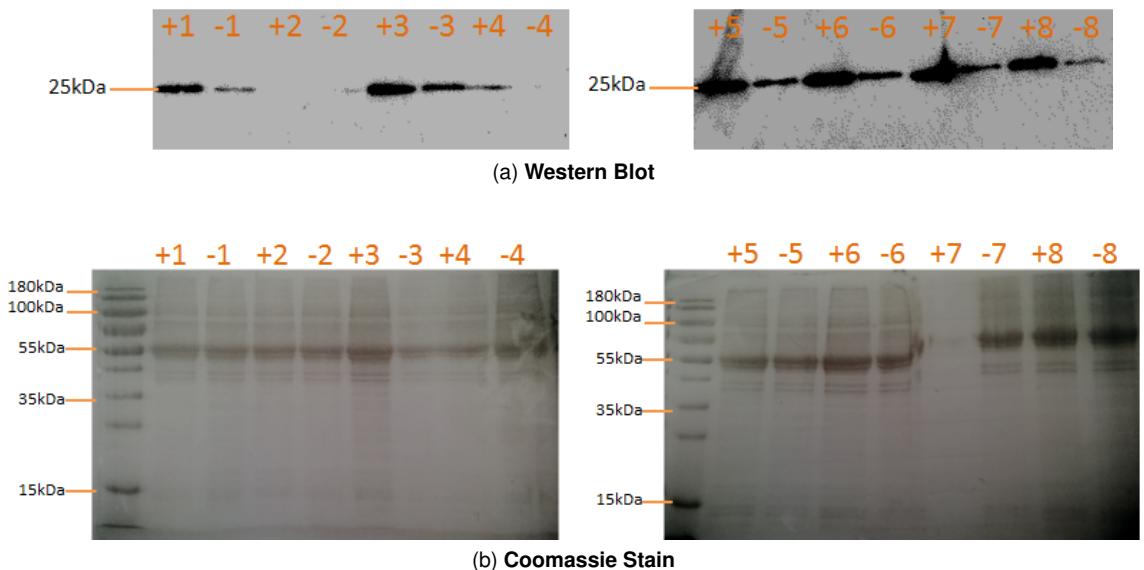


Figure 4.24: Assessing the strength of protein expression of eight different BSF *T. brucei* clones bearing the pUnstable plasmid and uninduced (-) or induced (+) with 1 μ g/ml tetracycline. Following protein preparation, analysis is carried out via a) western blot detecting for the eGFP signal and b) corresponding coomassie stain. Clones labelled as '-1', '-2', etc. are uninduced. Clones labelled as '+1', '+2', etc. are induced.

4.4.2 Transfection of BSF and PCF *T. brucei* with DNA encoding the unstable Goodwin oscillator

The unstable Goodwin oscillator, pUnstable, was transfected into both BSF and PCF trypanosomes. The BSF transfection had an efficiency of 5.5×10^{-6} . Following incubation, eight positive clones were picked into appropriate media, glycerol stocks prepared and the strength of protein expression of each clone analysed to pick the clone with the strongest fluorescent signal, as per the stable oscillator analysis.

Results, seen in Figure 4.24, show that there is basal constitutive transcription, as seen in eGFP signals detected even when transfected BSF *T. brucei* clones were not induced with tetracycline. The corresponding coomassie stain indicates that protein samples were prepared consistently. This is in clear contrast to pStable clones, wherein eGFP was only detected via western blots when tetracycline was present. However, there is a consistently clear difference in signal strength between induced and uninduced pUnstable clones. This indicates that an element of repression via TetR protein is still taking place.

Based on signal strength, the first BSF *T. brucei* clone bearing the pUnstable plasmid was taken forward for further analysis. Although there is a basal level of protein synthesis, it is still possible for oscillatory protein production to occur. The oscillator can function as a tuned GRN when compared to pStable, as per mechanisms in Section 4.1.2. The detected eGFP signals in uninduced clones can be attributed to the appended Myc tag and Ub-L sequences preventing the correct binding of TetR to the operator or to the protein sample being harvested during a peak phase of the oscillatory cycle. Further investigation via protein analysis and microscopy carried out to analyse this will be discussed later.

Following the insertion of pUnstable to PCF *T. brucei* cells, four positive clones were picked, trans-

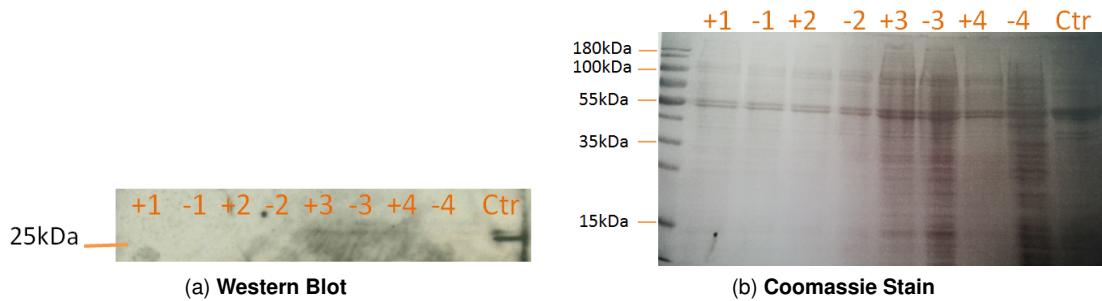


Figure 4.25: Assessing the strength of protein expression of four different PCF *T. brucei* clones bearing the pUnstable plasmid and uninduced (-) or induced (+) with 1 μ g/ml tetracycline. Following protein preparation, analysis is carried out via a) western blot detecting for the eGFP signal and b) corresponding coomassie stain. Clones labelled as '-1', '-2', etc. are uninduced. Clones labelled as '+1', '+2', etc. are induced. The control is a protein sample of the constitutively expressed pConstitutive plasmid.

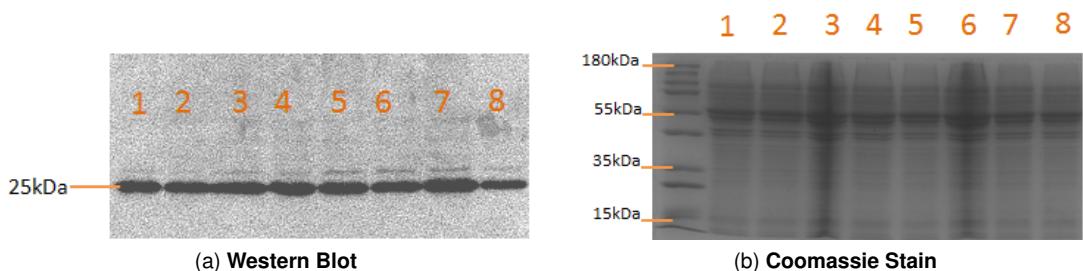


Figure 4.26: Assessing the strength of protein expression of eight different BSF *T. brucei* clones bearing the pConstitutive control plasmid which express protein constitutively. Following protein preparation, analysis is carried out via a) western blot detecting for the eGFP signal and b) corresponding coomassie stain.

lating to an efficiency of 2×10^6 cells. Glycerol stocks were prepared and the strength of protein expression of each clone analysed via western blot and a coomassie stain, with results seen in Figure 4.25. No eGFP signal was detected, either in the presence or absence of tetracycline, in spite of correct protein sample preparation (as seen by coomassie results). The plasmid was successfully transfected since the consistent use of antibiotic selection throughout the experiments would have resulted in poor cell growth. The lack of eGFP protein can either be attributed to the mechanisms of the plasmid itself not functioning within a PCF setting or the plasmid being inserted at a weakly expressing recombinant locus. Nonetheless, a clone was picked at random for microscopy imaging and phenotype characterisation.

4.4.3 Transfection of BSF *T. brucei* with the pConstitutive control plasmid

In parallel to the above transfections, the pConstitutive plasmid, which expresses Ub-L-eGFP constitutively, was transfected into BSF cells. This was done so as to use pConstitutive clones as positive controls throughout experiments and to check whether the P_{RRNA} promoter (which is also present in the stable and unstable oscillator plasmids) is functional and allows for constitutive and uniform gene expression. Following transfection, of the nine clones which were analysed for eGFP signal strength via western blot and coomassie stain, as seen in Figure 4.26, clone 7 gave a strong signal and hence was used in all subsequent analysis.

4.5 Discussion

Assembly and transfection of the stable Goodwin oscillator.

Following the assembly and transfection of the plasmid encoding the stable Goodwin oscillator GRN, the following conclusions can be made:

1. Successfully constructed the stable oscillator device, pStable, using a combination of four assembly strategies which involved multiple recombinant DNA and ligation techniques. Throughout, standard, three-fragment, non-directional and Gibson ligations were used to assemble the TetR-regulated plasmid, with the Gibson ligation method proving to be the more efficient approach.
2. Developed three other engineered plasmids as a result of the assembly process; pUbSir (resulting in constitutive expression of Ub-Sir2rp3 fusion protein), ptetR_UbSir (resulting in TetR-regulated expression of Ub-Sir2rp3 fusion protein) and pConstitutive (resulting in constitutive expression of Ub-eGFP fusion protein), all having the same pRP backbone as pStable. Progress can be seen throughout the assembly process in that a more standardised cloning practice was adopted as the four different strategies were implemented. Thus, for example, DNA prepared from a miniprep elution was replaced by DNA eluted via a maxiprep which results in higher quality material. Similarly, in later strategies, restriction enzymes in double digests were applied sequentially rather than simultaneously so as to increase digest efficiency.
3. Successfully integrated by recombination the stable Goodwin oscillator, pStable, into the genome of both BSF and PCF *T. brucei* strains. In order to maximise the probability of observing fluorescent oscillations, the clone with the strongest eGFP signal when induced with 1 μ g/ml tetracycline was selected for the phenotyping of the oscillator.

Assembly and transfection of the unstable Goodwin oscillator.

Following the assembly and transfection of the plasmid encoding the unstable Goodwin oscillator GRN, the following conclusions can be made:

1. The unstable Goodwin oscillator, pUnstable, was synthesised at Eurogentec Ltd. The plasmid was delivered in less than six weeks following the placement of the order. The construction of the stable device, on the other hand, took eighteen months. In light of the difference in timescales and the increasingly low costs for purchasing bespoke plasmids (Carlson, 2014), the practicality of assembling plasmids externally is highlighted.
2. Transfected the plasmid encoding the unstable oscillator into both BSF and PCF *T. brucei* strains. Analysis of transfected BSF clones showed that eGFP was detected even when no tetracycline was present. This can be either due to leaky expression, indicating that the appended Myc and Ub-L genes are preventing the correct binding of TetR to its operator or due to protein being prepared during a peak phase of the oscillatory cycle.
3. Demonstrated that, in theory, oscillations can still take place in pUnstable protein expression

dynamics. When transfected cells were induced with tetracycline, the eGFP signal flagged up was stronger than that in uninduced protein samples, indicating that an element of non-linear TetR repression is still being effected in a GRN with time-delay present, thus allowing for oscillations to manifest.

4. Transfected BSF *T. brucei* cells with pConstitutive, which expresses eGFP protein constitutively. This will be used as a control during phenotyping investigations.
5. The assembly of the pStable and pUnstable plasmids encoding for the Goodwin oscillator and the insertion of these plasmids within a *T. brucei* host is a novel development, with no similar engineered network constructs available. This will enable the investigation of *T. brucei* as a synthetic biology chassis.

Chapter 5

Establishing an experimental framework for measurement of oscillating fluorescence levels in viable *T. brucei* cells

Given the successful assembly of the stable and unstable Goodwin oscillators, the next objective was to develop an experimental set-up for the live cell imaging of single motile *T. brucei* cells. Imaging is to be done via brightfield and fluorescent microscopy, with images taken at regular intervals so as to enable the analysis of fluorescence levels of individual cells over time. Moreover, the imaging framework needs to allow for the imaging of cell populations, while also allowing for the tracking and identification of individual cells.

The challenge lies in that trypanosomes are motile cells going through swim-and-tumble phases and travelling up to $5.6\mu\text{m/s}$ (Weisse et al., 2012). Thus, unless their motility is constricted, cells will move in and out of focus and the field of vision during imaging and their paths will overlap with each other, as seen in Figure 5.1 which shows a series of images of BSF trypanosomes at 30 second intervals. This makes it difficult to distinguish between the paths of two cells over consecutive images.

The development of a framework for the imaging of isolated cells, rather than one which allows for the imaging of cell populations and tracking at the single cell level as in this case, was not considered as an option. Firstly, a single motile cell can still move in and out of focus resulting in poor quality images. Secondly, novel GRN characterisation within synthetic biology involves scanning a large number of cells. Imaging cells one at a time would make for an infeasible investigation in terms of time.

No similar studies have been undertaken in either the field of synthetic biology or the study of trypanosomes. Previous work in synthetic biology has focused mainly on the use of *E. coli* and mammalian chassis, with no emphasis made on the immobilisation of the cells. In trypanosome studies, on the other hand, cells are generally fixed, lysed, examined as a population or analysed *in vivo* in mice (see MacLean et al. (2013) for a review). More recently, a microfluidic device

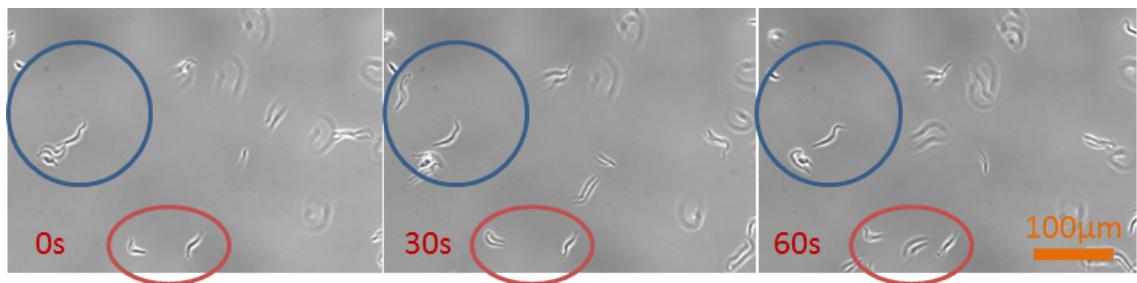


Figure 5.1: Images taken at 30 second intervals showing how free-swimming *T. brucei* mounted on 5mm multiwell slides move across the plane and in and out of focus, which limits cell tractability when comparing images. Images were taken using brightfield microscopy with an objective of 40X. Scale bar = 100 μ m.

has been developed for live cell imaging which embeds pillars of polydimethylsiloxane (PDMS) (Heddergott et al., 2012) which mimic red-blood cells, interfering with the cells' swimming route. However, this only results in stilted movement.

Table 5.1 lists the different methods and set-ups developed in this chapter. These can be categorised into two frameworks; those wherein cells are immobilised by encasing them within gel-like substances and those wherein cell movement is constricted through the use of a rigid microfluidic trapping chamber. Throughout, knowledge was drawn from imaging techniques used for trypanosomes and other live organisms such as zebrafish and roundworms, and then applied to wildtype BSF *T. brucei* cells. Each set-up was developed with the aim of satisfying three requirements:

1. Minimise motility of cells across the x-, y- and z- planes.
2. Maintain viability of cells for maximum amount of time. This is a challenge because BSF trypanosomes start to lose viability if they become immobilised (Price et al., 2010). They lose the ability to perform cytokinesis even though they continue to progress through multiple cell cycles (Broadhead et al., 2006). This results in misshapen clumps.
3. Be able to distinguish between cells within and between time-frames and avoid clumping of cells.

In order to assess each set-up, WT BSF cells which are less robust than their PCF counterparts (Price et al., 2010) were embedded accordingly and hourly cell-counts were taken to categorise the observed trypanosomes in one of three tracks, as seen in Figure 5.2:

1. Healthy cells which are motile or undulating at a fixed spot
2. Unhealthy cells which are clumped together but are still undulating (Kulkarni et al., 2013)
3. Shrunken, non-viable cells (Wang et al., 2010; Marinho et al., 2014)

Cells were observed using an inverted microscope at a 250X magnification, until the cell population was non-viable. At each time-point the number of motile/undulating cells, clumped but viable cells and non-motile shrunk cells within the field of vision were counted. The microscope was only switched on during counts so as to avoid overheating of cells.

	Material/ Technology used	Description	Section
1	Microfluidic cage devices	Use of microfluidic technology to develop trapping devices	Section 5.1
2	Agarose blocks	Use of agarose blocks to constrict movement	Section 5.2
3	Glass slides, Petri dishes and multiwell plates	Testing of different platforms to mount cells on	Section 5.3
4	Agar, agarose and xanthan gum solutions	Testing solutions as immobilisation agents	Section 5.4-5.6
5	Cygel™	Use as an immobilisation solution	Section 5.7

Table 5.1: Summary of the different set-ups, solutions and technology used to develop a framework for the measurement of fluorescence levels in viable *T. brucei* cells.

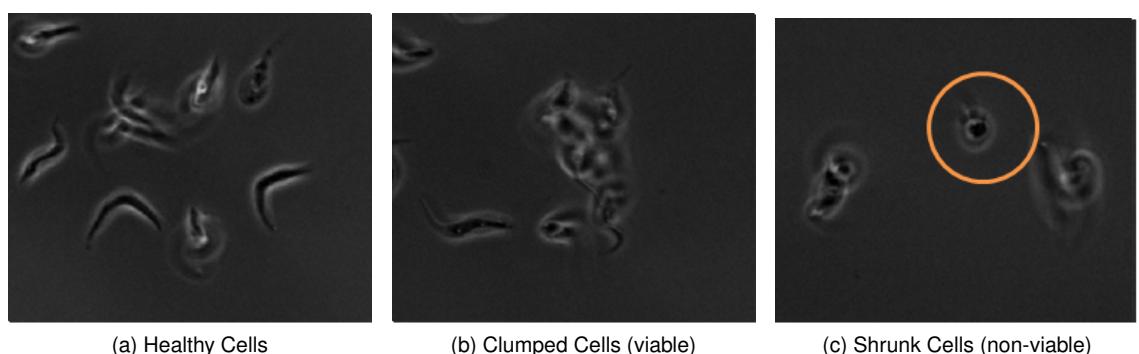


Figure 5.2: Images showing different BSF morphologies: a) Healthy viable cells, b) Unhealthy viable cells which clump together and undulate at a fixed spot and c) Non-viable cells which remain isolated but shrink and lose their original morphology. Images were taken using brightfield microscopy with an objective of 40X.

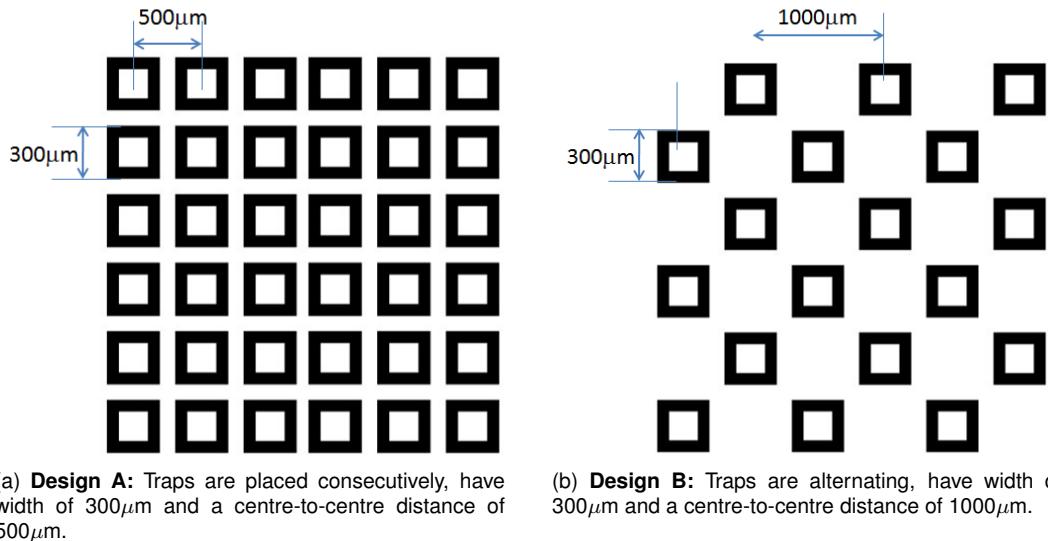


Figure 5.3: Microfluidic cage devices: CAD design of two different layouts of microfluidic cage devices used to trap trypanosomes.

5.1 Microfluidic cage devices

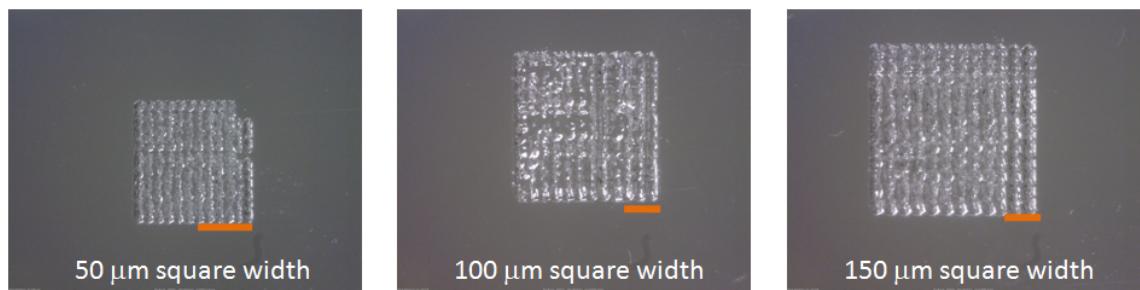
The technology of microfluidics is being merged with synthetic biology to address characterisation efforts in a resource-efficient way. Research has shown the potential of using microfluidic devices in precise analytical and chemical studies (San-Miguel & Lu, 2013) and for use as mechanical traps (Zhang et al., 2010; Aufderheide, 2008).

5.1.1 Design and fabrication of microfluidic traps

Microfluidics technology was used to develop a trapping device (Figure 5.3), consisting of multiple ‘cages’ which entrap single trypanosome cells while still allowing for the imaging of small populations of cells at once. Moreover, cells could be tracked since they were trapped within a bound area. In theory, cells would retain viability for a longer span of time due to unrestricted motility (Price et al., 2010). Although it is also implied that cells can swim in and out of focus, this can be controlled by making the traps shallow.

Trypanosomes measure 6-42 μm in length by 1-3 μm in width. Therefore, each cage required a surface area less than $150 \mu\text{m}^2$. However, laser technology available was not capable of satisfying these requirements. Instead, square traps of width 50 μm , 100 μm , 150 μm and 300 μm were designed and developed, as per Figure 5.3, which shows two different CAD designs using a square width of 300 μm . In Design A cages were placed consecutively next to each other, while in Design B cages were placed in an alternating pattern. The traps could be manufactured using a laser mark speed of 200 or 300 mm/s, resulting in two different well depths.

Both types of design were taken forward and manufactured as per Section 2.4.3.1. The designs were ablated onto PMMA, which is a widely used polymer for the use of microfluidic devices in cell culture (Ni et al., 2009). Figure 5.4 shows images of the manufactured products. Traps placed



(a) **Design A:** Manufactured microfluidic devices having a square width of $50\mu\text{m}$, $100\mu\text{m}$ or $150\mu\text{m}$. Wells are placed consecutively next to each other. Scale bar = $300\mu\text{m}$.



(b) **Design B:** Manufactured microfluidic device having a square width of $300\mu\text{m}$ and a centre-to-centre distance of 1mm . Wells are alternating with empty squares to allow for better definition. Scale bar = 2mm .

Figure 5.4: Microfluidic cage devices: Images of different microfluidic cage devices constructed having a) a consecutive pattern with square width of $50\mu\text{m}$, $100\mu\text{m}$ or $150\mu\text{m}$ or b) an alternating pattern with square width of $300\mu\text{m}$. In both cases, laser mark speed is of 300mm/s .

consecutively to each other (as per Design A, Figure 5.4a) with a square width of $50\mu\text{m}$ have cage walls which are not clearly defined. This is a consequence of the heat from the laser beam which causes an element of melting around the path it etches. However, the effect is less noticeable when etching larger wells with a width of $150\mu\text{m}$. Moreover, definition is improved when using an alternating well pattern as in Design B (Figure 5.4b).

Based on the above images, traps developed as per Design B were taken forward for testing with BSF cells. Specifically, traps with individual cages having width $50\mu\text{m}$, $100\mu\text{m}$ or $150\mu\text{m}$ and centre-to-centre distance of 1mm were used. Traps with individual cells having width $300\mu\text{m}$ were considered to be too large for the scope of this study. Moreover, both traps ablated with a laser speed of 200mm/s and 300mm/s were tested. Since the duration of ablation is unknown, the depth of the traps can not be calculated. In total, six devices were tested with BSF cells in order to assess which design satisfied the set-up criteria.

5.1.2 Testing of microfluidic traps using live *T. brucei* cells

BSF cells were grown to a density of 1×10^6 cells/ml. Following this, $50\text{-}100\mu\text{l}$ of cells was pipetted over the wells and left to stand for two minutes. Following this, a coverslip was placed on top and sealed using nail varnish. In all six devices tested, over 50% of trypanosomes retained viability for 8-9 hours, after which lack of oxygen and CO_2 , and pH changes led to cell death.

This period of cell viability is longer than that recorded for other tested methods and can be attributed in part to the cells retaining full motility and the larger volume of media, oxygen and CO_2 .

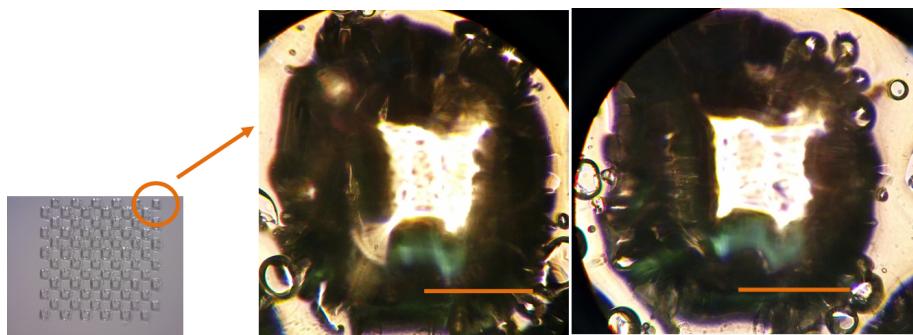


Figure 5.5: Microfluidic cage devices: The image on the left shows the device with wells having square width of $50\mu\text{m}$. The circled wells are amplified in the middle and right images using a magnification of 400X. Visibility is hampered due to the thick border around the perimeter of the well, the bubbles and small dark spots which are present. Scale bars = $20\mu\text{m}$.

supply available to trapped cells. The large wells also meant that multiple cells were inserted into each trap. This could be resolved by diluting the cells sufficiently so as to increase probability of having one or less cells per well. Images taken at 400X magnification showed surfaces of wells to be rough, dark and uneven, which makes the distinguishing of cells from the dark patches in the well difficult. Figure 5.5 shows an example of this, with images showing thick cell borders, dark patches and air bubbles.

5.2 Agarose blocks for entrapping live *T. brucei* cells

In the above method, microfluidic cages were used to trap cells within a bound area. The following method by Huang et al. (2014) adopts a different approach wherein cells are both trapped and immobilised between an agarose block and a glass slide. Similar experiments using 'agarose pads' (Kim et al., 2013) have been successfully carried out to entrap the roundworm *Caenorhabditis elegans*.

Following set-up as per Section 2.4.3.2, the device was incubated at room temperature. The set-up was assembled with no difficulties. Although cells were easy to distinguish due to low density, the cell population lost viability within two hours of preparation. With respect to the other framework criteria, the set-up did not restrict trypanosome motility sufficiently as to be able to track individual cells over time.

5.3 The use of glass slides, petri dishes and multiwell slides as bases for experimental set-up to immobilise live *T. brucei* cells

In a modified approach to the above method, cells can be encased within liquid substances having different viscosities and properties, which substances solidify to different degrees after incubation at room temperature to form encasing and immobilising gel matrices around the cells. For this

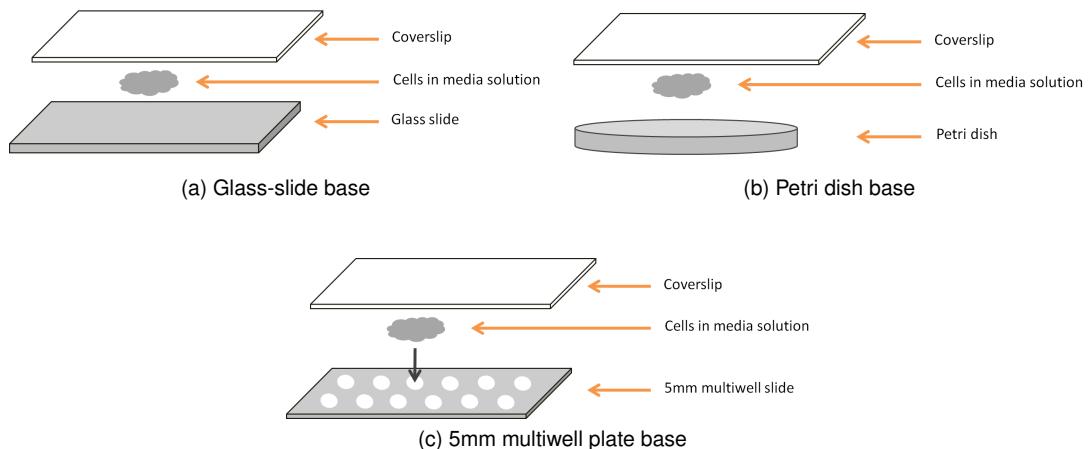


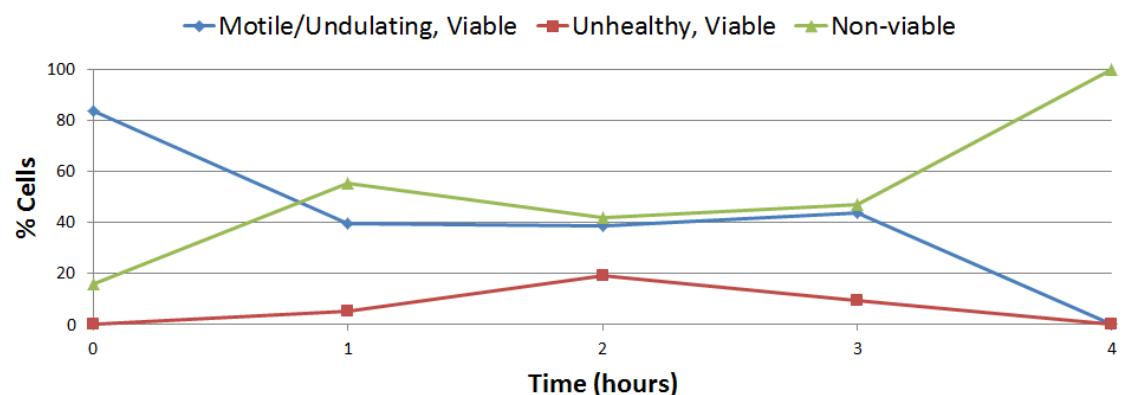
Figure 5.6: Diagram representing the experimental set-up to test the optimal base for the maintenance and imaging of viable *T. brucei* cells. The set-up consisted of: a) a glass slide, b) petri dish or c) 5mm multiwell slide, on top of which cells suspended in media were pipetted, covered with a cover slip and sealed to prevent spills.

purpose, the first requirement was to determine which of glass slides, Petri dishes or 5mm multiwell slides served as the best base for experimental set-ups. Cells were suspended in media with no immobilising agent applied, and constricted between two solid barriers; a coverslip and a glass slide or Petri dish or 5mm multiwell slide, as seen in Figure 5.6. In addition to determining the span of time over which cells retained viability, observations were made on ease of handling of each set-up.

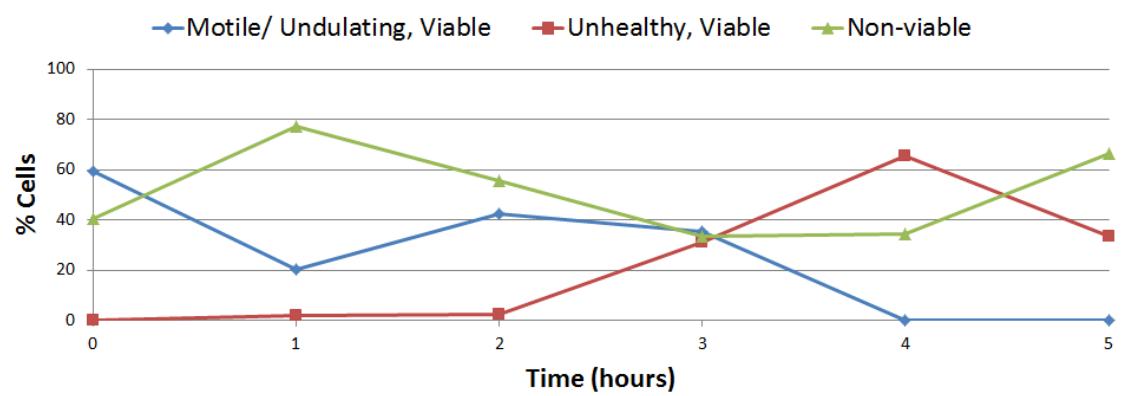
Cells grown to a density of 1×10^6 cells/ml in 10ml media were pelleted for 10 minutes at 2000rpm, the supernatant was removed and the cells were resuspended in media to a concentration of 1×10^7 cells/ml. $5\mu\text{l}$ of resuspended cells were pipetted onto a clean glass slide or 9cm Petri dish or multiwell slide, covered with a coverslip and sealed with nail varnish to comply with Health and Safety regulations.

Results of cell counts can be found in Figure 5.7 which plots the percentage of viable and non-viable cells mounted onto a Petri dish (Figure 5.7a) or glass slide (Figure 5.7b) over time. Results show that cells retained viability for a maximum of four hours when using either platform. In both cases, over 50% of cells show signs of being unhealthy (clumped) within the first hour. This can be attributed to a change in pH levels, insufficient supply of nutrients, oxygen and CO_2 and sub-optimal incubation conditions. Since no immobilising agent was applied, cells were unrestricted in their movement. Duplicates were not carried out, due to the short lifespan of cells. Even if additional runs of the study resulted in a longer lifespan, the variability made the method unreliable.

The multiwell slide set-up showed similar results (data not shown). However, the use of multiwell slides made for easier focusing under the microscope than when using Petri dishes and glass slides. Therefore, 5mm multiwell slides were used for subsequent studies.



(a) Set-up using a 9mm petri dish as a base



(b) Set-up using a glass slide as a base

Figure 5.7: Petri dish and glass slide set-up: Plot showing the percentage of viable healthy cells, viable unhealthy (clumped/misshapen) cells and non-viable cells over time which were encased using a) a Petri dish and coverslip set-up and b) a glass slide and coverslip set-up.

Concentration	Time post-preparation		
	0 mins	3 mins	7 mins
1% agarose	Good	Good	Good. Material is more viscous.
0.5% agarose	Good	Good	Good. Material slightly more viscous.
0.25% agarose	Good	Good	Good. Material's viscosity not affected.

Table 5.2: Agarose as an immobilisation agent: Handling properties of different concentrations of liquid agarose solutions at room temperature at 0, 3 and 7 minutes post preparation.

5.4 Agarose solutions as live *T. brucei* cell immobilisation agents

Agarose was the first immobilising agent to be tested on trypanosomes using the multiwell slide set-up. Agarose is a traditional cell-embedding method and is recommended by Price et al. (2010), who use 3% agarose, for long-term live trypanosome imaging. It has also been used in the imaging of live motile zebra fish embryos (Kaufmann et al., 2012; Renaud et al., 2011). Agarose has been used in the study of trypanosomes, although often in a different context, wherein cells are spread on it and are motile (Oberholzer et al., 2010; Carruthers & Cross, 1992) or else embedded in agarose for lysing (Inverso et al., 2010) or pulsed field gel electrophoresis (Leech et al., 2004).

5.4.1 Handling and viscosity properties of agarose solutions

Prior to assessing the effect of agarose solutions on motile trypanosomes, different agarose dilutions were prepared and their handling properties assessed. This was done in order to select those solutions which could be easily handled with pipettes, thus preventing technical barriers when setting up. 2% agarose was prepared in dH₂O and autoclaved. The solution was briefly heated until liquid, after which 1ml 1% agarose, 0.5% agarose and 0.25% agarose dilutions were prepared in media.

After preparation, each solution was pipetted up and down using P200 pipettes to see how easy it is to handle the material. This was done at 0, 3 and 7 minute intervals post preparation, to see how the solution changes over time and whether it solidifies. The 3 minute time point, in particular, is important because during the experiment involving cells, preparation and pipetting of the cells and agarose solution onto the multiwell slide is expected to take a similar amount of time. Results can be found in Table 5.2. The 0.25%, 0.5% and 1% agarose concentrations are relatively easy to handle at all time points. Viscosity of the 1% and 0.5% agarose concentrations was observed to increase over time. The 0.25% dilution, on the other hand, retained very low viscosity. This would be ineffective in immobilising *T. brucei*. Based on these results, the 1% and 0.5% agarose solutions were taken forward to assess how these concentrations affected cells in terms of motility, viability and tractability.

5.4.2 Effect of agarose solutions on *T. brucei* cell motility and viability

To analyse the effect of the 1% and 0.5% agarose solutions on trypanosomes, 10ml of cells at 1×10^6 cells/ml was centrifuged for five minutes at 2000rpm, after which the supernatant was removed and the cells were resuspended in media to a concentration of 1×10^7 cells/ml. The required agarose solution was then prepared, mixed gently and allowed to cool for two minutes. Immediately after this, $2\mu\text{l}$ of cells was pipetted onto a multiwell slide and overlaid with $2\mu\text{l}$ of the agarose solution to form an encasing dome, after which a coverslip was placed on top, and the edges sealed using nail varnish. The assembly was then incubated at room temperature and hourly cell counts taken to categorise cells as motile and viable or unhealthy but viable or non-viable. The experiment was carried out in duplicates.

Figure 5.8 shows that over 50% of trypanosomes embedded in 0.5% agarose (Figure 5.8a) lose healthy viability by the two hour time-mark. All cells lose healthy viability by the three hour time-mark due to suboptimal nutrient supply and aeration conditions.

Over 90% of cells overlaid with 1% agarose (Figure 5.8b) lose healthy viability after the three hour time point. Cells are sensitive to the preparation and fixing process, with less than 70% of cells being viable at the first observation time-point. A large element of variability was observed between duplicate readings (over 65% in some cases).

5.5 Agar solutions as live *T. brucei* cell immobilisation agents

The embedding of *T. brucei* cells within agar is a less common practice than the use of agarose. It has, however, been used to embed other motile and non-motile organisms as diverse as the bacterial *Pseudomonas aeruginosa* (Pitt & Bradley, 1975) and *Lactococcus lactis* (Floury et al., 2013) and the trypanosomal strain *Trypanosoma cruzi* (Wittner et al., 1982). In these studies, agar was used at a concentration of 0.4-0.6% (Pitt & Bradley, 1975; Wittner et al., 1982).

5.5.1 Handling and viscosity properties of agar solutions

In order to avoid technical barriers during the cell mounting process, different agar dilutions were tested for ease of handling, similarly to the previous cases. 2% agar was prepared in dH₂O and autoclaved, after which it was briefly heated until liquid and 1ml 1% agar, 0.5% agar and 0.25% agar dilutions were prepared in media.

Following preparation, volumes were pipetted using P200 pipettes at 0, 3 and 5 minute intervals. Table 5.3 shows that the 1% agar solution solidifies too quickly for use within a real-time setting. Based on this, the two lower concentrations (0.5% and 0.25% agar) were taken forward to be tested with *T. brucei* to assess how these concentrations affected cells in terms of motility, viability and tractability.

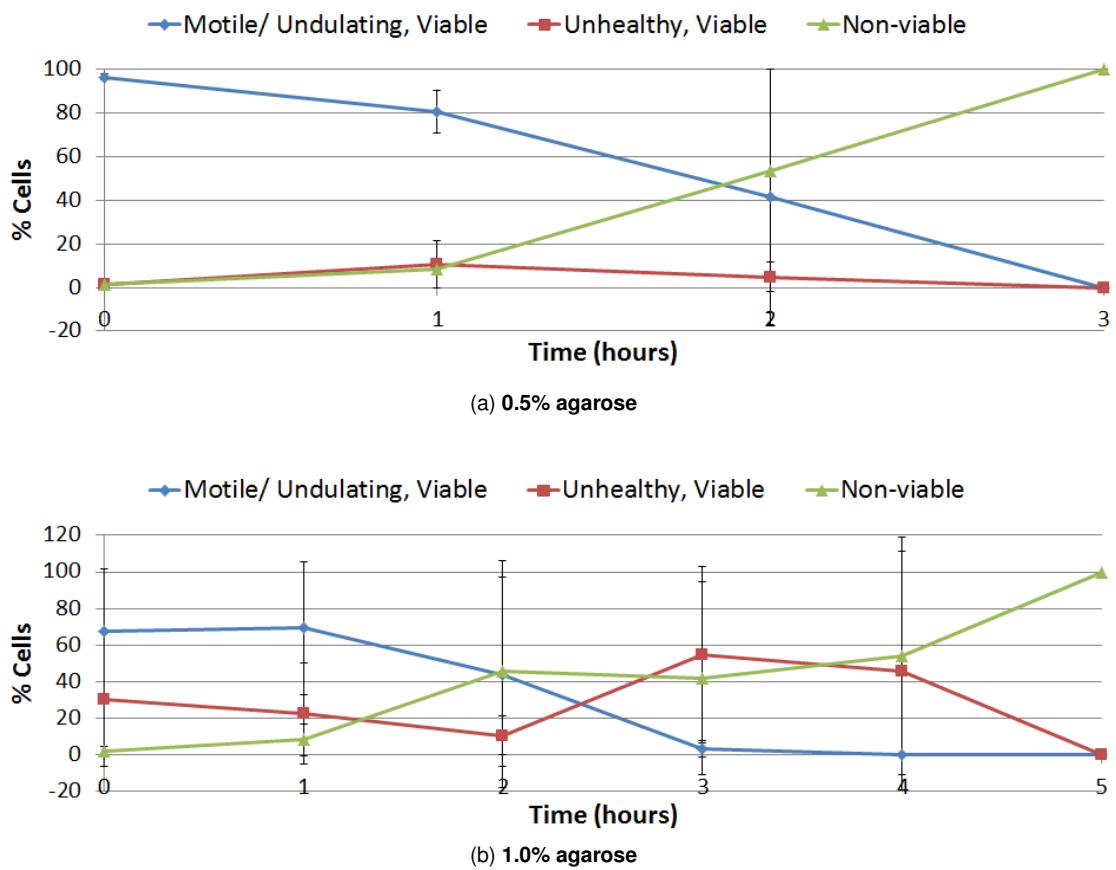


Figure 5.8: Agarose as an immobilisation agent: Plot showing the percentage of viable healthy cells, viable unhealthy (clumped/misshapen) cells and non-viable cells over time which were overlaid with a) 0.5% agarose and b) 1% agarose. Error bars represent standard deviation. The experiment was carried out in duplicates.

Concentration	Time post-preparation		
	0 mins	3 mins	7 mins
1% agar	Good	Medium has solidified. Can not be pipetted.	Viscosity is unchanged. Can not be pipetted.
0.5% agar	Good	Material starts to solidify. Requires slow pipetting.	Viscosity is unchanged. Requires slow pipetting.
0.25% agar	Good	Good	Good. Material's viscosity not affected.

Table 5.3: Agar as an immobilisation agent: Handling properties of different concentrations of liquid agar at room temperature at 0, 3 and 7 minutes post preparation.

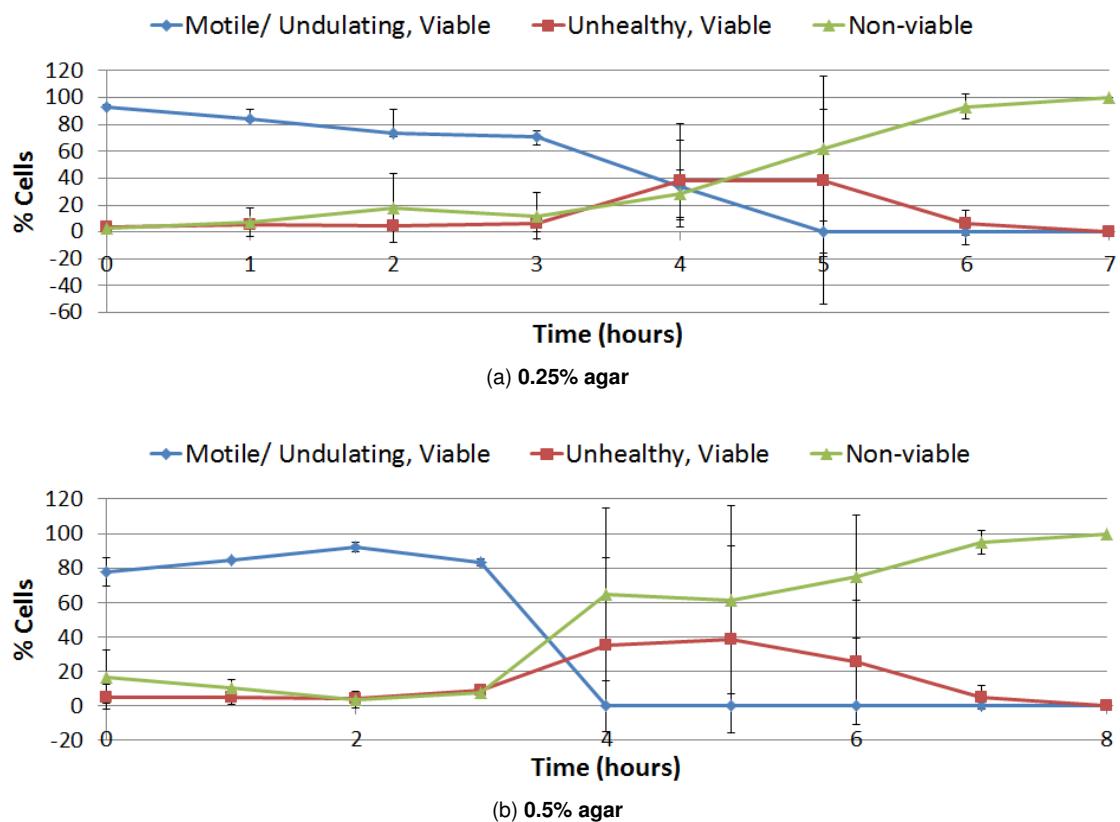


Figure 5.9: Agar as an immobilisation agent: Plot showing the percentage of viable healthy cells, viable unhealthy (clumped/misshapen) cells and non-viable cells over time which were overlaid with a) 0.25% agar and b) 0.5% agar. Error bars represent standard deviation. The experiment was carried out in duplicates.

5.5.2 Effect of agar solutions on *T. brucei* cell motility and viability

To observe the effect of agar on embedded trypanosomes, cells were set-up as per Section 5.4.2, replacing the use of agarose with agar at the appropriate dilution. Cell count results are represented in Figure 5.9 which shows that over 50% of cells fixed in 0.25% agar retained healthy viability for three hours. This was followed by a one hour period during which the biggest percentage (39%) of cells were unhealthy (clumped/misshapen) but motile and viable. After this, the majority of cells (over 60%) turned non-viable, but it took over six hours for all cells to lose viability. For *T. brucei* fixed in 0.5% agar, a high percentage (over 80%) maintained healthy viability for a three-hour time-frame, after which there was a sharp drop over a one hour period. A small percentage of cells (5%) retained (unhealthy) viability for over seven hours.

Error bars representing standard deviation indicate a large element of variability following Hour 4. However, prior to this, an error of less than 10% was recorded for almost all time-points across the three tracks. This adds an element of reliability to the agar method. As with agarose solutions, over 7% of cells lost viability during the preparation and fixing process. Although cells embedded in 0.5% agar were more sensitive to the setting-up process than when using 0.25% agar, the former outperforms the latter due to the high percentage of cells (+ 80%) which retain viability for the first three hours of the experiment. In conclusion, the agar set-up addresses the criteria for long-term viability with cells viable for up to five hours post-immobilisation. Cell motility, however, was not fully restricted, resulting in limited cell tractability.

5.6 Xanthan gum solutions as live *T. brucei* cell immobilisation agents

Xanthan gum is a polysaccharide which retains stability over a range of temperatures and pH values (Pace Pereira Lima et al., 2012). It has been used as an encasing agent in plant tissue imaging as an alternative to agar (Jain & Babbar, 2006).

5.6.1 Handling and viscosity properties of xanthan gum solutions

Prior to testing its effect on trypanosome cells, different concentrations of xanthan gum were prepared to assess viscosity and handling properties. The research by Jain and Babbar (2006) did not provide information on the concentration of xanthan gum used. 1% xanthan gum solution was prepared in media and diluted to 0.5% and 0.25% solutions in media for 1ml volumes. No heating of xanthan gum was required prior to dilutions, since the solution does not solidify at room temperature. This also implies that there was no need for observations of pipetting properties at multiple time points following preparation, since the viscosity is not expected to change with time as in the previous cases. When pipetted using P200 pipettes, the 1% xanthan gum solution was too viscous to pipette well. However, both 0.5% and 0.25% xanthan solutions were easy to handle, with viscosity being neither too low as for the solution to be too liquid and neither too high as to prevent correct pipetting. Hence, both solutions were taken forward to test on trypanosomes.

5.6.2 Effect of xanthan gum solutions on *T. brucei* cell motility and viability

In order to observe the effect of xanthan gum on embedded trypanosomes, cells were prepared as per Section 5.4.2, replacing the use of agarose with xanthan gum and omitting heating steps. Hourly cell count results can be seen in Figure 5.10, which shows the percentage of viable and non-viable trypanosomes over time.

The plots for the percentage of viable cells overlaid with 0.25% xanthan gum (Figure 5.10a) and 0.5% xanthan gum (Figure 5.10b) show that in both cases, cells were sensitive to the preparation and embedding process, with c.90% and 71% of cells encased in 0.25% and 0.5% xanthan gum respectively being viable and healthy on the first count. Less than 50% of cells immersed in 0.25% xanthan gum were viable after one hour of incubation, however, they then retained the same level (c. 40%) of healthy viability for a period of roughly four hours, after which there was a rise in the number of unhealthy (clumped/misshapen) but motile and viable cells for a period of one hour, before the population became non-viable. This pattern was not noted in the other solutions (except when using 40% CygelTM), wherein cells tended to lose healthy viability at a roughly steady rate. On the other hand, over 50% of cells embedded in 0.5% xanthan gum lost viability after the two hour time-mark and no healthy cells were observed after the three hour mark. A large variability is present across all tracks and time-points, as represented by plot error bars.

In conclusion, the 0.25% xanthan gum set-up satisfies the criteria for maintaining long-term viability but not the criteria for minimal motility and maximal tractability, with unrestricted cellular movement observed throughout. 0.5% xanthan gum does not satisfy any of the three criteria.

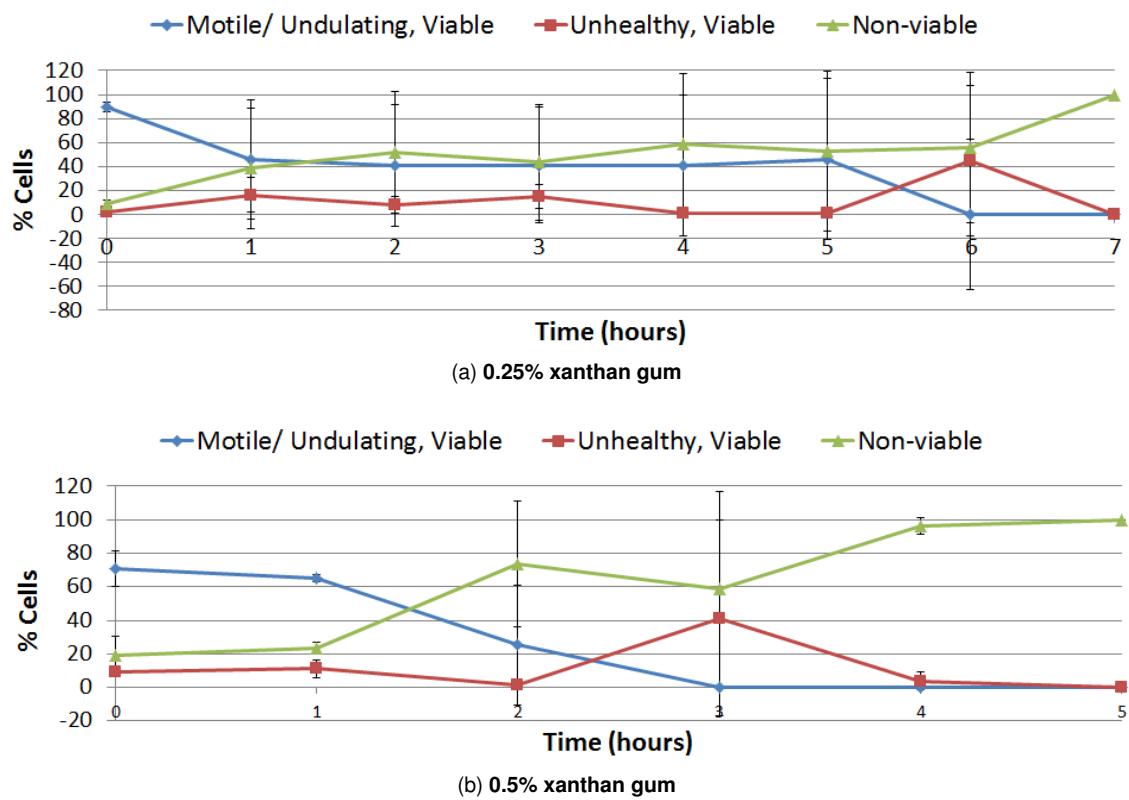


Figure 5.10: Xanthan gum as an immobilisation agent: Plot showing the percentage of healthy and unhealthy (clumped/misshapen) motile viable and non-viable cells over time which were overlaid with a) 0.25% xanthan gum and b) 0.5% xanthan gum. Error bars represent standard deviation. The experiment was carried out in duplicates.

5.7 CygelTM as a live *T. brucei* cell immobilisation agent

CygelTM is a clear thermoreversible product which is liquid below room temperature and solid otherwise. Price et al. (2010) used two different CygelTM solutions, one in its original format and the other supplemented with 10mM glucose to immobilise both BSF and PCF *T. brucei* for up to three hours. The former was found to be less toxic to cells and hence was selected for use in this study.

5.7.1 Handling and viscosity properties of CygelTM solutions

Handling properties and viscosity of various CygelTM dilutions were assessed to select two concentrations to be tested on cells, based on the ease of handling with pipettes at different time-points following preparation. This would ensure that preparation and pipetting technical difficulties would not compromise the set-up. Initially, CygelTM was placed on ice for two minutes as per supplier instructions, after which 1ml 100%, 80%, 60%, 40% and 20% CygelTM dilutions were prepared in dH₂O. 40X PBS was supplemented to a 2.5% (v/v) concentration. The diluted volumes were kept on ice.

Once all solutions were prepared, each was pipetted up and down using P200 pipettes to observe how easy it is to handle the material after two and five minute intervals of incubation at room temperature. Results, as seen in Table 5.4 show that all solutions can be pipetted without difficulty on ice, although the 100% and 80% concentrated solutions require slow handling due to a high vis-

Concentration	Time post-preparation		
	On ice	2 mins	7 mins
100% Cygel™	Good. Required slow pipetting.	Good. Required slow pipetting.	Material is solidified.
80% Cygel™	Good. Required slow pipetting.	Good. Required slow pipetting.	Material is solidified.
60% Cygel™	Good	Good	Good. Material more viscous due to minor solidifying.
40% Cygel™	Good	Good	Good. Material more viscous due to minor solidifying.
20% Cygel™	Good	Good	Good. Material's viscosity not affected.

Table 5.4: Cygel™ as an immobilisation agent: Handling and pipetting properties of different concentrations of Cygel™ solutions when placed on ice, after two minutes of incubation at room temperature and after seven minutes of incubation at room temperature.

cosity. All liquids retained a similar viscosity after two minutes of incubation at room temperature. After a further five minute interval, the 100% and 80% solutions were solidified and had a gel-like consistency. The 60% and 40% solutions were noticeably more viscous than before, while the 20% consistency retained low viscosity.

Based on this knowledge, the 80% and 40% Cygel™ solutions were taken forward for testing with BSF *T. brucei*. The two solutions had a different consistency and viscosity, yet could both be handled after two minutes of incubation time. The next step was to assess how these concentrations affected cells in terms of motility, viability and tractability.

5.7.2 Effect of Cygel™solutions on *T. brucei* cell motility and viability

Cells were prepared as per Section 5.4.2, replacing the use of agarose with 80% and 40% Cygel™, which was prepared as per supplier instructions. Cell count results can be seen in Figure 5.11, which shows the percentage of viable and non-viable trypanosomes over time when they are encased in either 80% or 40% Cygel™. Results show that the application of 80% Cygel™ (Figure 5.11b) is unreliable due to two main factors: there is a large variability present between duplicate runs as seen from the error bands representing standard deviation, and cells remain healthy for only one hour. This is in line with observations from Price et al. (2010) who report a 90% drop in BSF viability after just one hour. In terms of minimising motility, trypanosomes were not observed to move freely, instead undulating within a fixed spot.

Cells embedded in 40% Cygel (Figure 5.11a) retained viability for a six hour period with over 50% of cells viable for the first four hours. In addition, there is a smaller amount of variability than when cells were fixed with 80% Cygel (maximum variability between duplicate readings is 50% when using 80% Cygel™ and 20% when using 40% Cygel™). When comparing the effect of 40% Cygel™ on *T. brucei* motility with the effect of the previous immobilisation agents, trypanosomes were noticeably more restricted in movement, resulting in increased cell tractability. The same process was carried out using Cygel™ dilutions prepared in media instead of dH₂O. No changes

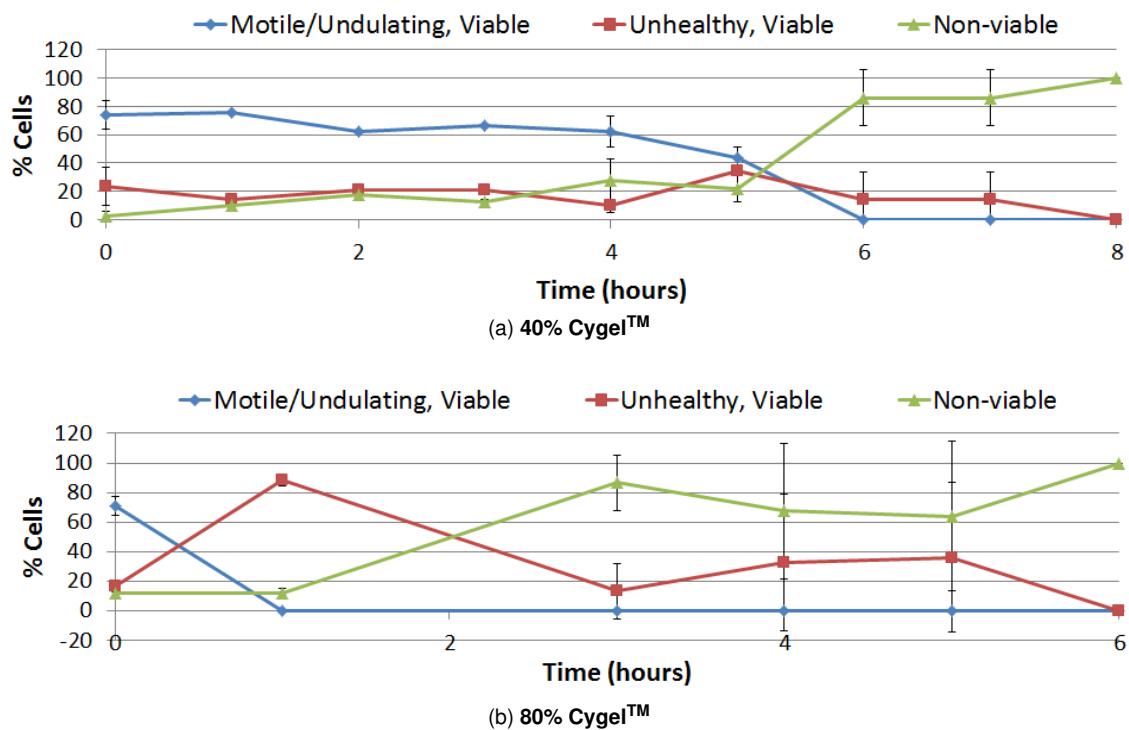


Figure 5.11: Cygel™ as an immobilisation agent: Plot showing the percentage of viable healthy cells, viable unhealthy (clumped/misshapen) cells and non-viable cells over time which were overlaid with a) 40% Cygel™ and b) 80% Cygel™. Error bars represent standard deviation. The experiment was carried out in duplicates.

were noted from the previous results.

In conclusion, the 40% Cygel™ immobilisation solution satisfied the three framework criteria of minimising motility while retaining cell viability and tractability. The solution had the best performance out of the investigated entrapping and immobilisation agents. The 40% Cygel™ set-up was taken forward for further screenings, as will be discussed next.

5.7.3 Optimal cell density of live *T. brucei* cells in Cygel™

The above investigations focused on the use of different materials, set-ups and technologies to establish an experimental framework by which live cell microscopy imaging of trypanosomes can be carried out while maintaining cell viability and tractability plus minimising cell motility. The following assay was developed to complement this framework by optimising the cell preparation method used prior to assembling the set-up. Specifically, the assay aims to determine the optimal density to which cells are resuspended during preparation. An optimal cell density selection improves tractability conditions and maximises the number of cells imaged while preventing overcrowding and the inability to distinguish between cells due to an overly dense population.

BSF cells were grown in 50ml media to 1×10^6 cells/ml, after which the culture was centrifuged for 10 minutes at 2000rpm, the supernatant removed and the pellet resuspended to a concentration of 10×10^7 cells/ml. From this, dilutions of 5×10^7 cells/ml, 2×10^7 cells/ml, 1×10^7 cells/ml and 5×10^6 cells/ml were prepared in media and embedded in 40% Cygel™. Afterwards, cells were

observed at 250X magnification and imaged with a Nexus 5 phone (Google Inc., CA, U.S.A). The images were used to take cell counts.

Figure 5.12 shows that cells at a density of 5×10^6 cells/ml (Figure 5.12a) and 1×10^7 cells/ml (Figure 5.12b) are too sparse, with less than 40 cells noted within each frame. Cells at a density of 2×10^7 cells/ml (Figure 5.12c), on the other hand, increase sharply in number from an average of 36 cells or less within the field of view to 235 ± 18 cells. For cells at a density of 5×10^7 cells/ml (Figure 5.12d), the average number of cells in the field of view was 393 ± 24 cells. For cells at a density of 10×10^7 cells/ml (Figure 5.12e), cells were dense and clumped together with large sparse patches in between, making it difficult to count and distinguish the cells from each other.

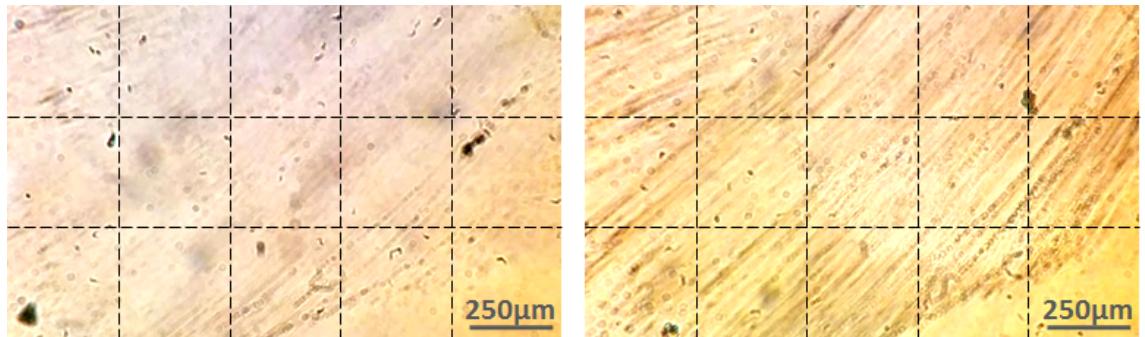
A non-linear increase between cell counts for cells at a density of 1×10^7 cells/ml and 2×10^7 cells/ml was noted. Further analysis was carried out to compare observed cell counts to expected counts. The expected number of cells within the field of view was calculated as follows:

- A well on a 5mm multiwell slide has a 5mm diameter and hence, total surface area is calculated to be 19.63mm^2 .
- When viewing the slide at a 250X magnification, the field of view of the above images was estimated to be roughly $0.99\text{mm} \times 1.67\text{mm}$ or 1.65mm^2 , as based on specifications of the SP-95-1 Inverted microscope (Brunel Microscope Ltd., Wiltshire, U.K.).
- Based on the above calculations, the field of view amounts to 8.4% of the total surface area of the well.
- The expected number of cells which should be within the field of view amount to 8.4% of the cell population pipetted.

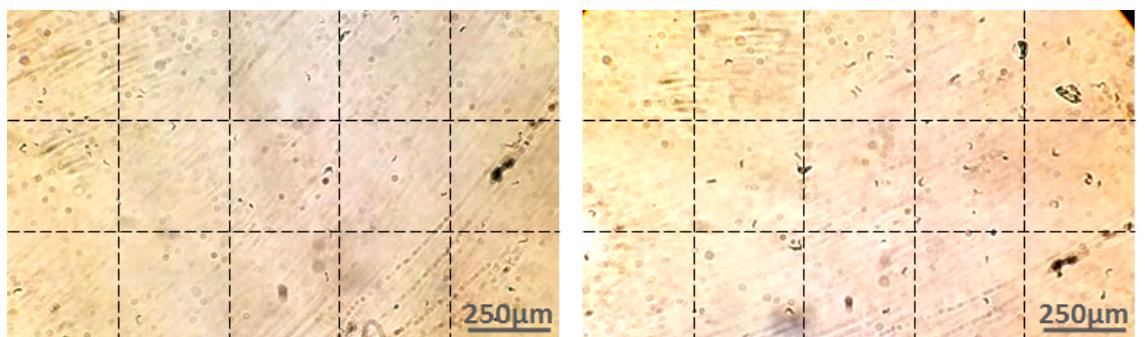
Calculations are based on a number of assumptions. Firstly, it was assumed that all volumes used were exact and no pipetting errors were made. Secondly it was assumed that the resuspended cell solution following centrifugation was homogeneous. Results are presented in Table 5.5, which shows a large difference between expected and observed cell counts. This can be attributed to a number of reasons. The pipetted volume is too large for the well and is pushed out when the coverslip is placed on to the slide. Moreover, Cygel™ has a higher viscosity than the cell solution, which means that it sinks to the bottom, resulting in a non-homogeneous Cygel™ and cell solution such that a larger proportion of the cell solution than Cygel™ is pushed out by the coverslip.

One solution is to mix cells with the Cygel™ solution prior to pipetting onto the slide. This will help to retain more cells within the well area. This method was adopted when imaging trypanosomes, as will be seen in the coming chapter. Pipetting a lower volume than $4\mu\text{l}$ of cell and Cygel™ solutions was not considered a viable option, since it leads to rapid drying of the well, compromising cell viability in less than an hour.

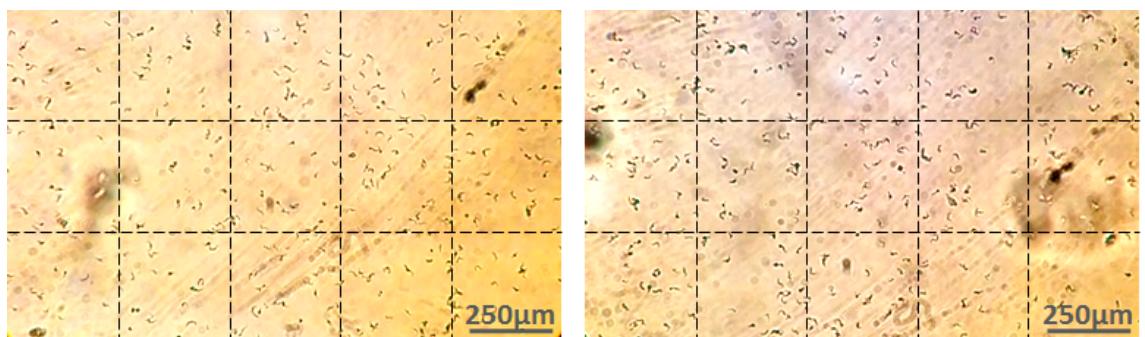
Following assessment of the optimal cell density of trypanosomes embedded in Cygel™, *T. brucei* at a concentration of 5×10^7 cells/ml were used for live cell microscopy efforts. This leads to a dense field of view but prevents cell clumping.



(a) Cells at a density of 5×10^6 cells/ml. Average cell count is 25 ± 1 cells.

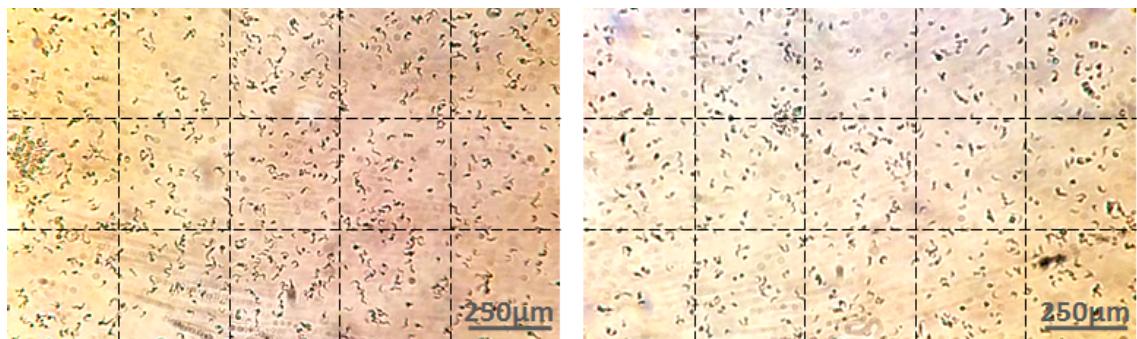


(b) Cells at a density of 1×10^7 cells/ml. Average cell count is 36 ± 4 cells.

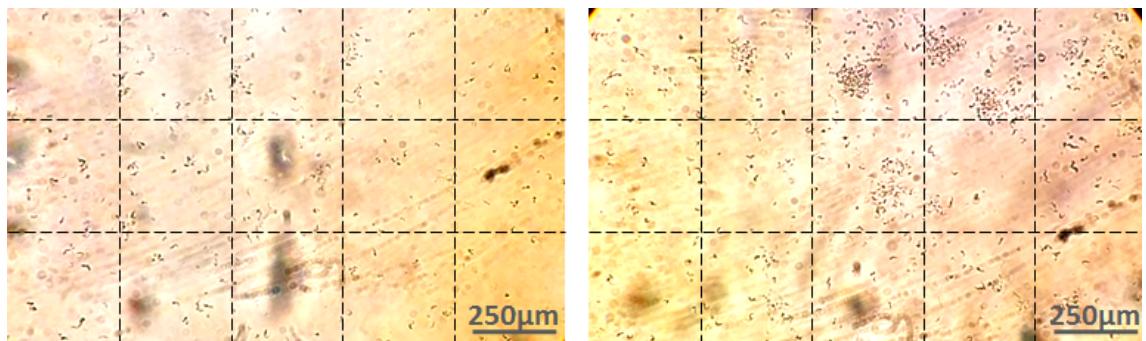


(c) Cells at a density of 2×10^7 cells/ml. Average cell count is 235 ± 18 cells.

Figure 5.12: Cell density screening assay: Images of live BSF *T. brucei* embedded in 40% Cygel™ at a concentration of a) 5×10^6 cells/ml, b) 1×10^7 cells/ml, c) 2×10^7 cells/ml. Cells were observed at 250X magnification. The experiment was carried out in duplicates. Scale bar = 250 μm.



(d) Cells at a density of 5×10^7 cells/ml. Average cell count is 393 ± 24 cells.



(e) Cells at a density of 10×10^7 cells/ml. Cells are clumped together. An accurate cell count of the number of cells within vision can not be carried out.

Figure 5.12: **(Continued)** Cell density screening assay: Images of live BSF *T. brucei* embedded in 40% Cygel™ at a concentration of d) 5×10^7 cells/ml, e) 10×10^7 cells/ml. Cells were observed at 250X magnification. The experiment was carried out in duplicates. Scale bar = $250 \mu\text{m}$.

5.8 Discussion

Following investigations to develop an experimental framework which enables the measurement of fluorescence levels in viable *T. brucei* cells, the following conclusions can be made:

1. Established an experimental framework using 40% Cygel™ to form an encasing and immobilising matrix among the cells. This set-up will be used in the microscopy imaging of live *T. brucei* cell populations bearing the stable and unstable Goodwin oscillators in order to analyse oscillations in fluorescence at the single cell level.
2. Out of the two sets of frameworks tested to either entrap or immobilise *T. brucei*, the use of Cygel™ satisfied the three framework criteria best. The experimental set-up was required to minimise cell motility while maximising tractability and maintaining cell viability. In the 40% Cygel™ set-up, *T. brucei* were viable for up to eight hours post-embedding and cell motility was observed to be reduced, thus allowing for increased cell tractability.
3. Established that the Cygel™ set-up can be optimised by using *T. brucei* at a concentration of 5×10^7 cells/ml. When looked at via the microscope this results in a dense field-of-view but prevented clumping of cells.

Density (cells/ml)	Number of cells pipetted	Expected number of cells within the field of view	Average number of cells observed within the field of view
5×10^6	10,000	840	25
1×10^7	20,000	1,681	36
2×10^7	40,000	3,361	235
5×10^7	100,000	8,403	393
10×10^7	200,000	16,807	N.A

Table 5.5: The cell density assay: Analysis of the expected and observed cell counts of *T. brucei* embedded in 40% CygelTM, depending on the cell density being used.

Chapter 6

Characterisation of the unstable Goodwin oscillator phenotype in *T. brucei* cells

6.1 Preliminary investigations of the unstable Goodwin oscillator

Following the transfection of BSF and PCF *T. brucei* with the unstable oscillator, pUnstable, preliminary characterisation was carried out in order to parameterise gene network components and assess whether they are functioning as expected. This, in turn, will inform both mathematical modelling and microscopy imaging. The studies implemented are outlined in Table 6.1. Analysis was primarily done using BSF *T. brucei* bearing the pUnstable network. Results are assumed to hold for PCF clones since the pUnstable GRN does not use mechanisms which are dependent on the form (BSF or PCF) of the trypanosome.

6.1.1 Growth profile of *T. brucei* cells bearing the unstable Goodwin oscillator

The first characterisation assay carried out on transfected *T. brucei* was to map a growth profile for BSF and PCF clones in order to analyse whether the unstable Goodwin oscillator plasmid im-

Assay	Section
1 Establish whether the pUnstable plasmid inhibits trypanosomal cell growth	Section 6.1.1
2 Establish functional activity of TetR regulator	Section 6.1.2
3 Determine the half-life of the TetR and eGFP protein and determine whether Ub-L moiety leads to destabilised TetR and eGFP protein	Section 6.1.3

Table 6.1: List of preliminary characterisation assays carried out to investigate function of the main components in the pUnstable oscillator in its *T. brucei* host.

pacts on cell growth. This is important because a slower cell growth compromises the potential of the host cell as a viable synthetic biology chassis, since the insertion of more complex plasmids would result in increasingly poor growth and unstable lines.

The assay was carried out over 72 hours on both induced and uninduced BSF and PCF trypanosomes, since *T. brucei* doubling times vary depending on the form (Lee et al., 2007). Wild-type *T. brucei* cells were used as a control. Initially, cells were diluted to the required density (1×10^6 cell/ml for BSF and 2×10^6 cells/ml for PCF cells) in 10ml of media. If required, $1 \mu\text{g}/\text{ml}$ tetracycline was added as an inducer. Cells were then left to incubate for 24 hours, after which cell counts were taken and the dilution-(induction)-growth-count process was repeated twice. The experiment was carried out in duplicates. The growth profile was plotted as both a saw-tooth curve with standard deviation error bars and as a cumulative growth curve of the mean growth.

6.1.1.1 Growth profile of BSF *T. brucei* cells bearing the unstable Goodwin oscillator

Following a 72-hour growth profiling, growth patterns of BSF pUnstable trypanosomes (Figure 6.1) do not show a clear distinction between wildtype and induced/uninduced transfected BSF *T. brucei*. This is also reflected in the corresponding estimated cell doubling times and exponential growth functions as calculated from the mean cumulative growth curves, seen in Figure 6.1c.

6.1.1.1.1 Use of Mann-Whitney statistical analysis for comparison of growth rate data. Following these observations, further in-depth analysis was carried out to assess whether the observations hold under statistical and scientific scrutiny. In order to evaluate whether the BSF cumulative growth rates are statistically significantly different from each other, the Mann-Whitney U test was implemented (Mann & Whitney, 1947). This test is used to compare two independent and non-parametric datasets (i.e. datasets which are not normally distributed) and deduce whether they are significantly different from each other, as seen in Wang and Bushman (2006), Vardi et al. (2001) and Hamamoto et al. (2015) who use the test to compare the growth patterns of different populations. The test works by rejecting one of two hypothesis which are set; the null hypothesis (H_0) which states that the difference between the median growth rates of the two populations is zero and the alternative hypothesis (H_1) which states that the difference is not zero.

In this case, the test was used to compare the median growth of induced or uninduced transfected *T. brucei* to wildtype cells. The following two sets of hypothesis were tested, to compare the median of cumulative growth of *uninduced transfected* cells to wildtype cells (Test 1) and to compare the cumulative growth of *induced transfected* cells to wildtype cells (Test 2):

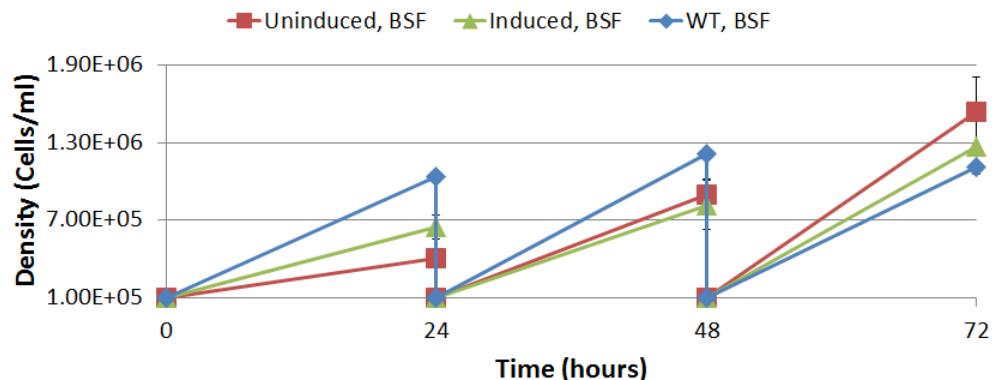
Test 1

H_0 : The difference in median between the cumulative growth of **BSF** wildtype cells and **uninduced** BSF cells transfected with pUnstable is **zero**

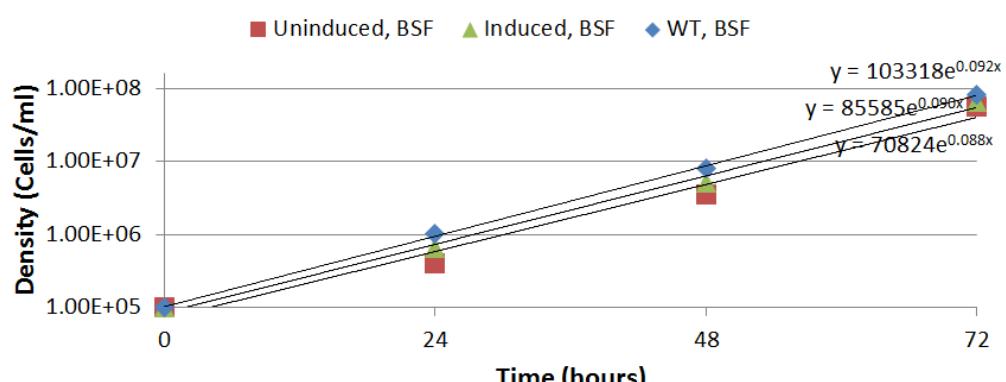
H_1 : The difference in median between the cumulative growth of **BSF** wildtype cells and **uninduced** BSF cells transfected with pUnstable is **not zero**

Test 2

H_0 : The difference in median between the cumulative growth of **BSF** wildtype cells



(a) Saw-tooth growth rate profile of BSF *T. brucei*



(b) Cumulative growth rate profile of BSF *T. brucei*. x is the time in hours.

Strain	Cell doubling time	Exponential fitted curve
Wildtype BSF	7.53 hrs	$y = 103318.e^{0.092t}$
Uninduced, transfected BSF	7.87 hrs	$y = 70824.e^{0.088t}$
Induced, transfected BSF	7.70 hrs	$y = 85585.e^{0.09t}$

(c) Cell doubling time for BSF *T. brucei*: Estimated cell doubling times and exponential fitted curves for wildtype and pUnstable-transfected BSF *T. brucei*. t is the time in hours.

Figure 6.1: Growth rate profiles and doubling times of induced/uninduced BSF *T. brucei* bearing pUnstable represented by a: a) saw-tooth curve profile, b) semi-logarithmic cumulative growth profile, c) doubling times and fitted exponential curves based on cumulative growth data. Induced cells were treated with 1 μ g/ml tetracycline. The plot axes show Density in cells/ml versus Time in hours. Error bars represent standard deviation. The experiment was carried out in duplicates.

and **induced** BSF cells transfected with pUnstable is **zero**

H_0 : The difference in median between the cumulative growth of **BSF** wildtype cells and **induced** BSF cells transfected with pUnstable is **not zero**

The test requires that the sample data from both populations is ranked in a single monotonic order. For each population, the ranks are added up and a test statistic U is calculated using the formula:

$$U = N_1 \cdot N_2 + N_L \cdot \frac{N_L + 1}{2} - T_L$$

where N_1 is the sample size of the first population, N_2 is the sample size of the second population, N_L is the number of samples in the population which had the largest sum of ranks and T_L is the largest sum of ranks from the two populations' sums. The U statistic is then compared to a critical U value, U_{crit} , the value of which depends on the population size. For this test to hold, there must be more than five data points per population (Sheskin, 2003). Given that the growth curve assay was carried out over 72 hours, only four cumulative growth data points were deduced from experiment data. In order to generate more values, a data point representing cell concentration at the 92nd hour was inferred from the fitted exponential curve for each track (Figure 6.1c), for a total of five data points, after which the Mann-Whitney U test was applied.

The data points used and test results are presented in Table 6.2 which shows that for Test 1, the U statistic was larger than $U_{crit}=2$, indicating that there is not enough evidence to reject the null hypothesis that the median growth rates of the wildtype and uninduced pUnstable BSF cells are not different from each other. A similar analysis was carried out for wildtype cells and induced pUnstable clones in Test 2 (Table 6.2). The U statistic was calculated to be 10.5 which is again bigger than $U_{crit}=2$, indicating there is not enough evidence to reject the hypothesis that growth profiles of wildtype BSF cells and induced pUnstable BSF clones are not different from each other. Hence, the pUnstable oscillator plasmid does not have a negative impact on BSF cell growth.

6.1.1.2 Growth profile of PCF *T. brucei* cells bearing the unstable Goodwin oscillator

Growth profiling was also carried for PCF *T. brucei* transfected with pUnstable over a 72-hour period. The growth curves for PCF wildtype and induced/uninduced transfected cells (Figure 6.2a and 6.2b) highlight a difference between the uninduced pUnstable-transfected cells and the remainder of the tracks. In fact, estimated doubling times (Figure 6.2c) have a wide range of values. In order to test whether these differences hold under statistical scrutiny, the following two hypothesis were tested using the Mann-Whitney U test:

Test 3

H_0 : The difference in median between the cumulative growth of **PCF** wildtype cells and **uninduced** PCF cells transfected with pUnstable is **zero**

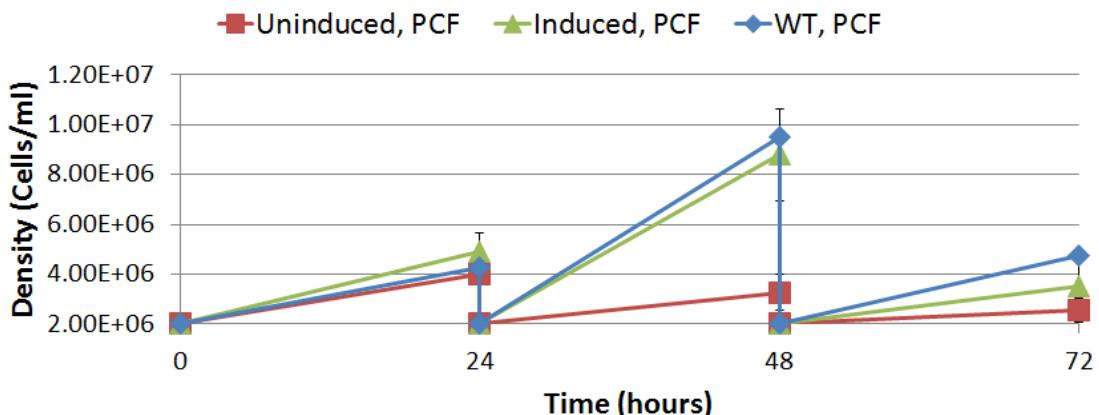
H_1 : The difference in median between the cumulative growth of **PCF** wildtype cells and **uninduced** PCF cells transfected with pUnstable is **not zero**

Test 4

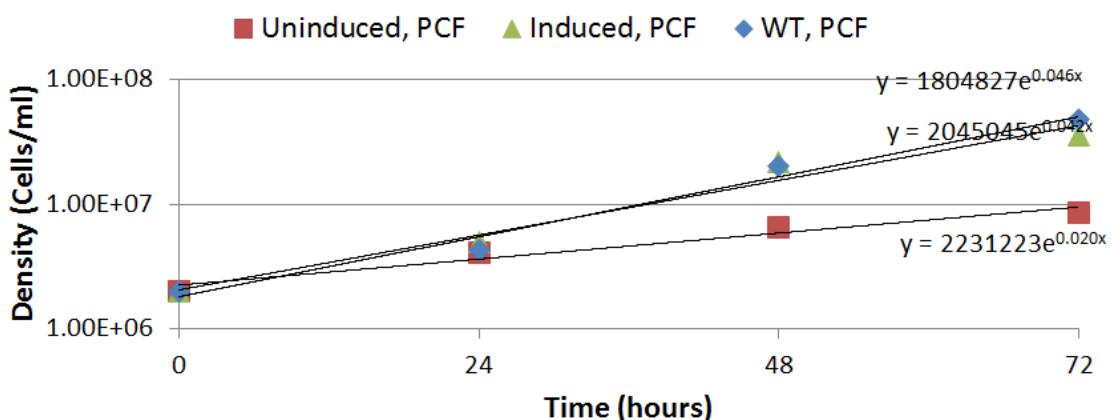
H_0 : The difference in median between the cumulative growth of **PCF** wildtype cells and **induced** PCF cells transfected with pUnstable is **zero**

Test 1: Comparing the cumulative growth of BSF wildtype cells with that of <i>uninduced</i> cells transfected with pUnstable		
Data	Wildtype (Cells/ml) 1×10^5 1.05×10^6 1.26×10^7 1.4×10^8 1.39×10^9 (Estimated)	Uninduced, transfected (Cells/ml) 1×10^5 4.05×10^5 3.62×10^6 5.58×10^7 3.3×10^8 (Estimated)
U statistic	11.5	
H_0 rejected	No	
Test 2: Comparing the cumulative growth of BSF wildtype cells with that of <i>induced</i> cells transfected with pUnstable		
Data	Wildtype (Cells/ml) 1×10^5 1.04×10^6 1.26×10^7 1.4×10^8 1.39×10^9 (Estimated)	Induced, transfected (Cells/ml) 1×10^5 6.45×10^5 5.26×10^6 6.68×10^7 4.84×10^8 (Estimated)
U statistic	10.5	
H_0 rejected	No	

Table 6.2: Comparison of growth rates of BSF *T. brucei* cells bearing the uninduced/induced unstable Goodwin oscillator expression cassette and unmodified wildtype cells via the Mann-Whitney U test, which tests whether cumulative growth rates for transfected cells are significantly different from wildtype BSF cells. The table shows the experiment and estimated data points of the cell counts in Cells/ml which were used for the analysis to calculate the U statistic. This is compared to U_{crit} at the 0.05% confidence level which has value 2. The table also prints the outcome whether to reject the Null hypothesis, H_0 .



(a) Saw-tooth growth rate profile of PCF *T. brucei*



(b) Cumulative growth rate profile of PCF *T. brucei*. *x* is the time in hours.

Strain	Cell doubling time	Exponential fitted curve
Wildtype PCF	15.07 hrs	$y = 1804827.e^{0.046t}$
Uninduced, transfected PCF	34.66 hrs	$y = 2231223.e^{0.02t}$
Induced, transfected PCF	16.50 hrs	$y = 2045045.e^{0.042t}$

(c) Cell doubling times and exponential fitted curves for PCF *T. brucei*: Estimated cell doubling times for wildtype and pUnstable-transfected PCF *T. brucei*. *t* is the time in hours.

Figure 6.2: Growth rate profiles and doubling times of induced/uninduced PCF *T. brucei* bearing pUnstable represented by a: a) saw-tooth curve profile, b) semi-logarithmic cumulative growth profile, c) doubling times and fitted exponential curves based on cumulative growth data. Induced cells were treated with 1 μ g/ml tetracycline. The plot axes show Density in cells/ml versus Time in hours. Error bars represent standard deviation. The experiment was carried out in duplicates.

Test 3: Comparing the cumulative growth of PCF wildtype cells with that of uninduced cells transfected with pUnstable		
Data	Wildtype (Cells/ml) 2×10^6 4.25×10^6 2.02×10^7 4.79×10^7 1.49×10^8 (Estimated)	Induced, transfected (Cells/ml) 2×10^6 4.03×10^6 6.59×10^6 8.4×10^6 1.15×10^8 (Estimated)
U statistic	9.5	
H₀ rejected	No	
Test 4: Comparing the cumulative growth of PCF wildtype cells with that of induced cells transfected with pUnstable		
Data	Wildtype (Cells/ml) 2×10^6 4.25×10^6 2.02×10^7 4.79×10^7 1.49×10^8 (Estimated)	Uninduced, transfected (Cells/ml) 2×10^6 4.9×10^6 2.15×10^7 3.77×10^7 1.52×10^7 (Estimated)
U statistic	11.5	
H₀ rejected	No	

Table 6.3: Comparison of growth rates of PCF *T. brucei* cells bearing the uninduced/induced unstable Goodwin oscillator expression cassette and unmodified wildtype cells via the Mann-Whitney U test, which tests whether cumulative growth rates for transfected cells are significantly different from wildtype PCF cells. The table shows the experiment and estimated data points of the cell counts in Cells/ml which were used for the analysis to calculate the U statistic. This is compared to U_{crit} at the 0.05% confidence level which has value 2. The table also prints the outcome whether to reject the Null hypothesis, H_0 .

H_1 : The difference in median between the cumulative growth of **PCF** wildtype cells and **induced** PCF cells transfected with pUnstable is **not zero**

Calculations of the U statistic from the Mann-Whitney test based on experimental and inferred data-points, as seen in Table 6.3, show that both null hypotheses, which state that the difference in median growth between wildtype and induced/uninduced transfected PCF trypanosomes is zero, should not be rejected, indicating that observed divergent growth patterns were outliers. Thus, similarly to BSF pUnstable-transfected cells, PCF pUnstable clones were found to have no impact on cell growth. Note that the cell doubling time for PCF cells is much larger than for BSF cells, in line with expectations (Pays, 1999).

6.1.2 Functional activity of the TetR regulation - Expression dynamics of the unstable Goodwin oscillator expression cassette in *T. brucei* cells following tetracycline addition and removal

This study focuses on the effect of tetracycline on pUnstable. In particular, the study establishes the length of time over which tetracycline remains effective as an inducer. This serves to confirm whether protein expression inhibition is regulated by TetR or otherwise, since protein expression

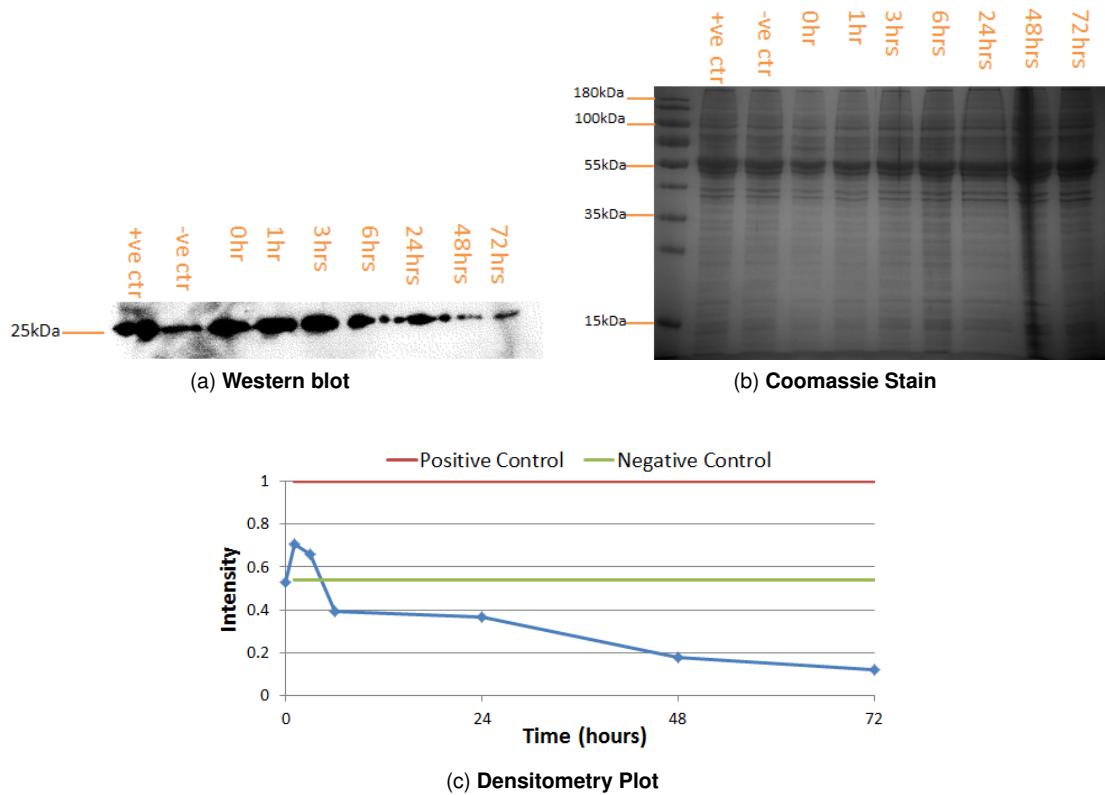


Figure 6.3: eGFP expression levels over a 72-hour period of BSF *T. brucei* bearing pUnstable following 1 μ g/ml tetracycline induction and removal. Analysis is carried out via a a) western blot detecting for the eGFP signal, b) corresponding coomassie stain and c) densitometry plot quantifying the amount of eGFP protein detected in the western blot using the positive control as a reference point of intensity 1.0. The positive control is a pConstitutive protein sample, while the negative control is an uninduced pUnstable protein sample.

should be suspended as the tetracycline molecule degrades, TetR binds to the operator instead of tetracycline and inhibition resumes.

1 μ g/ml tetracycline was applied to a BSF culture at a concentration of 1 \times 10⁵cells/ml and incubated until the population has grown to a density of 1 \times 10⁶cells/ml, after which cells were harvested at 2000rpm for 10 minutes, the supernatant removed and the pellet resuspended in an equal volume of pre-warmed media. The harvest, removal of supernatant and resuspension process was then repeated. This step removes extracellular tetracycline while intracellular tetracycline remains and diffuses back out of cells at an unknown rate. Following resuspension, samples were taken and protein prepared at the following time-points: 0hr, 1hr, 3hr, 6hr, 24hr, 48hr and 72hr. Samples were analysed via a western blot detecting for the eGFP protein and a corresponding coomassie stain.

Results, as seen in Figure 6.3, show that following the removal of extracellular tetracycline, the fluorescent signal from the pUnstable GRN loses strength over time. In particular, there is a pronounced decrease in signal strength following the three hour time-point with another significant decrease in signal strength between the 24 and 48 hour time-point. This serves to confirm the functional activity of the TetR regulator.

6.1.3 Half-life of the destabilised TetR and eGFP proteins in *T. brucei* cells bearing the unstable Goodwin oscillator

The next assay to be carried out was a cycloheximide-based assay to determine the half-lives of the destabilised TetR and eGFP proteins and input the data into mathematical modelling. In addition, knowledge of the eGFP half-life will help inform whether the protein is sufficiently destabilised as to exhibit oscillations due to TetR inhibition or whether it is too stable to reflect regulatory influence within the network.

Cycloheximide serves to disrupt translation processes and hence, synthesis. However, degradation reactions remain unaffected. Thus, following the addition of cycloheximide to a sufficiently dense BSF culture, the number of TetR and reporter proteins within the cells remain constant except for those which are degraded via natural processes. It follows that the period over which the protein signal in strength decreases by half corresponds to the protein half-life. The assay was carried out using transfected BSF cells which had been induced with tetracycline for a 24-hour period so as to counteract inhibitory regulation and allow for maximal constitutive protein expression to occur. This implies that decrease in protein signal reflects only degradation processes and not oscillatory cycles.

6.1.3.1 Half-life of the destabilised, Myc-tagged TetR protein in *T. brucei* cells bearing the unstable Goodwin oscillator

In the pUnstable design, a Myc tag was added upstream of the TetR protein to quantify TetR indirectly and avoid using anti-TetR antibodies. Results from the western blot detecting the Myc-tag were blank (results not shown). The positive control used in the western blot, however, did flag up a signal, indicating that the blot was performed correctly. The coomassie stain showed that samples were prepared correctly. The lack of Myc-tag detected is unlikely to be due to no Myc protein being synthesised. The difference in signal strength between induced and uninduced samples (Figure 4.24, Figure 6.3) indicate that tetracycline is binding to TetR protein and, by proxy, that TetR protein is being synthesised. It follows that the Myc tag, which is within the same cassette as TetR and shares the same start and stop codons, is also being expressed.

Note that detecting for the Ub-L protein which is upstream of TetR would flag up Ubiquitin signals from the organism itself and not just the pUnstable plasmid. The use of TetR-binding antibodies proved ineffective, as will be discussed in more detail in Section 7.1.4.1. No conclusions were made regarding the half-life of TetR and the destabilising effect of Ub-L.

6.1.3.2 Half-life of the destabilised eGFP protein in *T. brucei* cells bearing the unstable Goodwin oscillator

In order to determine the half-life of the Ub-L destabilised eGFP protein in pUnstable, a western blot and coomassie stain were run using cycloheximide-treated protein samples, with results as in Figure 6.4 indicating that the destabilised eGFP protein has a half-life of 2-3 hours. A clear difference can be seen between eGFP signals of cycloheximide-treated and untreated samples (Figure 6.4a and 6.4b respectively), indicating that the change in the treated samples is either due to the

effect of cycloheximide or incorrect protein samples. However, the corresponding coomassie stains show consistent samples, indicating that the decrease in protein quantity detected in Figure 6.4a is due to the effect of cycloheximide rather than incorrect sample preparation. Due to basal expression levels within pUnstable this data was not plugged into mathematical modelling and instead, data gathered from pStable was used. Further investigation concerning the eGFP half-life is discussed in Section 7.1.4.2.

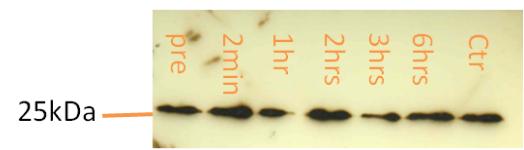
6.1.4 Evaluation of the unstable oscillator phenotype following preliminary investigations

Following the above phenotyping assays, the following findings and observations were made:

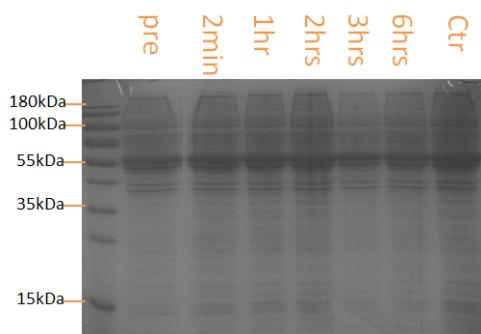
1. Preliminary characterisation of the pUnstable oscillator via western blots and coomassie stains has shown that the main components of the pUnstable network are functional and that TetR and eGFP protein is successfully synthesised. Promoters successfully initiate transcription: detected eGFP levels indicate that the P_{RRNA} promoter is functional while clones cultured in the presence of Hygromycin indicate that the P_{EP1} promoter is also working. There is a basal constitutive expression such that the pUnstable network can be used as a partial replacement for tetracycline molecules since it allows for TetR and eGFP continuous expression.
2. BSF *T. brucei* transfected with pUnstable and induced with tetracycline flag up a stronger eGFP protein signal than uninduced transfected *T. brucei*, indicating that TetR regulation is being implemented and that the tetracycline operator is functional.
3. Intracellular tetracycline can induce maximal expression within the plasmid for over three hours, after which inhibition resumes as the tetracycline molecules degrade.
4. The pUnstable plasmid does not impair cell growth in transfected BSF or PCF *T. brucei*.
5. The Ub-L gene upstream of eGFP has a destabilising effect. The half-life of eGFP within pUnstable is c. 3 hours long when it usually is over 24 hours long (Corish & Tyler-Smith, 1999). Quantitative mathematical modelling work concerning this will be discussed later.
6. The Ub-L gene was added upstream of the *tetR* gene in order to destabilise the TetR protein, while the Myc tag was added upstream of this fusion in order to quantify TetR protein more easily when using western blots. Further investigation is needed to determine how Myc and Ub-L are affecting TetR.
7. In conclusion, the components of the pUnstable GRN allow for oscillations to take place. Since the cell population is unlikely to oscillate in synchrony (Elowitz & Leibler, 2000; Stricker et al., 2008; Danino et al., 2010; Tigges et al., 2010) and since western blots have limited sensitivity, oscillations are unlikely to be detected via protein sample analysis. Therefore fluorescent microscopy imaging was carried out in order to observe fluorescent patterns, as will be discussed next.



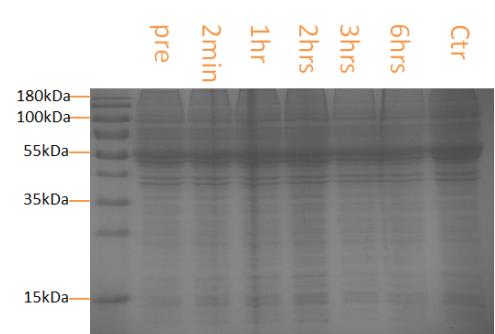
(a) Western blot detecting for eGFP in protein samples from induced pUnstable-transfected BSF *T. brucei* **treated** with cycloheximide



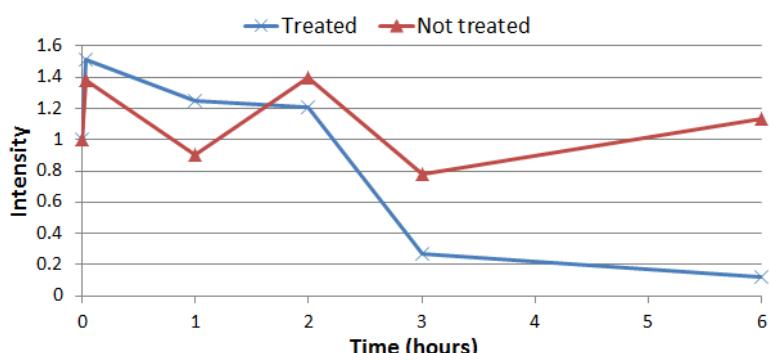
(b) **Control:** Western blot detecting for eGFP in protein samples from induced pUnstable-transfected BSF *T. brucei* **not treated** with cycloheximide



(c) Coomassie stain for protein samples from induced pUnstable-transfected BSF *T. brucei* **treated** with cycloheximide



(d) **Control:** Coomassie stain for protein samples from induced pUnstable-transfected BSF *T. brucei* **not treated** with cycloheximide



(e) Densitometry plot showing the signal intensity extracted from the western blots above of eGFP protein from pUnstable-transfected cells either treated or untreated with cycloheximide.

Figure 6.4: Deducing the half-life of the **destabilised eGFP protein** in BSF *T. brucei* bearing pUnstable and induced with $1\mu\text{g/ml}$ tetracycline. Protein samples were prepared at specific time-points following cycloheximide addition and then analysed via western blot (a) detecting for eGFP and a corresponding coomassie stain (c). In parallel, a control assay (b, d) was carried out wherein no cycloheximide was added to the culture. e) shows a densitometry plot quantifying the amount of eGFP in both treated and untreated pUnstable-transfected cells detected in the western blots. The plot uses the eGFP intensity from the measurement taken prior to the addition of the cycloheximide or blank ('pre') as value 1.0.

Strain	Plasmid introduced	Conditions	Section
1	BSF	pConstitutive	-
2	BSF	pUnstable	Uninduced
3	BSF	pUnstable	Induced with tc_{thres}
4	PCF	pUnstable	Uninduced
5	PCF	pUnstable	Induced with tc_{thres}

Table 6.4: The pConstitutive and pUnstable-transfected BSF and PCF *T. brucei* clones which were imaged using live-cell fluorescent and brightfield microscopy to observe fluorescent patterns and characterise the phenotype of pUnstable. ‘ tc_{thres} ’ is the tetracycline concentration above which protein expression across the cell population is induced.

6.2 Characterisation of the phenotype of the unstable Goodwin oscillator

Following assessment of the main pUnstable components, the next step was to observe fluorescent patterns in pUnstable-transfected BSF and PCF *T. brucei* populations via live-cell fluorescent microscopy. Cells were immobilised in 40% Cygel™ at a density of 5×10^7 cells/ml as per Section 5.7. Microscopy was carried out in two hour segments at two minute intervals, using an objective of 20X. Brightfield microscopy was used as a control, with fluorescent and brightfield images taken consecutively. This was done firstly to confirm that observed fluorescence peaks are a result of protein dynamics within the cells rather than debris and secondly so as to facilitate manual cell tracking.

After imaging, fluorescent background correction was applied uniformly across all image stills. The relative fluorescence between stills was not affected. Images of the cell populations were then collated and analysed as a video in order to detect and quantify changes in fluorescence levels over time in individual cells. Quantification was initially done via Volocity, which uses a combination of thresholding (Pal & Pal, 1993), edge detection (Gonzalez & Woods, 2002) and watershed segmentation (Beucher & Lantuejoul, 1979, September) algorithms in order to select fluorescing cells (Hand et al., 2009). After, each quantified cell was tracked manually over the stills to check none had been misidentified. This was necessary due to an element of cell motility which can lead to incorrect cell tracking by Volocity. Table 6.4 lists the strains which were imaged and the conditions under which they were imaged.

6.2.1 Phenotype of pConstitutive cells in BSF *T. brucei*

Fluorescent microscopy imaging was carried out using pConstitutive control cells to analyse whether eGFP expression levels were detectable with fluorescent microscopy. Imaging showed that all observed cells were producing eGFP protein constitutively at detectable levels as seen in Figure 6.6 which shows stills of fluorescing pConstitutive cells at intermittent time intervals.

The next step was to quantify the effect of cell undulations on the strength of the fluorescent signal. Cells can undulate across the plane or within a fixed spot (Ralston et al., 2009), leading to changes in focus and hence, fluorescent intensity. Three cells were selected and their fluores-

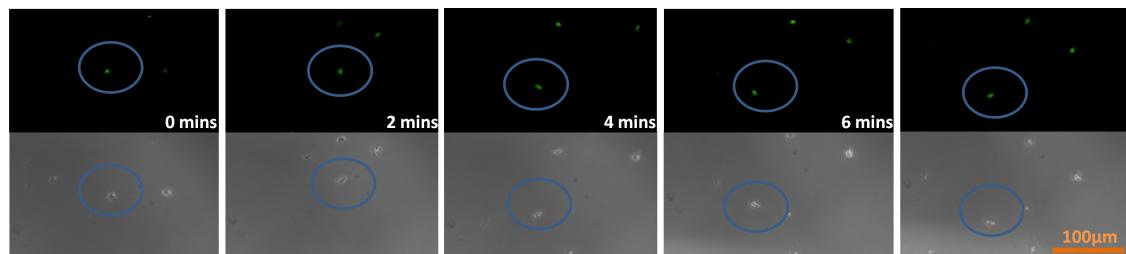


Figure 6.5: Images of pConstitutive-transfected BSF *T. brucei* over five consecutive time-points taken using a 20X objective. The top row shows fluorescent microscopy images while the bottom row shows corresponding brightfield microscopy images. The encircled cell is pConstitutive Cell 1 which was used to quantify variation in the fluorescent profile due to cell undulations. Scale bar = 100 μ m.

cence quantified over five time-points. Readings were taken over five time-points because cell motility meant that the cells moved out of the plane of vision after this period. As an example, Figure 6.5 shows brightfield and fluorescent images of pConstitutive Cell 1, one of the three cells used for this analysis. The mean and standard deviation of the fluorescent signal intensity within each cell were calculated as per Table 6.5. The range of mean fluorescent signals (32.6-74.3) in cells is attributed to cell-cell variability, which is a variation in the phenotypic output (for example, protein synthesis) of monoclonal cells due to intrinsic noise or deterministic processes (Snijder & Pelkmans, 2011; Pelkmans, 2012).

Although in theory, BSF cells bearing pConstitutive have a uniform fluorescent signal, the observed and quantified variation in signal strength (Table 6.5) indicates that cell undulations and intracellular sources of stochasticity are having an impact on the detected signal. The percentage of standard deviation varies between cells (2-28%). This results in an average 17% variation. Similar levels of GFP variation (up to 20%) were detected in control cells in Tigges et al. (2009). Although the effects of internal processes on cell output (Haupts et al., 1998; Vilar et al., 2002) are well documented, no information was found on the effect of cell motility (and microscope focus) on the detected output, making it difficult to distinguish between the effect of noise and motility.

In a final observation made, it was noted that photobleaching, temperature and cell environment during fluorescence microscopy affected trypanosomes differently when compared to observations made when using optical microscopy in Chapter 5. Specifically, the fluorescent microscopy environment led to decreased cell viability and increased motility.

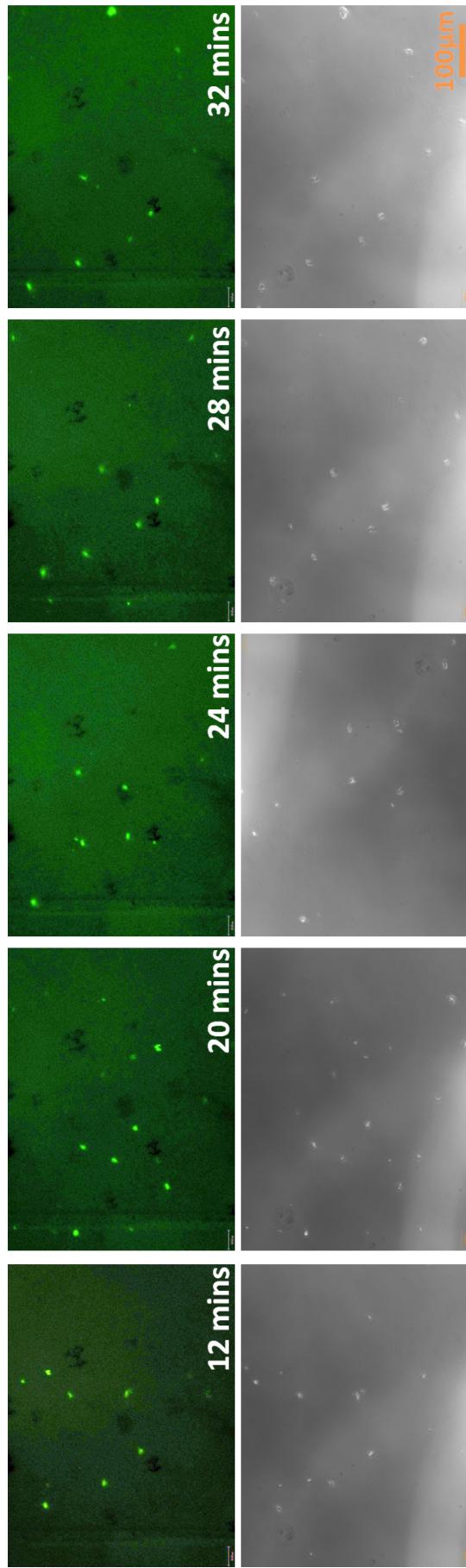


Figure 6.6: Images of *T. brucei* cells bearing the pConstitutive plasmid using an objective of 20X and taken at different time-points. The top row shows images taken using fluorescent microscopy while the bottom row shows corresponding brightfield images. Fluorescent images have been background corrected. Scale bar = 100 μ m.

Cell	Mean fluorescent signal	% standard deviation
pConstitutive Cell 1	73.84	22
pConstitutive Cell 2	74.29	2
pConstitutive Cell 3	32.57	28
Average	60.23	17

Table 6.5: Mean fluorescent signal intensity and percentage of standard deviation of three pConstitutive-transfected BSF *T. brucei* cells over five time-points. The fluorescent signal has no units.

6.2.2 Phenotype of the unstable Goodwin oscillator in BSF *T. brucei*

6.2.2.1 Imaging of uninduced BSF *T. brucei* bearing the unstable Goodwin oscillator

pUnstable-transfected BSF cells were imaged as per the above method to observe fluorescent patterns and establish whether pUnstable-transfected cells were producing eGFP protein periodically. Imaging and subsequent analysis showed cells to be fluorescing at a basal level, as seen in Section 6.1. No significant changes were detected in the fluorescence intensity levels (results not shown) over the two hour span. This is attributed in part to cell motility and low cell density, making for limited cell tracking. This will be further discussed in the coming chapter. Throughout, brightfield microscopy confirmed that cells were not shrunk or burst.

6.2.2.2 Imaging of BSF *T. brucei* bearing the unstable Goodwin oscillator and induced with tc_{thres} to increase the probability of detecting oscillations

A second image acquisition study was carried out on pUnstable-transfected BSF cells wherein, prior to microscopy, cells were induced with tetracycline at the threshold tetracycline concentration, $tc_{thres} = 10\text{fg/ml}$ for 24 hours. This is the threshold concentration such that when cells are induced with tc_{thres} , only a proportion of cells fluoresce. If used at a higher concentration, tetracycline induces detectable expression across the whole population. See Section 7.1.2 for the original assay. Following this, cell preparation and imaging was as described above. This was done in order to control for the possibility that changes in fluorescence were too weak to be detected via microscopy and quantitation.

The hypothesis behind the addition of tetracycline works as follows, with reference to Figure 6.7. When no tetracycline (Tetracycline = 0) is added to the culture of transfected *T. brucei* cells, TetR and eGFP protein is synthesised. TetR then binds to the operator, Tet Op, resulting in inhibition. As TetR degrades via natural processes, the cycle restarts itself. This oscillatory cycle has been described in more detail in Section 4.1.2. Due to low eGFP quantity, fluorescence is not detected by microscopy. On the other hand, if the culture of transfected *T. brucei* is induced with a high concentration of tetracycline (Tetracycline $> tc_{thres}$), TetR and eGFP is synthesised and TetR binds to the tetracycline molecules, to which it has a greater affinity than Tet Op. This results in constitutive expression until the tetracycline molecules are degraded. Thus, no oscillations occur during this period.

Finally, if tetracycline is added at a concentration of tc_{thres} (Tetracycline = tc_{thres}), tetracycline molecules applied to the cell culture are dispersed among the cells and a proportion of them becomes intracellular (Step 1 in bottom box in Figure 6.7). Thus, when expression occurs (Step 2) and a TetR protein is synthesised, it binds to the tetracycline molecules present instead of the operator (Step 3), thus allowing for further expression of both *tetR* and eGFP genes and an increase in fluorescent levels (Step 4). This process is repeated until the number of TetR proteins exceeds the number of tetracycline molecules (Step 5).

At this stage, TetR binds to the tet operator and expression is inhibited (Step 6). This is followed by a period where TetR and eGFP proteins degrade due to natural processes (Step 7), after which the oscillatory cycle re-initiates. Thus, tetracycline at the tc_{thres} concentration serves to increase fluorescence levels, the amplitude of oscillations and in turn the probability of observing oscillations via microscopy. Although the above explanation does not take basal levels of transcription into account, the hypothesis still holds.

Following imaging of tc_{thres} induced pUnstable-transfected BSF *T. brucei*, quantified cells showed no significant change in fluorescence intensity (results not shown). Similarly to results for uninduced pUnstable-transfected BSF cells, trypanosomes retained an element of motility, making for limited tractability. This does not disprove the hypothesis behind adding tetracycline to the culture, but is an indication that further cells need to be imaged under increased immobilisation conditions.

6.2.3 Phenotype of the unstable Goodwin oscillator in PCF *T. brucei*

Apart from pUnstable-transfected BSF cells, induced and uninduced pUnstable-transfected PCF clones were also imaged over a two hour segment in order to observe fluorescence patterns. Fluorescence was not detected in individual cells (images not shown). This reflects findings in transfection results in Section 4.4 wherein western blot results did not flag up eGFP signals in pUnstable-transfected PCF clones, even when induced. Further studies, including assessing the point of insertion of the plasmid on the genome following transfection, are required to analyse these results.

6.2.4 Evaluation of the unstable oscillator phenotype following live cell microscopy imaging

Following the imaging of *T. brucei* cells having the pConstitutive and pUnstable plasmids, the following results and observations were made:

1. BSF *T. brucei* transfected with pConstitutive express eGFP continuously and at detectable levels. The fluorescence signal varies by an average of 17% over five readings due to cell motility, undulations and intracellular stochasticity. This will be used when analysing changes in fluorescence levels in cells having the pUnstable or pStable plasmid, to help indicate whether observed fluctuations were true oscillations or noise. However, the imaging of pConstitutive transfects over a time span of two hours would allow for better comparability

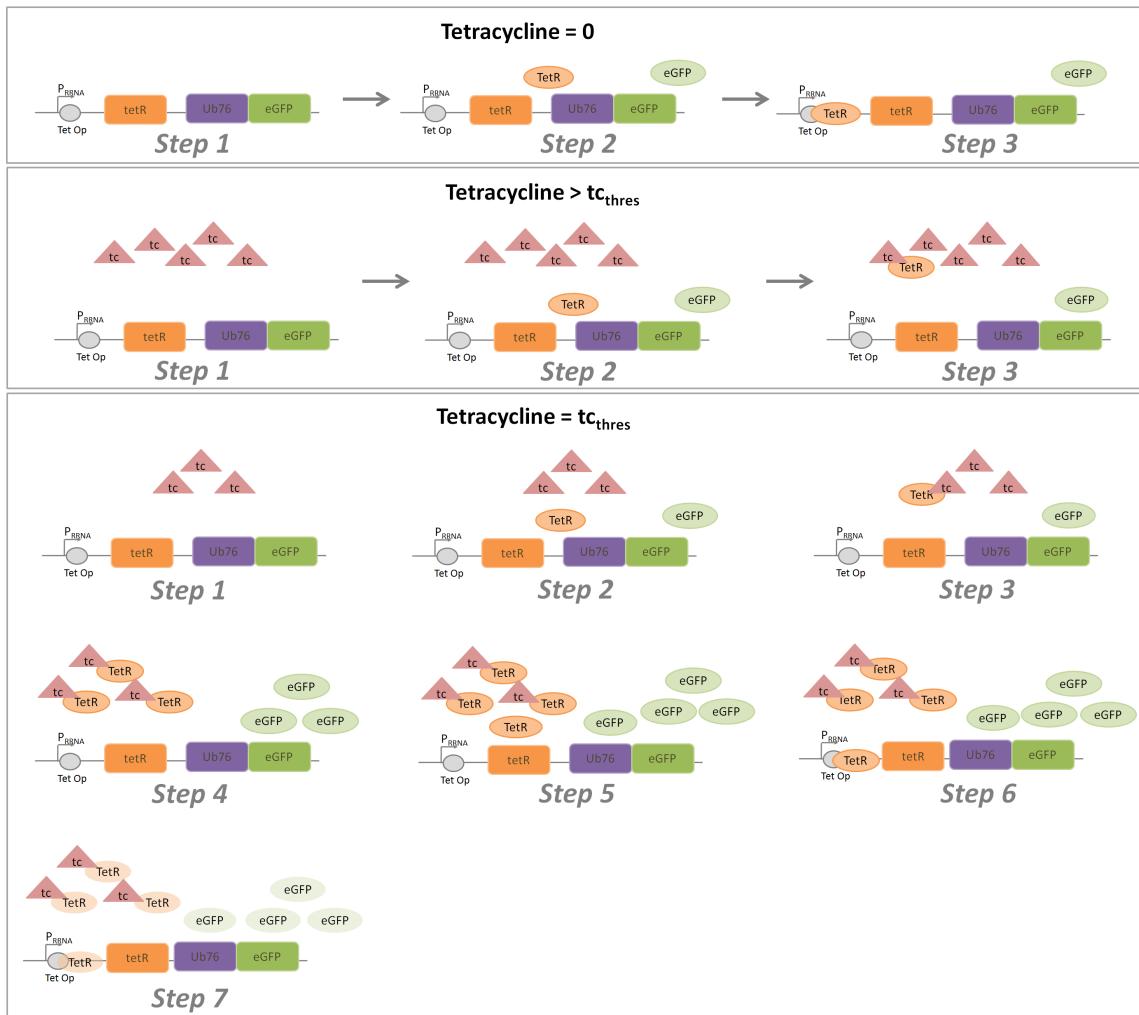


Figure 6.7: Diagram explaining how different concentrations of tetracycline molecules affect protein dynamics in *T. brucei* cells bearing the Goodwin oscillator, leading to either undetected fluorescence oscillations, detected fluorescence oscillations or increasing fluorescence levels. P_{RRNA} is the promoter, $Tet Op$ is the tetracycline dependant operator, $tetR$ is the $tetR$ gene, $Ub-L-eGFP$ is the Ub-L and eGFP gene moiety, $TetR$ is the TetR protein, $eGFP$ is the eGFP protein, tc is tetracycline and tc_{thres} is the threshold tetracycline concentration.

Tetracycline = 0: When no tetracycline is added to the culture (Step 1), TetR and eGFP protein is expressed at time 0 (Step 2). TetR binds to $Tet Op$, resulting in protein synthesis inhibition (Step 3). TetR and eGFP degrade naturally (not shown), after which the cycle re-starts. Throughout, eGFP levels are not sufficiently high to be detected.

Tetracycline > tc_{thres} : When tc at a higher concentration than tc_{thres} is added to the culture (Step 1), TetR and eGFP is expressed at time 0 (Step 2). TetR binds to tc , towards which it has a greater affinity than $Tet Op$ (Step 3). Due to an excess of tc , all TetR proteins bind to it, preventing inhibition and resulting in increasing fluorescence levels but no oscillations.

Tetracycline = tc_{thres} : When tc at a concentration of tc_{thres} is added to the culture (Step 1), TetR and eGFP is expressed at time 0 (Step 2). TetR binds to tc (Step 3). This continues until all tc is bound (Step 4), at which stage TetR binds to $Tet Op$ (Step 5,6) and inhibits further synthesis. TetR and eGFP degrade via natural processes (Step 7), after which the oscillatory cycle re-starts. The increase in eGFP quantities prior to inhibition increases the probability that fluorescence levels are detected.

of results to pStable and pUnstable transfects, wherein imaging is taken over two hour segments. Thus, results can only be used as an indication and it is recommended that further imaging of pConstitutive is carried out.

2. Imaging of induced and uninduced *T. brucei* bearing pUnstable showed non-significant changes in fluorescence signals, in part due to motility issues and due to the PCF pUnstable clones not issuing a signal. The optimisation of the imaging set-up will go a long way towards controlling this issue as it would enable the imaging of a larger number of transfected clones at lower motility.
3. pUnstable components are functioning and interacting as expected, as seen in characterisation assays carried out on BSF and PCF *T. brucei* transfectants. Moreover, simulations of protein dynamics (discussed further on) indicate that the GRN dynamics are fluctuating. This indicates that observations from microscopy imaging do not eliminate that pUnstable functions as an oscillator and that further work should be carried out.

6.3 Discussion

Following characterisation of the unstable Goodwin oscillator phenotype in *T. brucei* cells via protein analysis and microscopy imaging, the following conclusions can be made:

1. Characterised a novel synthetic GRN encoding an unstable Goodwin oscillator within a *T. brucei* chassis. This is the first synthetic oscillator to be characterised within a trypanosomal platform.
2. Assessed functionality of the main pUnstable network components. Promoters and operators were functional and the network successfully synthesised destabilised eGFP, TetR and antibiotic resistant protein.
3. Deduced that the pUnstable network has a basal level of *tetR* and Ub-L-eGFP transcription. This implies that the pUnstable network can be used in alternative applications to pStable where basal transcription levels are required or as an alternative to induction by low levels of tetracycline. This results in a more efficient and feasible process while still allowing for oscillations to occur.
4. Confirmed that the pUnstable network is regulated via autonomous TetR regulation which can be reversed via the addition of tetracycline. When added at a 1 μ g/ml concentration, tetracycline has a lasting effect of over three hours.
5. Established that the pUnstable network does not impact on either BSF or PCF cell growth. This is important in the assessment of *T. brucei* as a synthetic biology platform as it shows that the cells retain stability following the integration of extracellular plasmid.
6. Characterised the pConstitutive plasmid phenotype in BSF *T. brucei* via microscopy imaging and quantitation. Although based on a limited data set, cells were shown to express eGFP constitutively.
7. Quantified variation in fluorescent signal due to cell undulations and intracellular stochasticity over a ten minute time-span as 17% of mean observed fluorescent signal.
8. Imaged and analysed induced and uninduced BSF and PCF *T. brucei* bearing the unstable Goodwin oscillator. Fluorescent oscillations were not observed under the conditions studied. This does not exclude pUnstable from functioning as an oscillator, especially since

preliminary analysis and mathematical modelling have indicated that the GRN can function as an oscillator (irrespective of whether it is genuine or stochastic). Further development of the experimental set-up and imaging technology will enable a wider set of conditions to be scanned for observable oscillatory behaviour.

Chapter 7

Characterisation of the stable Goodwin oscillator phenotype in *T. brucei* cells via experimentation and mathematical modelling

Following characterisation of the pUnstable plasmid, the next step was to characterise the pStable oscillator (Figure 7.1) phenotype in *T. brucei* cells. This involved parameterising gene network components and dynamics via protein analysis and microscopy imaging and the development of a mathematical model for the synthetic trypanosomal Goodwin oscillator.

7.1 Preliminary investigation of the stable Goodwin oscillator

Prior to the imaging of transfected *T. brucei* clones, characterisation was carried out to analyse function of the main components of the pStable plasmid. An overview of these studies can be found in Table 7.1. Analysis was mainly carried out on BSF *T. brucei* bearing the pStable network. Results are assumed to hold for PCF clones since the pStable GRN does not use mechanisms

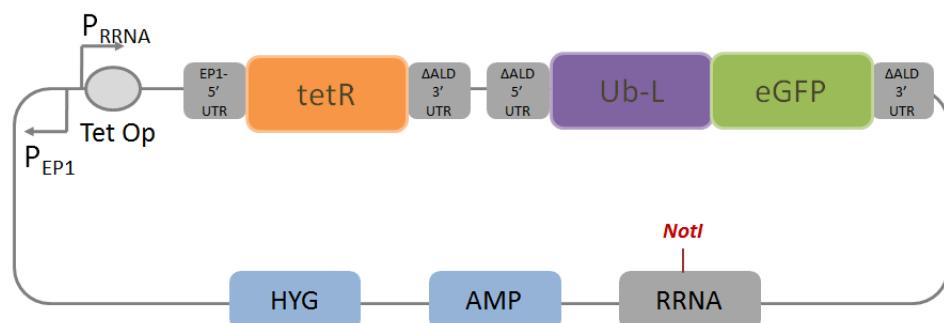


Figure 7.1: Diagram showing the stable Goodwin oscillator, pStable. The main components within the network are the P_{RRNA} and P_{EP1} promoters, the *tetR* cassette, the *Ub-L*-eGFP fusion gene, the tetracycline operator Tet Op; the 5' and 3' UTRs for both *tetR* and *Ub-L*-eGFP; the Hygromycin (Hg) and Ampicillin (Amp) selectable markers and the rRNA gene targeting fragment having the *NotI* restriction site.

Assay	Section
1 Establish whether the pStable plasmid inhibits trypanosomal cell growth	Section 7.1.1
2 Establish the minimum tetracycline induction threshold for pStable	Section 7.1.2
3 Establish functional activity of TetR regulator	Section 7.1.3
4 Determine the half-life of the TetR and eGFP protein and determine whether Ub-L-eGFP moiety leads to destabilised eGFP protein	Section 7.1.4

Table 7.1: List of preliminary characterisation assays carried out to investigate function of the main components of the pStable oscillator network in transfected *T. brucei*.

which are dependent on the form (BSF or PCF) of the trypanosome.

7.1.1 The growth profile of *T. brucei* transfected with the Goodwin oscillator

The first study was a growth assay to determine whether trypanosomes transfected with pStable have stunted growth when compared to wildtype *T. brucei*. The assay was carried out as per Section 6.1.1, using wildtype *T. brucei* as a control. The experiment was carried out in duplicates and the growth profile represented in both the saw-tooth and cumulative growth curve format.

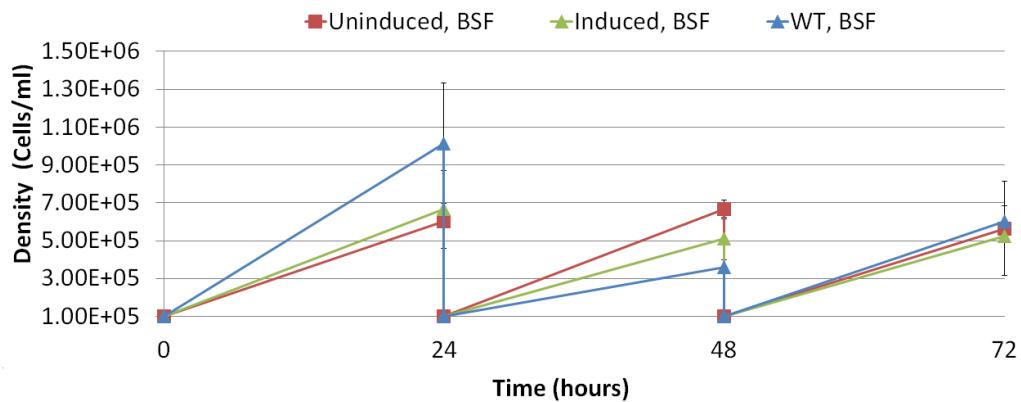
7.1.1.1 Growth profile of BSF *T. brucei* cells bearing the stable Goodwin oscillator

Following a 72-hour growth profiling, the saw-tooth growth curve (Figure 7.2a) shows no distinct patterns for the growth rate of the three different BSF tracks (uninduced pStable-transfected clone, induced pStable-transfected clone, control wildtype). However, a significant amount of variety is observed between the duplicate counts within each track with the standard deviation being as large as 44% of the mean cell count (see the 72 hour count, uninduced cell line). The estimated cumulative growth curves (Figure 7.2b) and cell doubling time estimates (Figure 7.2c) of the *mean* values of the three tracks also indicate undistinctive growth patterns.

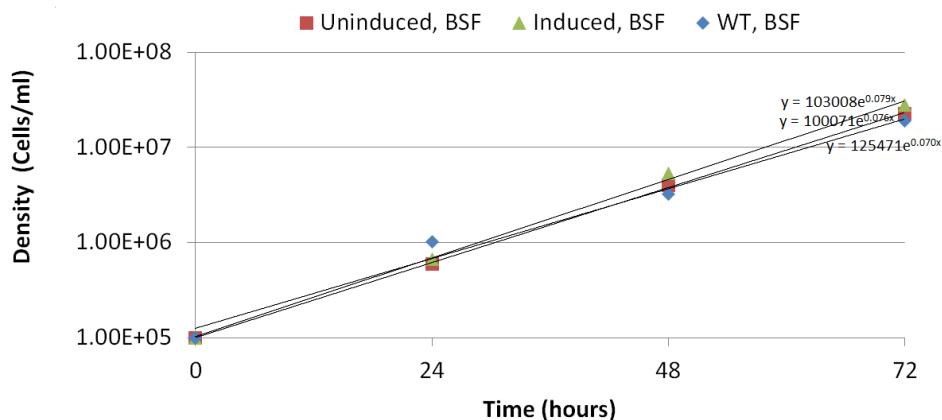
7.1.1.1.1 Use of Mann-Whitney statistical analysis for comparison of growth rate data. In order to test whether these observations hold under mathematical scrutiny, the Mann-Whitney U test was used to compare the wildtype growth patterns to growth patterns of induced/uninduced pStable-transfected trypanosomes. The test was previously described and used in Section 6.1.1.1.1 for the analysis of growth rates of *T. brucei* cells having the unstable Goodwin oscillator cassette. The Mann-Whitney test serves to check whether the similarity, or lack thereof, of median growth rates of transfected and non-transfected cells hold under statistical scrutiny. The following tests were carried out:

Test 1

H_0 : The difference in median between the cumulative growth of **BSF** wildtype cells and **uninduced** BSF cells transfected with pStable is **zero**



(a) Saw-tooth growth rate profile of BSF *T. brucei*



(b) Cumulative growth rate profile of BSF *T. brucei*. *x* is the time in hours.

Strain	Cell doubling time	Exponential fitted curve
Wildtype BSF	9.90 hrs	$y = 125471.e^{0.07t}$
Uninduced, transfected BSF	9.12 hrs	$y = 100071.e^{0.076t}$
Induced, transfected BSF	8.77 hrs	$y = 103008.e^{0.079t}$

(c) Cell doubling time of BSF *T. brucei*: Estimated cell doubling times and exponential fitted growth curve functions, where *t* is the time in hours, for BSF wildtype and pStable-transfected *T. brucei*.

Figure 7.2: Growth rate profiles and doubling times of induced/uninduced BSF *T. brucei* bearing pStable represented by a: a) saw-tooth curve profile, b) semi-logarithmic cumulative growth profile, c) doubling times and fitted exponential curves based on cumulative growth data. The axes show Density in cells/ml versus Time in hours. Error bars represent standard deviation.

Test 1: Comparing the cumulative growth of BSF wildtype cells with that of uninduced BSF <i>T. brucei</i> transfected with pStable		
Data	Wildtype (Cells/ml) 1×10^5 1.015×10^6 3.654×10^6 2.192×10^7 1.040×10^8 (Estimated)	Uninduced, pStable clone (Cells/ml) 1×10^5 6×10^5 3.99×10^6 2.254×10^7 1.475×10^8 (Estimated)
U statistic	11.5	
H₀ rejected	No	
Test 2: Comparing the cumulative growth of BSF wildtype cells with that of induced BSF <i>T. brucei</i> transfected with pStable		
Data	Wildtype (Cells/ml) 1×10^5 1.015×10^6 3.654×10^6 2.192×10^7 1.040×10^8 (Estimated)	Induced, pStable clone (Cells/ml) 1×10^5 6.65×10^5 3.39×10^6 1.78×10^7 2.026×10^8 (Estimated)
U statistic	11.5	
H₀ rejected	No	

Table 7.2: Comparison of growth rates of BSF *T. brucei* cells bearing the uninduced/induced stable Goodwin oscillator expression cassette and unmodified wildtype cells via the Mann-Whitney U test, which tests whether cumulative growth rates for transfected cells are significantly different from wildtype BSF cells. The table shows the experiment and estimated data points of the cell counts in Cells/ml which were used for the analysis to calculate the U statistic. This is compared to U_{crit} at the 0.05% confidence level which has value 2. The table also prints the outcome whether to reject the Null hypothesis, H_0 .

H_1 : The difference in median between the cumulative growth of **BSF** wildtype cells and **uninduced** BSF cells transfected with pStable is **not zero**

Test 2

H_0 : The difference in median between the cumulative growth of **BSF** wildtype cells and **induced** BSF cells transfected with pStable is **zero**

H_1 : The difference in median between the cumulative growth of **BSF** wildtype cells and **induced** BSF cells transfected with pStable is **not zero**

In order to implement the Mann-Whitney test, five data points per population were required. Thus, an additional data point was inferred using the exponential fitted curve to estimate cell growth at the 96th hour, as seen in Table 7.2 which also shows the test outcome. In both cases, the U statistic is larger than the critical U value. Hence, there is not enough evidence to reject the null hypothesis (H_0) that the difference between wildtype and pStable transfected cells is zero.

The doubling time of the wildtype BSF trypanosomes was estimated to be 9.90 hours in this assay and 7.53 hours in the previous growth assay for BSF *T. brucei* bearing the pUnstable oscillator (Section 6.1.1.1). In both studies, the same methods and wildtype strains were used. This is indicative of cell-cell variability (Snijder & Pelkmans, 2011).

7.1.1.2 Growth profile of PCF *T. brucei* cells bearing the stable Goodwin oscillator

A similar study to the above was carried out on PCF *T. brucei* to determine a growth profile for induced and uninduced pStable-transfected PCF clones. Results for the saw tooth and cumulative growth curve can be seen in Figures 7.3a and 7.3b, both of which indicate that uninduced transfected cells have a higher growth rate than the other tracks. This is also reflected in the estimated doubling times (Figure 7.3c). In order to analyse whether this divergence is statistically significant, the Mann-Whitney U test was applied to the PCF cumulative data set to analyse statistically whether the median growth rate of PCF wildtype cells differs significantly from the median growth of induced and uninduced transfected PCF cells, as per the following hypothesis:

Test 3

H_0 : The difference in median between the cumulative growth of wildtype **PCF** cells and **uninduced** PCF cells transfected with pStable is **zero**

H_1 : The difference in median between the cumulative growth of wildtype **PCF** cells and **uninduced** PCF cells transfected with pStable is **not zero**

Test 4

H_0 : The difference in median between the cumulative growth of **PCF** wildtype cells and **induced** PCF cells transfected with pStable is **zero**

H_1 : The difference in median between the cumulative growth of **PCF** wildtype cells and **induced** PCF cells transfected with pStable is **not zero**

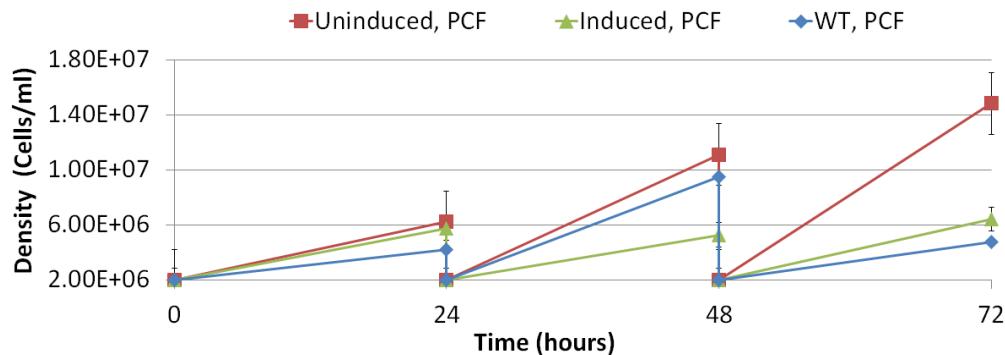
Experimental and inferred values used and test results for the U statistic can be found in Table 7.3. As with the above BSF results, the U statistic for both tests is larger than the critical U statistic, $U_{crit}=2$. Thus, both induced and uninduced transfected PCF cumulative growth profiles do not differ significantly from the wildtype PCF cumulative growth rate, indicating that the observed difference in PCF growth rates was due to outliers.

7.1.2 The minimum tetracycline induction concentration, tc_{thres} , required to induce pStable in *T. brucei* cells

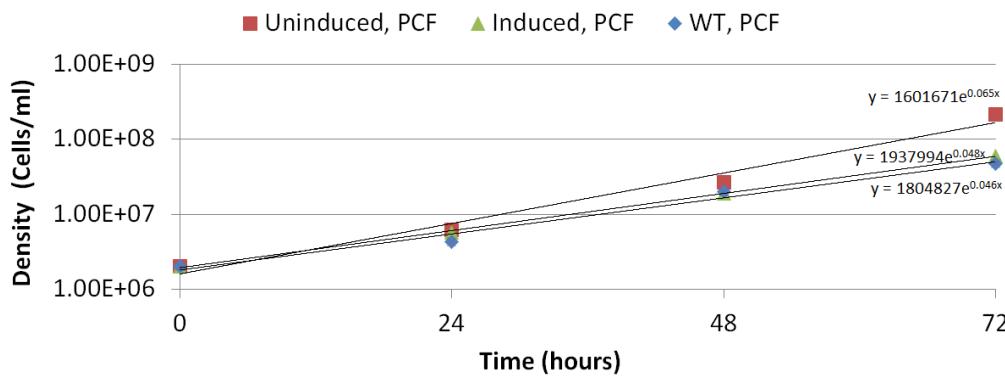
Following transfection, pStable-transfected BSF and PCF clones were shown to be inducible using $1\mu\text{g}/\text{ml}$ tetracycline (Figures 4.22a and 4.23a). In this assay, the minimum tetracycline concentration, tc_{thres} , which induces pStable was established for use during microscopy imaging of pStable (and pUnstable) clones, as discussed in further detail in Section 6.2.2 and 7.3.

Initially, four flasks of pStable-transfected BSF *T. brucei* were diluted to 1×10^5 cells/ml in HMI-9 media. Each volume was induced with $0-1\mu\text{g}/\text{ml}$ tetracycline and incubated for 24 hours, after which protein was isolated. Samples were analysed via a western blot detecting for the eGFP signal and a corresponding coomassie stain, as seen in Figure 7.4. eGFP signal is detected when the culture is induced with $2.5\text{ng}/\text{ml}$ tetracycline or higher. The antibody flagged up both a $\sim 27\text{kDa}$ band corresponding to the eGFP signal and an unknown $\sim 37\text{kDa}$ band. This band was not noted in other blots.

In order to confirm results, a second titration assay was carried out wherein the eGFP signal was analysed *in vivo* via fluorescent microscopy instead of western blots. This also served to investi-



(a) Saw-tooth growth rate profile of PCF *T. brucei*



(b) Cumulative growth rate profile of PCF *T. brucei*. *x* is the time in hours.

Strain	Cell doubling time	Exponential fitted curve
Wildtype PCF	15.00 hrs	$y = 1804827.e^{0.0462t}$
Uninduced, transfected PCF	10.66 hrs	$y = 1601671.e^{0.065t}$
Induced, transfected PCF	14.44 hrs	$y = 1937994.e^{0.048t}$

(c) Cell doubling time of PCF *T. brucei*: Estimated cell doubling times and exponential growth rates, where *t* is the time in hours, for wildtype and pStable-transfected PCF *T. brucei*.

Figure 7.3: Growth rate profiles and doubling times of induced/uninduced PCF *T. brucei* bearing pStable represented by a: a) saw-tooth curve profile, b) semi-logarithmic cumulative growth profile, c) doubling times and fitted exponential curves based on cumulative growth data. Induced cells were treated with 1 μ g/ml tetracycline. The plot axes show Density in cells/ml versus Time in hours. Error bars represent standard deviation. The experiment was carried out in duplicates.

Test 3 Comparing the cumulative growth of PCF wildtype cells with that of uninduced PCF transfected with pStable		
Data	Wildtype (Cells/ml) 2×10^6 4.25×10^6 2.02×10^7 4.79×10^7 1.49×10^8 (Estimated)	Uninduced, transfected (Cells/ml) 2×10^6 6.23×10^6 3.46×10^7 2.57×10^8 8.21×10^8 (Estimated)
U statistic	9.5	
H₀ rejected	No	
Test 4 Comparing the cumulative growth of PCF wildtype cells with that of induced PCF transfected with pStable		
Data	Wildtype (Cells/ml) 2×10^6 4.25×10^6 2.02×10^7 4.79×10^7 1.49×10^8 (Estimated)	Induced, transfected (Cells/ml) 2×10^6 5.75×10^6 1.52×10^7 4.91×10^7 1.94×10^8 (Estimated)
U statistic	11.5	
H₀ rejected	No	

Table 7.3: Comparison of growth rates of PCF *T. brucei* cells bearing the uninduced/induced stable Goodwin oscillator expression cassette and unmodified wildtype cells via the Mann-Whitney U test, which tests whether cumulative growth rates for transfected cells are significantly different from wildtype PCF cells. The table shows the experiment and estimated data points of the cell counts in Cells/ml which were used for the analysis to calculate the U statistic. This is compared to U_{crit} at the 0.05% confidence level which has value 2. The table also prints the outcome whether to reject the Null hypothesis, H_0 .

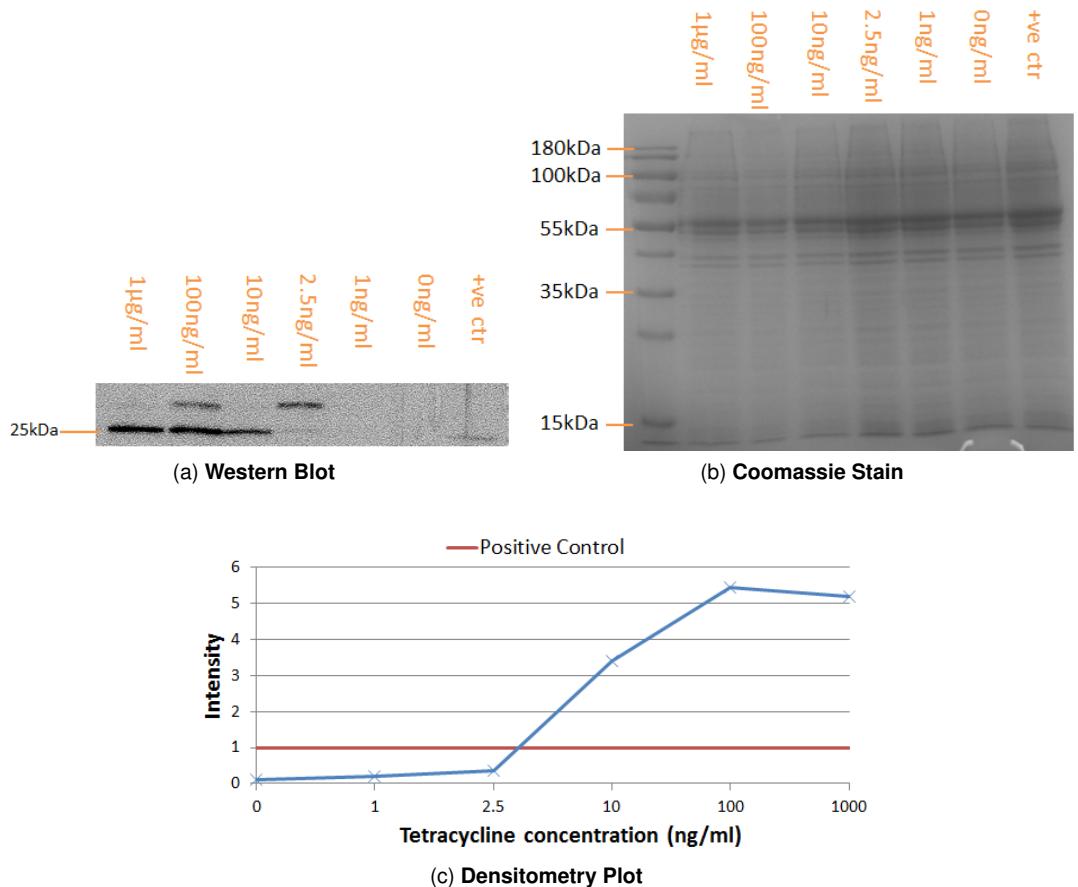


Figure 7.4: Titration assay carried out to deduce the tetracycline induction threshold, tc_{thres} , of the pStable network. Analysis was carried out via a) western blot detecting for the eGFP signal and b) corresponding coomassie stain of a pStable-transfected BSF clone induced with different levels of tetracycline in the $0-1\mu\text{g}/\text{ml}$ range. c) shows a densitometry plot quantifying the amount of eGFP detected in the western blot using the positive control as a reference point of intensity 1.0. The positive control shows protein from the constitutively expressed pConstitutive network.

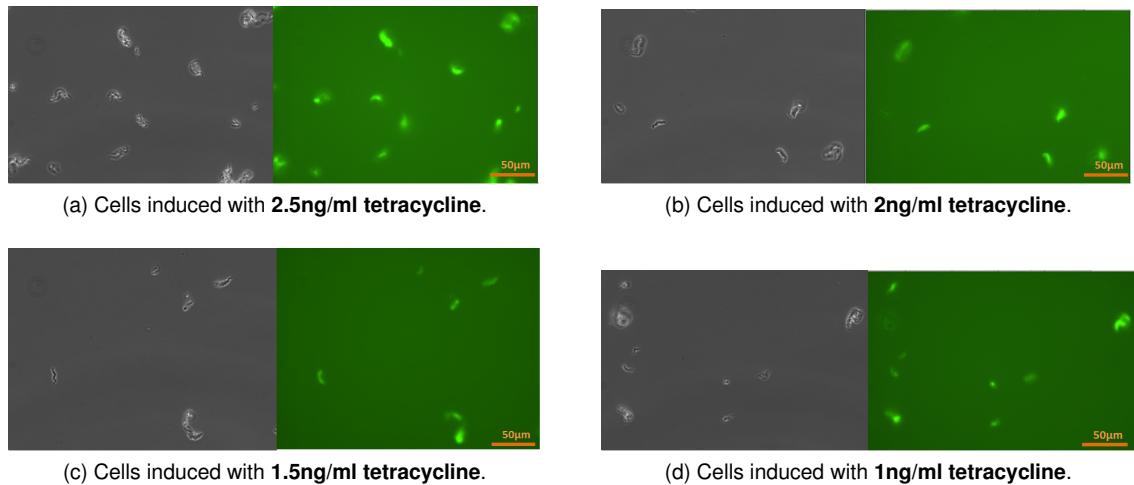


Figure 7.5: Second titration assay carried out to deduce the tetracycline induction threshold, tc_{thres} , for pStable. Analysis is carried out via live cell imaging of a pStable-transfected BSF clone induced with a) 2.5ng/ml b) 2ng/ml c) 1.5ng/ml d) 1ng/ml of tetracycline. Each image shows brightfield (left) and fluorescent (right) images of the same plane using a 40X objective. Scale bar = $50\mu\text{m}$.

gate whether the 37kDa band observed in the western blot was due to contamination. Similarly to the previous titration assay, cells were induced with 1-2.5ng/ml of tetracycline and incubated for 24 hours. Following this, cells were embedded in 40% Cygel as per Section 5.7. Fluorescent microscopy was carried out using a 40X objective. In parallel, brightfield microscopy using a 40X objective was used as a control measure to verify that observed fluorescence is being emitted from cells rather than debris. Microscopy results in Figure 7.5 show that whereas induction using 1ng/ml tetracycline gave no detectable signal when the eGFP protein was selected for on a western blot (Figure 7.4), microscopy images showed cells induced with 1ng/ml to be fluorescing. This is due to microscopy fluorescence detection being more sensitive than western blotting. Thus, further microscopy-based analysis using lower tetracycline concentrations was carried out in order to deduce the tetracycline induction threshold, tc_{thres} .

For the third titration assay, a similar process as above was implemented, inducing pStable-transfected BSF cells with 1fg/ml-750pg/ml tetracycline. Results of brightfield and fluorescent imaging (not presented) show that cell populations induced with over 10fg/ml tetracycline fluoresce, indicating an excess of tetracycline molecules relative to TetR protein numbers. However, in populations induced with 10fg/ml of tetracycline only a margin of the cells fluoresce, with other cells remaining uninduced.

This is attributed to intracellular stochasticity and in particular, the dispersion of tetracycline among the population. Specifically, in some cells there are more tetracycline molecules than TetR proteins, resulting in (temporary) observed fluorescence due to (temporary) constitutive expression. This occurs because TetR proteins bind to the excess tetracycline molecules instead of the tet operator, since they have a greater affinity for the former. This allows for protein expression. This continues until the number of TetR proteins exceeds the intracellular tetracycline molecules, at which stage the TetR protein binds to the tet operator and TetR inhibition takes place, as explained previously in Figure 6.7. In the remainder of the cells there is an excess of TetR proteins rather than tetracycline molecules, resulting in no modified expression patterns and no fluorescence observed. In both cases, oscillatory protein production in the pStable network can still take

place. Following these findings, the threshold induction $t_{c_{thres}}$ is taken to be 10fg/ml.

7.1.3 Functional activity of the TetR regulator - Expression dynamics of the stable Goodwin oscillator expression cassette in *T. brucei* cells following tetracycline addition and removal

The next step was to establish the length of time over which tetracycline remains effective as an inducer for pStable. This in turn serves to inform whether protein expression inhibition in the GRN is regulated by TetR or otherwise, since protein expression should be suspended as the tetracycline molecules degrade and TetR binds to the operator instead. pStable-transfected BSF clone were induced with 1 μ g/ml tetracycline, incubated for 24 hours and harvested in order to remove extracellular tetracycline molecules. Following this, protein samples were prepared at 0hr, 1hr, 3hr, 6hr, 24hr, 48hr and 72hr time-points.

Samples were analysed via a western blot detecting for the eGFP signal and corresponding coomassie stain, as seen in Figure 7.6. Following removal of tetracycline, the eGFP signal retains strength as a result of intracellular tetracycline for over 6 hours. After a 24 hour period, only residual levels of the eGFP signal remain. After 72 hours, no eGFP was detected. The coomassie stain confirms that protein samples were prepared correctly. Therefore, this assay has shown that the effect of tetracycline on pStable can be reversed after a 24-hour period. Moreover, expression patterns are in line with TetR inhibitory regulation, serving to confirm the functional activity of the TetR regulator.

7.1.4 Half-life of TetR and the destabilised eGFP protein in *T. brucei* cells having the stable Goodwin oscillator

For the final characterisation on pStable components, the half-life of TetR and eGFP proteins was deduced in order to plug the data into mathematical modelling to develop a hypothetical profile for the plasmid phenotype via simulations. In addition, knowledge of the eGFP half-life will help inform whether the protein is sufficiently destabilised as to exhibit oscillations due to TetR inhibition or whether it is too stable to reflect regulatory influence within the network. This is similar in scope of the cycloheximide-based assay carried out on pUnstable. The change in protein quantity over a six hour period following cycloheximide addition was recorded via a western blot detecting for the appropriate protein and a corresponding coomassie stain.

7.1.4.1 Half-life of the TetR protein in *T. brucei* cells bearing the stable Goodwin oscillator

For the detection of the TetR protein, both antibodies used (Anti-Tet Repressor Antibody (EMD Millipore Co., MA, U.S.A) and Anti-Tet Repressor Antibody (Sigma-Aldrich, Munich, Germany)) gave inconclusive results (not shown) when samples were run on a western blot. No bands were detected in spite of multiple runs of the experiment. The coomassie stain showed strong uniform bands throughout, indicating that protein was prepared correctly. For mathematical modelling, literature sources were referred to for an estimate of the TetR half-life, as will be discussed in

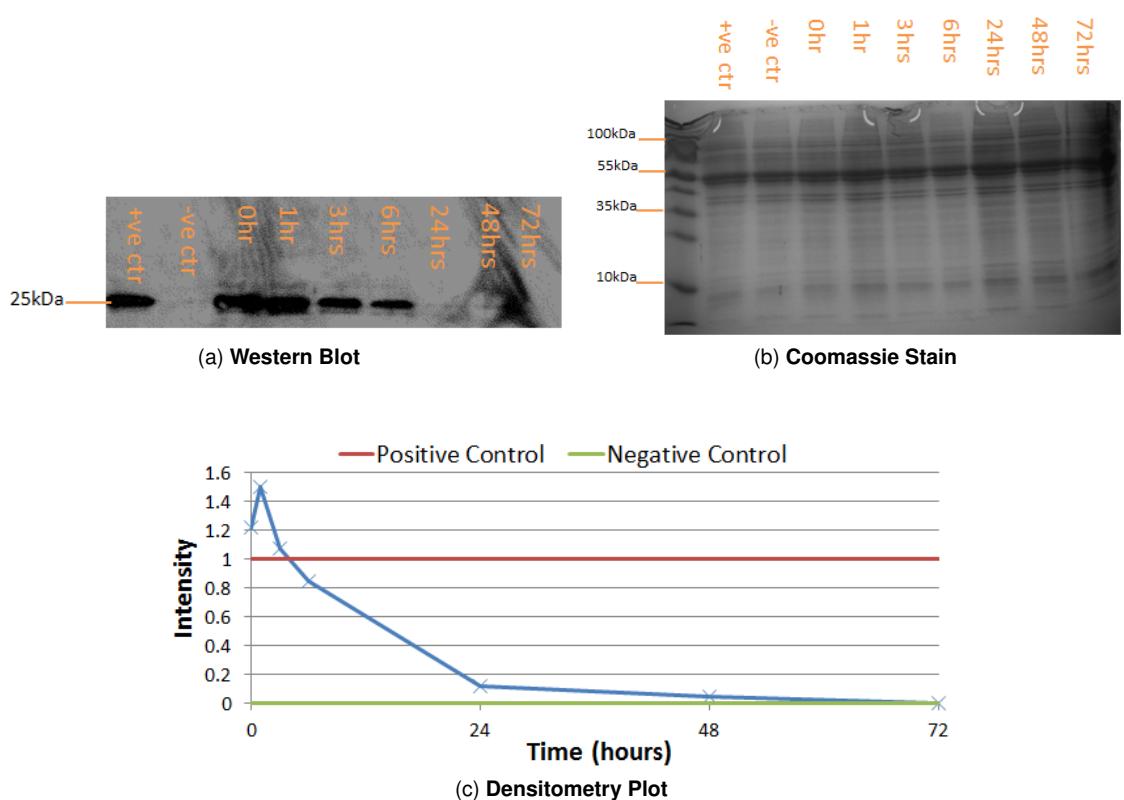


Figure 7.6: eGFP expression levels over a 72-hour period of BSF *T. brucei* bearing pStable following 1 μ g/ml tetracycline induction and removal. Analysis is carried out via a a) western blot detecting for the eGFP signal, b) corresponding coomassie stain and c) densitometry plot quantifying the amount of eGFP protein detected in the western blot using the positive control as a reference point of intensity 1.0. The positive control is a pConstitutive protein sample, while the negative control is an uninduced pStable protein sample.

7.1.4.2 Half-life of the destabilised eGFP protein in *T. brucei* cells bearing the stable Goodwin oscillator

In order to assess whether Ub-L was exerting a destabilising influence over the eGFP protein, an assay was carried out to deduce the half-life of the reporter protein being expressed in pStable, as per previous studies. Then, by comparing results to the half-life of a control plasmid which has the *non-destabilised* eGFP protein, a deduction could be made as to whether Ub-L was being effective in pStable. As a control plasmid having the (non-destabilised) eGFP gene, a BSF 2T1 cell line (VSG221 expressing Tagged cell line, clone 1) was used (Alsfeld et al., 2005). This comparative analysis was only carried out with the pStable network, rather than both Goodwin oscillators, because basal transcription within pUnstable makes for a less straightforward distinction.

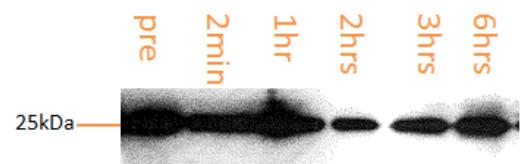
Figure 7.7 shows western blot results detecting for the eGFP protein in pStable protein samples. The protein half-life of eGFP is deduced to be between three and six hours. The corresponding Coomassie stain confirms that the reduction in signal strength is a result of decrease in protein quantity, rather than incorrect sample preparation. Additionally, eGFP signals of untreated protein samples (Figures 7.7b and 7.7d) do not show a reduction in signal strength over the six-hour timespan, indicating that the weakening signal observed in treated samples is a result of cycloheximide having an effect on the cell culture, rather than a process which occurs independently. eGFP band sizes, at c. 27kDa, indicate that the Ub-L-eGFP moiety is being cleaved *in vivo* as expected, as otherwise the signal would have a size of c. 38kDa (Dantuma et al., 2000).

The next step was to analyse whether the half-life of the reporter protein in pStable was different from that of non-destabilised eGFP expressed in the control 2T1 *T. brucei*. Results for the assay carried out on the 2T1 strain are found in Figure 7.8. In clear contrast to the eGFP signal within the cycloheximide-treated protein sample isolated from pStable-transfected BSF *T. brucei*, the non-destabilised eGFP signal in cycloheximide-treated 2T1 cells retains a strong band throughout the six hour observation period (Figure 7.8a). The corresponding coomassie stain (Figure 7.8c) shows that protein was prepared correctly. The blank assay (Figure 7.8b and 7.8d), wherein no cycloheximide was added to the induced 2T1 culture shows no clear difference in signal strength. Therefore, it is unclear whether observations are a result of the cycloheximide not having an effect in this assay run or whether they are a result of the eGFP being stable for over 24 hours, as is expected (Corish & Tyler-Smith, 1999).

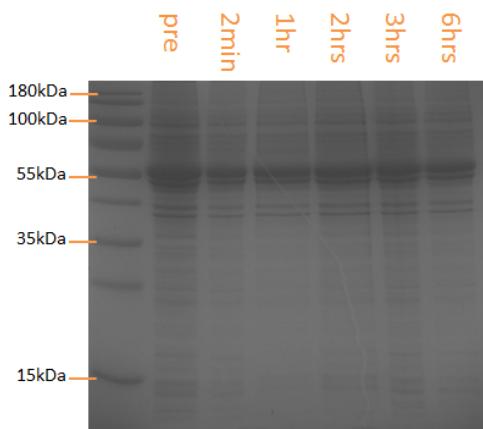
In order to control for this, a final assay was carried out wherein pStable-transfected and 2T1 BSF *T. brucei* were treated with cycloheximide *simultaneously* rather than in different runs, after which protein was isolated at one hour, three hours and six hours post treatment and run on a western blot to detect for the eGFP protein. Results in Figure 7.9 show a clear difference over time between the eGFP signal strength of the cycloheximide-treated 2T1 cell line and the eGFP signal strength of cycloheximide-treated pStable-transfected BSF clone. This indicates that the Ub-L gene is exerting an influence in the Ub-L-eGFP moiety, reducing the half-life of eGFP from its stable 24 hour half-life to 3-6 hours. This is in line with observations made when Ub-L was fused to β -gal in reticulocytes red blood cells, wherein β -gal half-life was reduced to 5.5 hours, as per the N-end rule (Gonda et al., 1989). The next step was to plug this data into a modelling



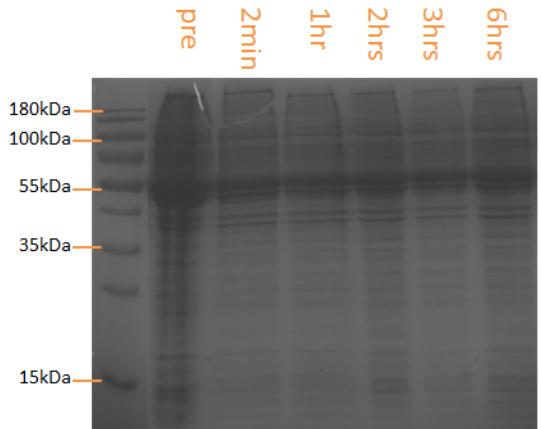
(a) Western blot detecting for eGFP in protein samples from induced pStable-transfected BSF *T. brucei* **treated** with cycloheximide



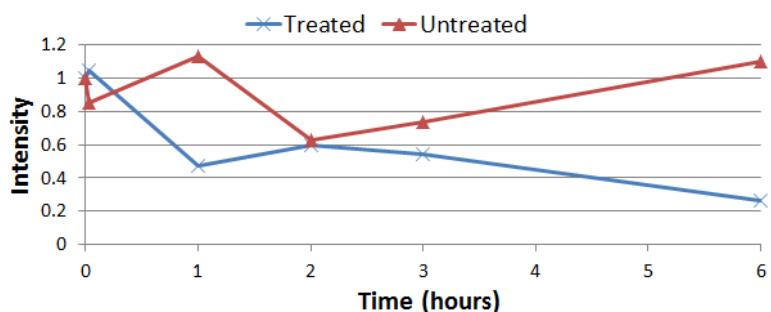
(b) **Control:** Western blot detecting for eGFP in protein samples from induced pStable-transfected BSF *T. brucei* **not treated** with cycloheximide



(c) Coomassie stain for protein samples from induced pStable-transfected BSF *T. brucei* **treated** with cycloheximide

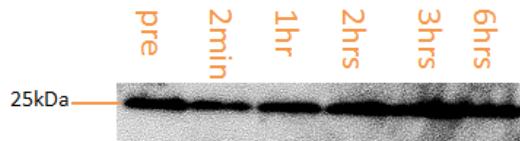


(d) Control: Coomassie stain for protein samples from induced pStable-transfected BSF *T. brucei* **not treated** with cycloheximide



(e) Densitometry plot showing the signal intensity extracted from the western blots above of eGFP protein from pStable-transfected cells either treated or untreated with cycloheximide.

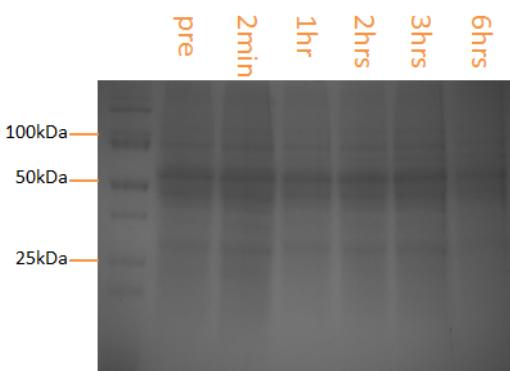
Figure 7.7: Deducing the half-life of the **destabilised eGFP protein** in BSF *T. brucei* bearing pStable and induced with 1 μ g/ml tetracycline. Protein samples were prepared at specific time-points following cycloheximide addition and then analysed via western blot (a) detecting for eGFP and a corresponding coomassie stain (c). In parallel, a control assay (b, d) was carried out wherein no cycloheximide was added to the culture. e) shows a densitometry plot quantifying the amount of eGFP in both treated and untreated pStable-transfected cells detected in the western blots. The plot uses the eGFP intensity from the measurement taken prior to the addition of the cycloheximide or blank ('pre') as value 1.0.



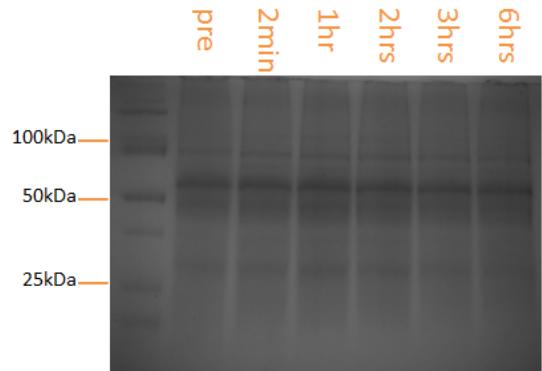
(a) Western blot detecting for eGFP in protein isolated from 2T1 BSF *T. brucei* **treated** with cycloheximide



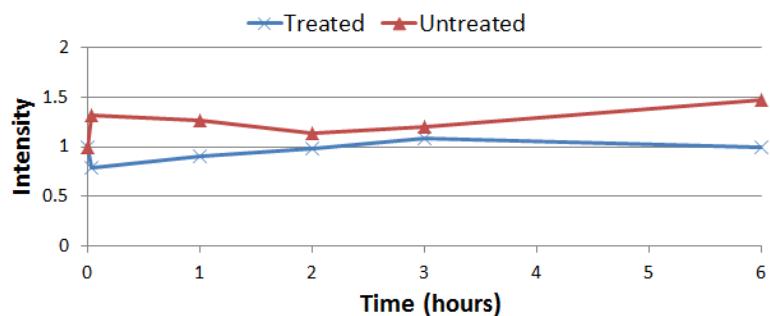
(b) **Control:** Western blot detecting for eGFP in protein isolated from 2T1 BSF *T. brucei* **not treated** with cycloheximide



(c) Coomassie stain for protein isolated from 2T1 BSF *T. brucei* **treated** with cycloheximide



(d) **Control:** Coomassie stain for protein isolated from 2T1 BSF *T. brucei* **not treated** with cycloheximide



(e) Densitometry plot showing the signal intensity extracted from the western blots above of eGFP protein isolated from 2T1 BSF cells either treated or untreated with cycloheximide.

Figure 7.8: Deducing the half-life of the **non-destabilised eGFP protein** in 2T1 BSF *T. brucei*. Protein samples were prepared at specific time-points following cycloheximide addition and then analysed via western blot (a) detecting for eGFP and a corresponding coomassie stain (c). In parallel, a control assay (b, d) was carried out wherein no cycloheximide was added to the culture. e) shows a densitometry plot quantifying the amount of eGFP in both treated and untreated 2T1 BSF clones detected in the western blots. The plot uses the eGFP intensity from the measurement taken prior to the addition of the cycloheximide or blank ('pre') as value 1.0.

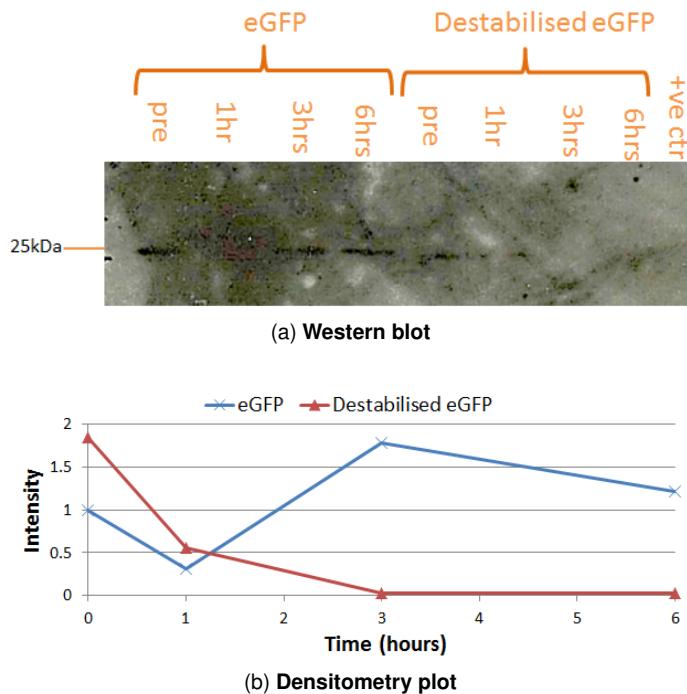


Figure 7.9: Comparing the half-life of non-destabilised and destabilised eGFP proteins in induced control 2T1 and pStable-transfected BSF clones respectively, via: a) protein samples which were prepared at specific time-points following cycloheximide addition and then analysed via western blot, b) densitometry plot quantifying the detected eGFP in the western blot. The plot uses the reading at time 0 of the non-destabilised eGFP as a reference point of value 1.0. The positive control is a sample from the constitutively expressed pConstitutive plasmid.

frame and assess whether the protein is sufficiently destabilised as to reflect expression patterns, as will be seen next.

7.1.5 Evaluation of the stable oscillator phenotype following preliminary analysis

Following characterisation and parameterisation of the pStable oscillator components and its phenotype within the *T. brucei* host, the following findings and observations were made:

1. Preliminary characterisation of the pStable oscillator via western blots and coomassie stains has shown that the main components of the network are functional and that TetR and eGFP protein is successfully synthesised. Promoters successfully initiate transcription: detected eGFP levels indicate that the P_{RRNA} promoter is functional while transfected clones cultured in the presence of Hygromycin indicate that the P_{EP1} promoter is also working.
2. No eGFP signal was detected unless the pStable-transfected *T. brucei* culture was induced with tetracycline. Moreover, once induced with $1\mu\text{g/ml}$ tetracycline, a strong eGFP signal was detected for over 6 hours, after which there was a decrease in signal strength and a resuming of expression inhibition. This indicates that TetR regulation is being implemented and that the tetracycline operator is functional. This is similar to findings for the pUnstable Goodwin oscillator.
3. The tetracycline induction threshold, tc_{thres} , for pStable-transfected BSF *T. brucei* at a con-

centration of 1×10^6 cells/ml is over 10fg/ml. In cultures induced with 10fg/ml of tetracycline only a margin of the cells fluoresce. This phenomenon allows for the tuning of expression dynamics and was used during fluorescent microscopy.

4. There is no statistically significant difference between growth rates of wildtype cells and pStable-transfected BSF and PCF *T. brucei*. This is similar to findings for *T. brucei* bearing the pUnstable Goodwin oscillator.
5. The half-life of the destabilised eGFP protein is 3-6 hours, a much lower value than the standard 24 hour half-life. This implies that the Ub-L gene fused to eGFP is flagging up the protein for degradation. This information will be plugged into mathematical modelling to simulate a profile for protein expression dynamics.
6. In conclusion, the components of the pStable GRN allow for oscillations to take place. Since the cell population is unlikely to oscillate in synchrony (Elowitz & Leibler, 2000; Stricker et al., 2008; Danino et al., 2010; Tigges et al., 2010) and since western blots have limited sensitivity, oscillations are unlikely to be detected via protein sample analysis. Therefore fluorescent microscopy imaging was carried out in order to observe fluorescent patterns.

7.2 Preliminary quantitative mathematical modelling and analysis of protein expression dynamics within the trypanosomal Goodwin oscillator

Preliminary quantitative mathematical modelling was carried out to develop a mathematical model of the trypanosomal Goodwin oscillator and run simulations to inform whether the network is likely to function as an oscillator within its *T. brucei* host. This is the first mathematical model of a synthetic GRN which has been developed for use within a trypanosomal platform. Previous mathematical analysis on *T. brucei* has focused on the epidemiological aspect, with studies modelling existing pathways within *T. brucei* (Kerkhoven et al., 2013; Verlinde et al., 2001; Gu et al., 2013), disease patterns (Medlock et al., 2013) and the modelling of *T. brucei* within a systems biology context (Bakker et al., 2010).

The following analysis uses half-life data based on preliminary analysis of the pStable Goodwin oscillator but does not exclude results from applying to the pUnstable network, as will be explained further on.

7.2.1 Screening of the Goodwin oscillator's parameter space using Gillespie stochastic simulations

The Goodwin oscillator was represented mathematically in terms of molecular components/species and related biochemical reactions (Monod, 1971). The main species of interest are TetR and eGFP protein. The model (i.e. set of species and reactions) was used to map out how TetR and eGFP protein quantities vary over time due to the interplay of biochemical reactions, such as transcription and degradation, which occur to and between the network components, such as the tet operator and TetR protein. The set of components which were modelled was kept fixed throughout the modelling process. On the other hand, reaction rates were treated as variable

parameters (Stricker et al., 2008; Fung et al., 2005; Danino et al., 2010).

A screening assay was developed in order to deduce the parameter values which lead to periodic TetR and eGFP protein expression dynamics. This is also known as parameter analysis. By assuming that the selected set of species and biochemical reactions are representative of the species and reactions which lead to oscillations, reaction rate parameters were varied across a wide range of values to vet which parameter combinations, if any, led to oscillations. Thus, the investigation consisted of multiple iterations wherein a combination of parameters was randomly selected and input into Gillespie simulations to generate a time series of the species' quantities via the ABC Sysbio software. This is a Python based software which implements Bayesian parameter inference, model selection and stochastic and deterministic simulations within dynamical systems including GRNs (Liepe et al., 2014; Filippi et al., 2013; Barnes et al., 2011b, 2011a).

In each iteration of the screening assay, each reaction rate value was selected from a uniformly distributed range of values based on information extracted from experiments or literature sources. *T. brucei*-specific data was not available in the literature available. Thus, generic uniform distributions with a wide range of values were set for each parameter, as explained further below. Once parameters were selected, the Direct Method Gillespie algorithm (Gillespie, 1977) was implemented. The Gillespie algorithm is a stochastic simulation algorithm (SSA) which uses probability measures to map out changes in species molecule numbers due to the biochemical reactions which occur over time. The SSA is simple in principle and has been used in various synthetic biology settings (Elowitz & Leibler, 2000; Stricker et al., 2008). Although there are several variations of the Gillespie algorithm which have been developed, the original SSA could be used to address the scope of this study. Following iterations, the parameter combinations, if any, which led to an oscillatory output for the TetR and eGFP proteins were selected.

By using the Gillespie SSA which bases the decision of *when* and *which* biochemical reaction next takes place on a probability measure, the screening assay introduced an element of randomness, reflecting the stochastic nature of the biological setting. This was considered to be more representative of the GRN than deterministic modelling, which can be considered as an average representation of stochastic trajectories (Kepler & Elston, 2001) and which has been shown to differ from stochastic simulations outside the thermodynamic limit (Stricker et al., 2008; Elowitz & Leibler, 2000), which conditions are not satisfied in this case (Gillespie, 1976).

7.2.1.1 Development of the mathematical model representing protein expression dynamics within the trypanosomal Goodwin oscillator

The first step was to translate the pStable Goodwin network into mathematical modules, specifically, into a set of species and reactions. The model is based on the theory that the quantity of each species within the GRN is in constant flux due to the multiple ongoing biochemical reactions and processes within the network (Monod, 1971). Although the aim is to map out protein trajectories for TetR and eGFP, several other components involved in protein synthesis have to be considered. The species which were modelled are listed in Table 7.4. While the selected components are not an exhaustive list of the components within pStable and pUnstable, they are representative of the species involved in the generation of TetR and eGFP oscillations. Note that, for example, the Myc gene in pUnstable was not modelled explicitly since it does not impact TetR and eGFP dynamics but only serves characterisation purposes.

Species	Description
1 mRNA _{TetR}	The product of <i>tetR</i> gene transcription
2 TetR	The product of mRNA _{TetR} translation
3 TetR ₂	TetR dimer
4 Op	The tet operator
5 TetR ₂ .Op	The complex formed when TetR ₂ is bound to Op
6 mRNA _{eGFP}	The product of eGFP gene transcription
7 eGFP	The product of mRNA _{eGFP} translation

Table 7.4: The species/components from the plasmid encoding for the Goodwin oscillator which were selected for preliminary quantitative modelling and trajectory simulations.

In order to represent interactions between the TetR-associated species (Species 1-5 in Table 7.4), biochemical reactions R_1-R_{10} listed in Table 7.5 were selected for modelling. The reaction rate parameters are as described per Table 7.6 which also lists the parameters' range of uniformly distributed values. In the model, the low binding rate of TetR to the operator in pUnstable is not modelled explicitly and is instead modelled implicitly in the low binding rates.

The rates are primarily based on values extracted from Tigges et al. (2010), wherein the authors developed a multi-component system based on positive and inhibitory regulation within a eukaryotic chassis. The rate of transcription is based on data from Stricker et al. (2008). Although detailed data of TetR systems within prokaryotic chassis is more readily available (Biliouris et al., 2011), it is likely to be less relevant than eukaryote-based data.

As with the set of species, the list of reactions is a representative but not exhaustive list of biochemical processes within the GRN. Figures 7.10a and 7.10b represent how the selected TetR species interact with each other as per reactions R_1-R_{10} listed in Table 7.5.

The same process as for the TetR family of components was carried out for eGFP-related species (Op, mRNA_{eGFP} and eGFP). The set of reactions $R_{11}-R_{14}$ modelled for the eGFP components are listed in Table 7.5, while the corresponding reaction rates are as per Table 7.6. The degradation rate of eGFP was determined experimentally and is set at 0.004min^{-1} . (Section 7.1.4). Although this data is based on pStable analysis, results are applicable to the pUnstable network as well. Figure 7.10c represents how the selected eGFP species interact with each other as per reactions $R_{11}-R_{14}$.

7.2.1.2 Differential equation representation of the mathematical model

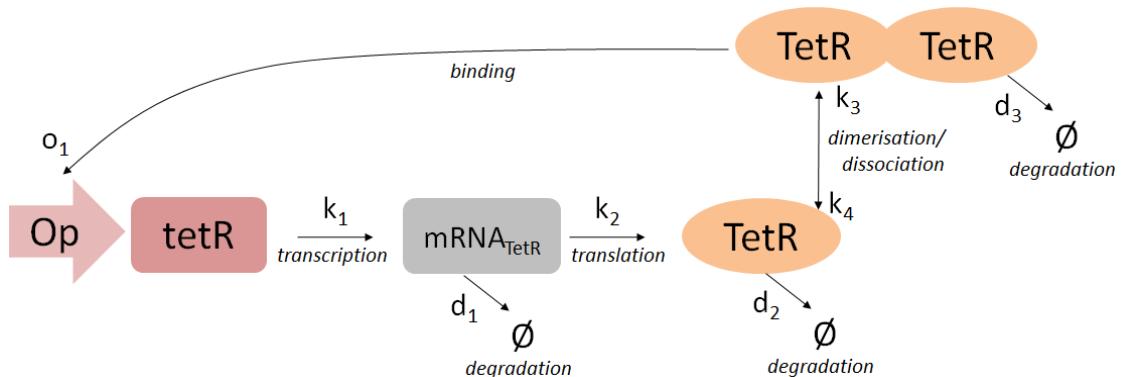
Prior to Gillespie simulations, the set of species and reactions were assembled to develop a set of ordinary differential equations (ODEs) which represent how each species evolves over time due to the interplay of the biochemical reactions which are being modelled. This set of ODEs was then coded in Systems Biology Markup Language (SBML) as per requirements of the ABC-Sysbio program running Gillespie simulations. The set of ODEs is a deterministic representation of the

		Reaction	Description
R_1	\emptyset	$\xrightarrow{k_1} mRNA_{tetR}$	Transcription
R_2	$mRNA_{tetR}$	$\xrightarrow{k_2} mRNA_{tetR} + TetR$	Translation
R_3	$TetR + TetR$	$\xrightarrow{k_3} TetR_2$	Dimerisation
R_4	$TetR_2$	$\xrightarrow{k_4} TetR + TetR$	Dissocation
R_5	$Op + TetR_2$	$\xrightarrow{o_1} Op.TetR_2$	Binding
R_6	$Op.TetR_2$	$\xrightarrow{o_2} Op + TetR_2$	Unbinding
R_7	$mRNA_{tetR}$	$\xrightarrow{d_1} \emptyset$	Degradation
R_8	$TetR$	$\xrightarrow{d_2} \emptyset$	Degradation
R_9	$TetR_2$	$\xrightarrow{d_3} \emptyset$	Degradation
R_{10}	$Op.TetR$	$\xrightarrow{d_4} Op$	Degradation
R_{11}	\emptyset	$\xrightarrow{k_1} mRNA_{eGFP}$	Transcription
R_{12}	$mRNA_{eGFP}$	$\xrightarrow{k_2} mRNA_{eGFP} + eGFP$	Translation
R_{13}	$mRNA_{eGFP}$	$\xrightarrow{d_1} \emptyset$	Degradation
R_{14}	$eGFP$	$\xrightarrow{d_5} \emptyset$	Degradation

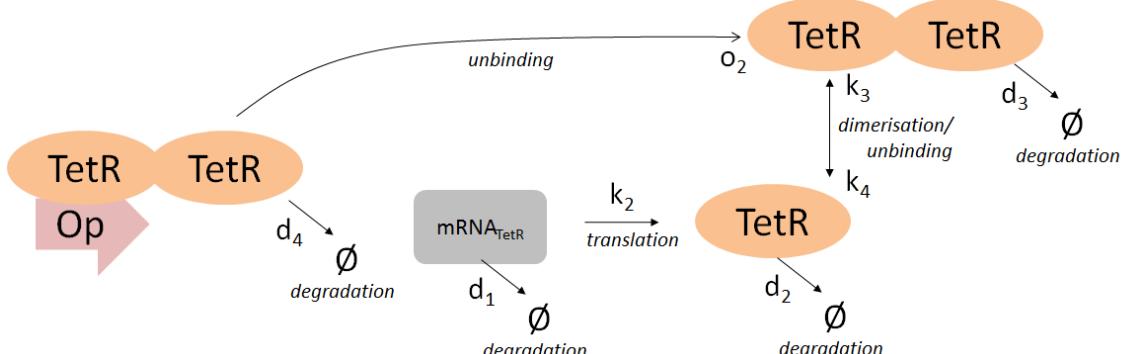
Table 7.5: The biochemical reactions selected for quantitative modelling of network dynamics. The species are as described in Table 7.4, while the parameter reaction rates k_i , o_i and d_i are as described in Table 7.6.

Rate	Description	Distribution
k_1	Rate of transcription of <i>tetR</i> and eGFP genes	$U(0.18mol.min^{-1}, 0.54mol.min^{-1})$
k_2	Rate of translation of $mRNA_{tetR}$ & $mRNA_{eGFP}$	$U(0.01min^{-1}, 0.03min^{-1})$
k_3	Rate of dimerisation of TetR	$U(0.007mol^{-1}.min^{-1}, 0.02mol^{-1}.min^{-1})$
k_4	Rate of dissociation of $TetR_2$	$U(0.0005min^{-1}, 0.0015min^{-1})$
o_1	Binding rate of $TetR_2$ to Op	$U(0.009mol^{-1}.min^{-1}, 0.027mol^{-1}.min^{-1})$
o_2	Unbinding rate of $TetR_2$ from Op	$U(0.00005min^{-1}, 0.00015min^{-1})$
d_1	Degradation rate of $mRNA_{tetR}$ and $mRNA_{eGFP}$	$U(0.008min^{-1}, 0.026min^{-1})$
d_2	Degradation rate of TetR	$U(0.011min^{-1}, 0.035min^{-1})$
d_3	Degradation rate of $TetR_2$	$U(0.011min^{-1}, 0.035min^{-1})$
d_4	Degradation rate of $TetR_2$ when bound to Op	$U(0.011min^{-1}, 0.035min^{-1})$
d_5	Degradation rate of eGFP	$0.004min^{-1}$

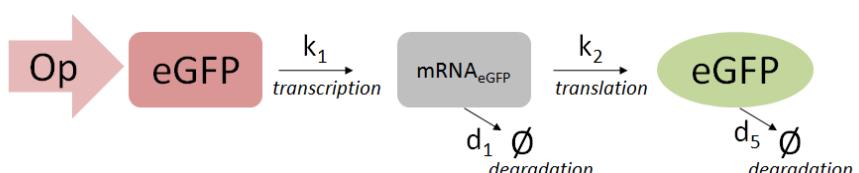
Table 7.6: The biochemical reaction rate parameters and their distributions, used in the modelling of the TetR and eGFP family of components within the trypanosomal Goodwin oscillator mathematical model. The rates are used as described in Table 7.5 and are based on values extracted from Tigges et al. (2010). ‘mol’ stands for molecules.



(a) Diagram showing the components and biochemical rates in the math model which correspond to **interactions between the TetR components when TetR₂ is not bound to Op.**



(b) Diagram showing the components and biochemical rates in the math model which correspond to **interactions between the TetR components when TetR₂ is bound to Op.** No transcription takes place during this stage.



(c) Diagram showing the components and biochemical rates in the math model which correspond to **interactions between the eGFP components when TetR₂ is not bound to Op.** If TetR₂ is bound to Op, only eGFP and mRNA_{eGFP} degradation reactions take place.

Figure 7.10: Diagram showing set of mathematically modelled components and biochemical reactions corresponding to: a) interactions between TetR₂ components when TetR₂ is not bound to Op, b) interactions between TetR₂ components when TetR₂ is bound to Op and c) interactions between eGFP components when TetR₂ is not bound to Op. The DNA species within the Goodwin oscillator which are modelled are mRNA_{TetR}, TetR, TetR₂, Op, TetR₂.Op, mRNA_{eGFP} and eGFP. The *tetR* and eGFP genes are represented in the diagram but are not modelled explicitly in the model. k_1 represents the rate of transcription, k_2 is the rate of translation, k_3 is the dimerisation rate, k_4 is the rate of dissociation, o_1 represents the binding rate of TetR₂ to Op, o_2 represents the reverse reaction rate and d_1-d_5 represent degradation rates.

model. Stochasticity is then introduced via simulations, effectively transforming the equations into the more realistic chemical master equation (Gillespie, 1976). Species are represented as $[x]$:

$$\frac{d[mRNA_{\text{tetR}}]}{dt} = k_1 \frac{[Op]}{S + [Op]} - d_1[mRNA_{\text{tetR}}] \quad (7.1)$$

$$\frac{d[TetR]}{dt} = k_2[mRNA_{\text{tetR}}] - 2k_3[TetR] \frac{[TetR - 1]}{2} + 2k_4[TetR_2] - d_2[TetR] \quad (7.2)$$

$$\frac{d[TetR_2]}{dt} = k_3[TetR] \frac{[TetR - 1]}{2} - k_4[TetR_2] - d_3[TetR_2] - o_1[TetR_2][Op] + o_2[TetR_2.Op] \quad (7.3)$$

$$\frac{d[Op]}{dt} = -o_1[TetR_2][Op] + o_2[TetR_2.Op] + d_4[TetR_2.Op] \quad (7.4)$$

$$\frac{d[TetR_2.Op]}{dt} = o_1[TetR_2][Op] - o_2[TetR_2.Op] - d_4[TetR_2.Op] \quad (7.5)$$

$$\frac{d[mRNA_{\text{eGFP}}]}{dt} = k_1 \frac{[Op]}{S + [Op]} - d_1[mRNA_{\text{eGFP}}] \quad (7.6)$$

$$\frac{d[eGFP]}{dt} = k_2[mRNA_{\text{eGFP}}] - d_5[eGFP] \quad (7.7)$$

It can be seen that all reactions except for dimerisation and transcription (seen in (7.1), (7.2), (7.3) and (7.6)) are modelled as mass action laws. Dimerisation takes into account the possible combinations of pairs of TetR which are used for the reaction, as seen in Gillespie (1976). Transcription is modelled as a Hill function, as seen in Elowitz and Leibler (2000), Fung et al. (2005) and Tigges et al. (2009) to name but a few. It has general form $f = \frac{[X]^n}{(S)^n + [X]^n}$ wherein X is the species of interest (in this case, Op) and n is the cooperativity coefficient of the species which in this case has value 1 (Sheinman & Kafri, 2012). This is based on the assumption that there is only one operator per cell due to the efficiency, or lack thereof, of trypanosome transfection. As discussed in Chapter 4, 2×10^7 cells were used per transfection. Of these, an average of 10 were successfully transfected and picked as clones i.e. the probability of inserting a single plasmid is 5×10^{-7} . It follows that the likelihood of a second plasmid being inserted into an already transfected cell (irrespective of whether they are independent events or not) is lower still. Finally, S is taken to be the amount of X (Op) which is required to have half of the binding sites occupied. Although not biologically relevant this was taken to be 0.5. Thus, using this representation, transcription is allowed only if Op is not bound to a TetR₂ molecule, otherwise $[Op]=0$ and the reaction is shut down. By relating transcription directly to whether Op is bound or not, rather than the amount of inhibitive TetR₂ molecules available, a direct estimate of ongoing transcription is generated.

7.2.1.3 Assumptions made for the analysis of the stable Goodwin oscillator behaviour in *T. brucei* cells

Several assumptions were made during the development of the mathematical model, in addition to those already listed above. Assumptions serve to simplify the complex biological environment, since reflecting all biochemical processes within the network would lead to a computationally infeasible model. The level of detail is similar to other synthetic biology studies carried out (Fung et al., 2005; Danino et al., 2010; Tigges et al., 2010). Firstly, gene expression was modelled using transcription and translation, rather than a set of more complex initiation-elongation-termination steps. Secondly, time-delay in protein synthesis was not modelled explicitly. Instead, it is inherently present in the transcription-translation(-dimerisation) process which results in a delay

between initiation and finalisation of the protein products. The rate of transcription of *tetR* and *eGFP* genes is assumed to be the same since it is initiated by the same promoter. The rate of translation for both mRNA_{TetR} and mRNA_{eGFP} is also considered equal. The same goes for the degradation rate of mRNA_{TetR} and mRNA_{eGFP} and the degradation of the TetR monomer, TetR dimer and bound TetR dimer (Stricker et al., 2008; Tigges et al., 2010).

Assumptions made by the Gillespie SSA are as discussed in Gillespie (1976) and Gillespie (1977).

7.2.2 Results and analysis of Gillespie time series simulations to inform screening analysis of the Goodwin oscillator parameter space

Having developed the mathematical model of the trypanosomal Goodwin oscillator, the next step was to carry out Gillespie time series simulations for the scope of parameter screening of the 11-dimensional parameter space specified above. The following section presents results from Gillespie simulations.

7.2.2.1 Time series trajectories of the Goodwin oscillator species

Gillespie time series simulations were set to model component dynamics over a time-span of 1000 minutes. The quantity of each species was recorded at 0.5 minute intervals. 500 simulations were carried out, implying 500 different parameter combinations were tested based on simple random sampling. Carrying out an exhaustive scan of the parameter space would have been infeasible since there are over 4×10^{100} possible parameter combinations based on the above uniform distributions and the power of the ABC-Sysbio program. The set of simulations can be considered to represent a population of 500 cells rather than just one cell sampled for 500 times. A decision was taken to carry out 500 simulations (as opposed to, for example, 1000 simulations) as a representative sample of the number of cells which are in the plane of view (average of 393 cells) during microscopy given that cells are embedded at a concentration of 5×10^7 cells/ml in 40% Cygel as per Section 5.7.

Unless otherwise specified, starting molecule numbers were set at 0 for all species except for the tet operator, Op, which was set at value 1. Figure 7.11 shows the 500 time series trajectories for mRNA_{TetR}, TetR, TetR₂, mRNA_{eGFP} and eGFP. The time series trajectories for Op and TetR₂.Op are not presented since they have a digital profile which switches between 0 and 1 which offers no new information. Since the plots were very dense and difficult to distinguish, 10 simulations were selected randomly and plotted in Figure 7.12.

The plots show that species are oscillating continually. Op and TetR₂.Op (not shown) switch between on and off states. On average, the tet operator, Op, is bound to a TetR₂ molecule, implying that on average, protein expression within the system is being inhibited. Protein dynamics of the same species are not synchronised over the different simulations, except for a calibration period at the beginning of the simulation, as reflected in the average time-series (black band). This was calculated via MATLAB by averaging the simulated time series data under consideration.

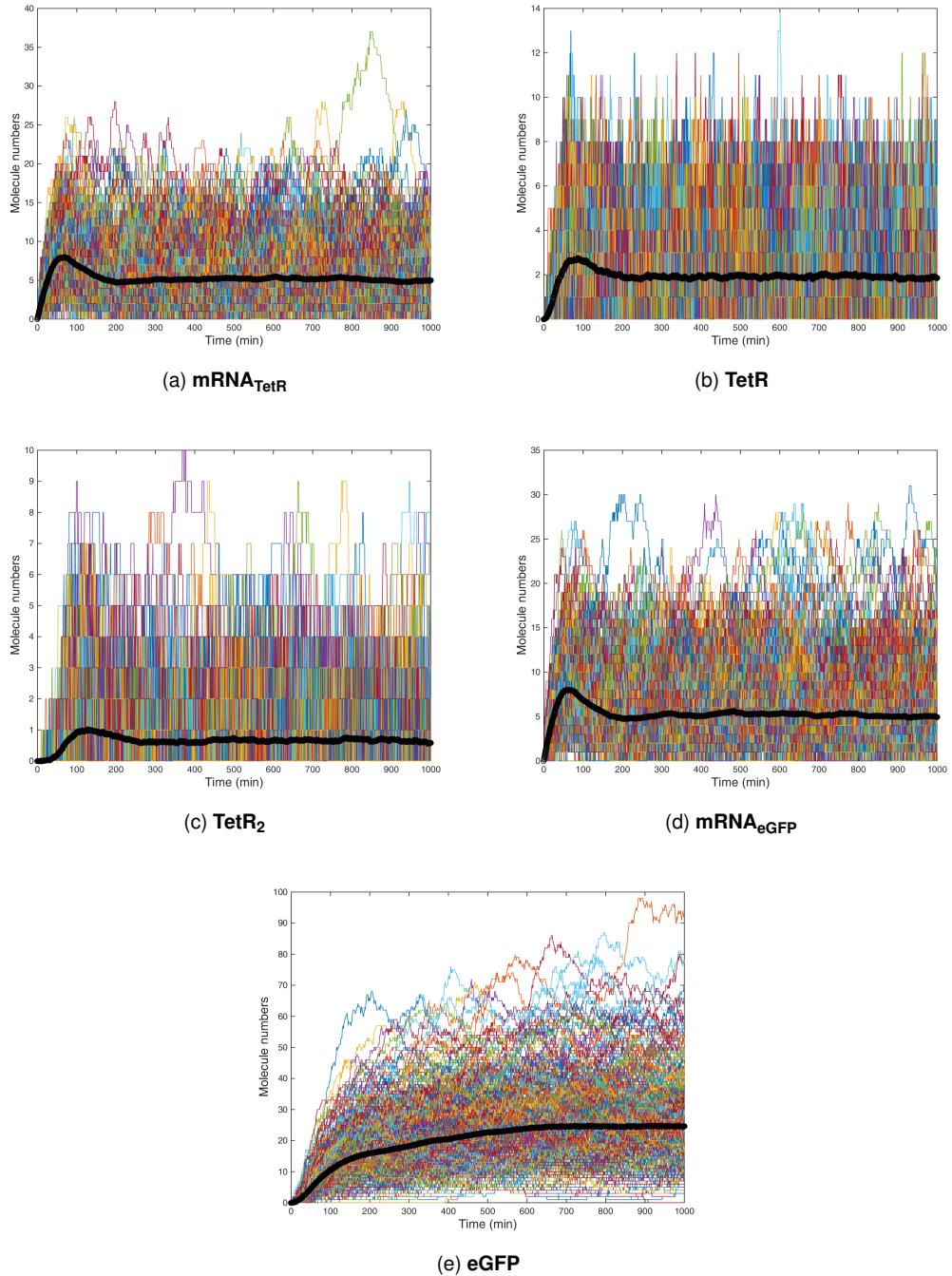


Figure 7.11: Time series trajectories of 500 Gillespie simulations over a time-span of 1000 minutes for the species: a) $\text{mRNA}_{\text{TetR}}$, b) TetR , c) TetR_2 , d) $\text{mRNA}_{\text{eGFP}}$ and e) eGFP . The axis show Number of molecules vs Time in minutes. The band in black represents the average number of species molecules. Starting conditions are set to 0 for all species except $\text{Op}=1$.

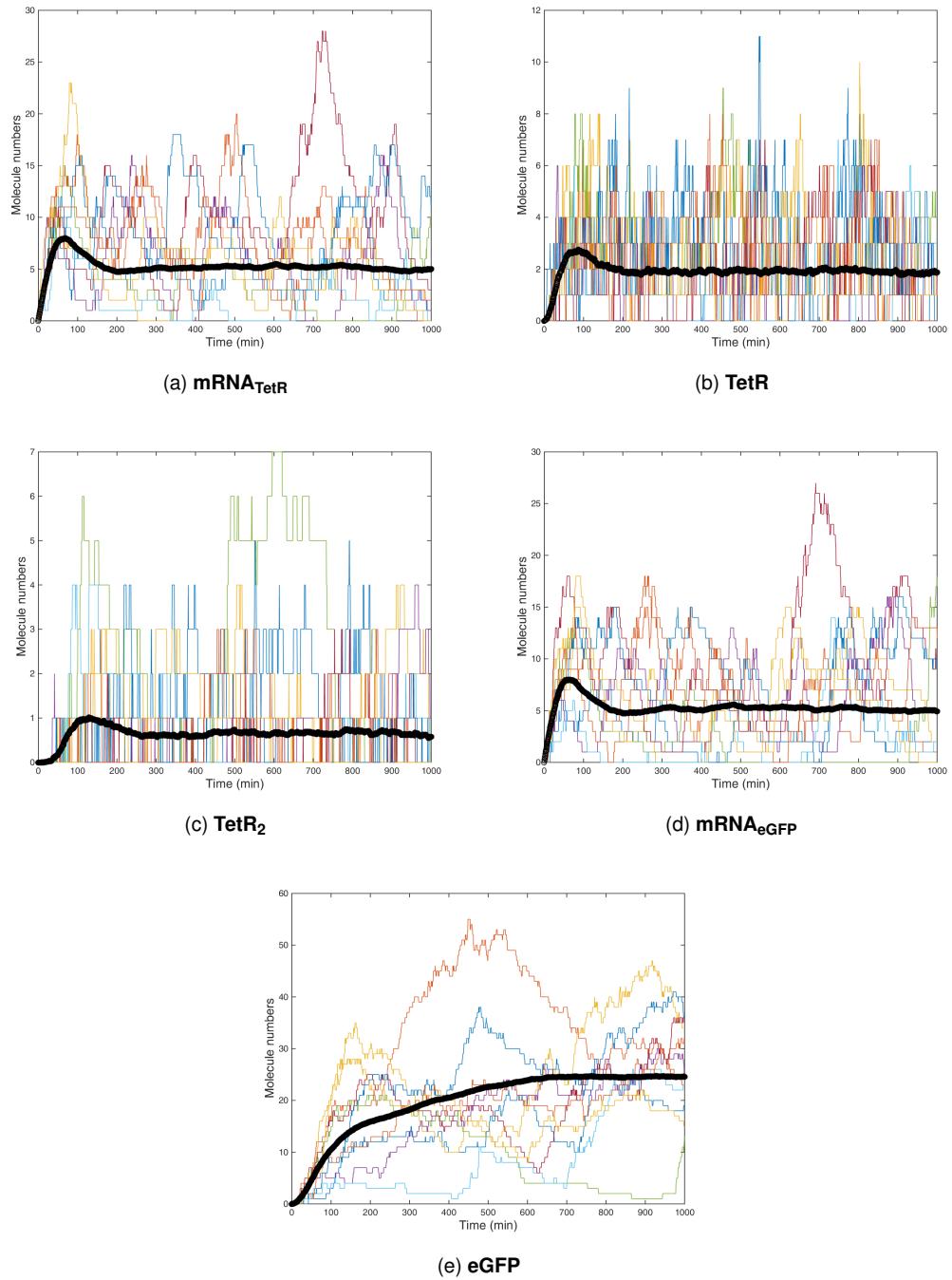


Figure 7.12: Time series trajectories of 10 randomly selected Gillespie simulations over a time-span of 1000 minutes for the species: a) $\text{mRNA}_{\text{TetR}}$, b) TetR , c) TetR_2 , d) $\text{mRNA}_{\text{eGFP}}$ and e) eGFP . The axis show Number of molecules vs Time in minutes. The band in black represents the average number of species molecules over time. Starting conditions are set to 0 for all species except $\text{Op}=1$.

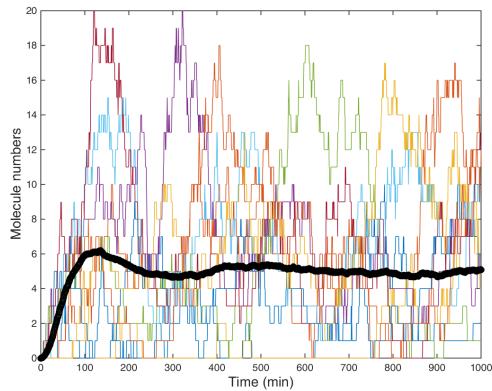


Figure 7.13: eGFP time series trajectories of 10 randomly selected Gillespie simulations over a time-span of 1000 minutes. The eGFP degradation rate was set to 0.02min^{-1} instead of 0.004min^{-1} . The axis show Number of molecules vs Time in minutes. The band in black represents the average number of species molecules over time. Starting conditions are set to 0 for all species except Op=1.

The type of fluctuations differ according to species; mRNA_{TetR} and mRNA_{eGFP} have similar profiles. TetR and TetR₂ have a lower amplitude and shorter period than their eGFP counterpart. All species, except for eGFP have sinusoidal oscillations. eGFP oscillations, on the other hand, are sinusoidal with an increasing trend as seen in Elowitz and Leibler (2000). The latter is attributed to the degradation rate of eGFP which does not allow for a sufficiently quick turnover of eGFP protein for the oscillation to reset itself, but instead builds on previous eGFP levels. In fact, when simulations were carried out using a higher eGFP degradation rate oscillations have a more uniform sinusoidal pattern and lower amplitudes, as seen in Figure 7.13. The plot shows 10 randomly selected simulations of eGFP quantity profiles, similarly to the above, using a degradation rate of 0.02min^{-1} instead of the experimentally determined 0.004min^{-1} .

7.2.2.2 Time series trajectories for different starting conditions

500 Gillespie simulations were carried out using different starting conditions, wherein initial quantities for the mRNA_{TetR}, TetR, TetR₂, mRNA_{eGFP} and eGFP species was set to 1 (not shown, results similar to Figure 7.11), 10 (not shown, results similar to Figure 7.11), a random number between 1 and 20 (Figure 7.14) and 50 (Figure 7.15). Throughout, Op was set to 1, while TetR₂.Op was set to 0. Finally, in Figure 7.16, all species were set to 0 except for TetR₂.Op which was set to 1. Each plot shows 10 randomly selected simulations which allow for visual assessment of expression dynamics of the mRNA and protein components.

Simulated trajectories repeatedly indicate that protein dynamics are independent of initial conditions. Following a brief calibration period, trajectories are similar to those observed in the original simulations (Figure 7.12). This is a positive result as it indicates that oscillations are likely to occur irrespective of the initial state of the population of cells.

7.2.2.3 Time series trajectories of TetR and eGFP proteins

The above sections presented results for expression dynamics throughout the device encoding the Goodwin oscillator. However, interest lay in TetR₂ and eGFP protein quantity trajectories.

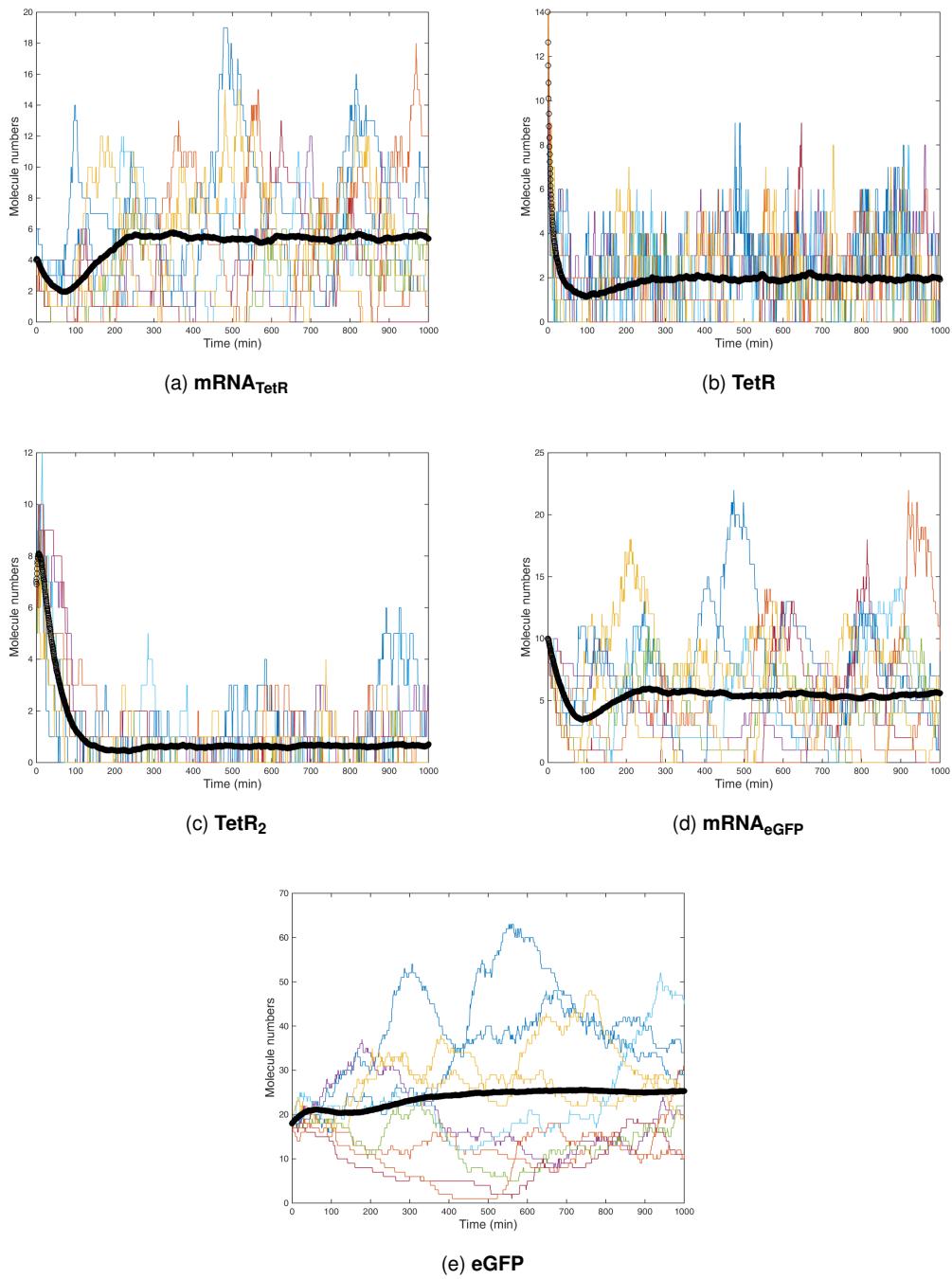


Figure 7.14: Starting molecule numbers set randomly from 1 to 20 for all species except $\text{Op}=1$, $\text{TetR}_2\text{-Op}=0$: Time series trajectories for 50 randomly selected Gillespie simulations over a time-span of 1000 minutes for the species: a) $\text{mRNA}_{\text{TetR}}$, b) TetR , c) TetR_2 , d) $\text{mRNA}_{\text{eGFP}}$ and e) eGFP . The axis show Number of molecules vs Time in minutes. The band in black represents the average number of molecules over time.

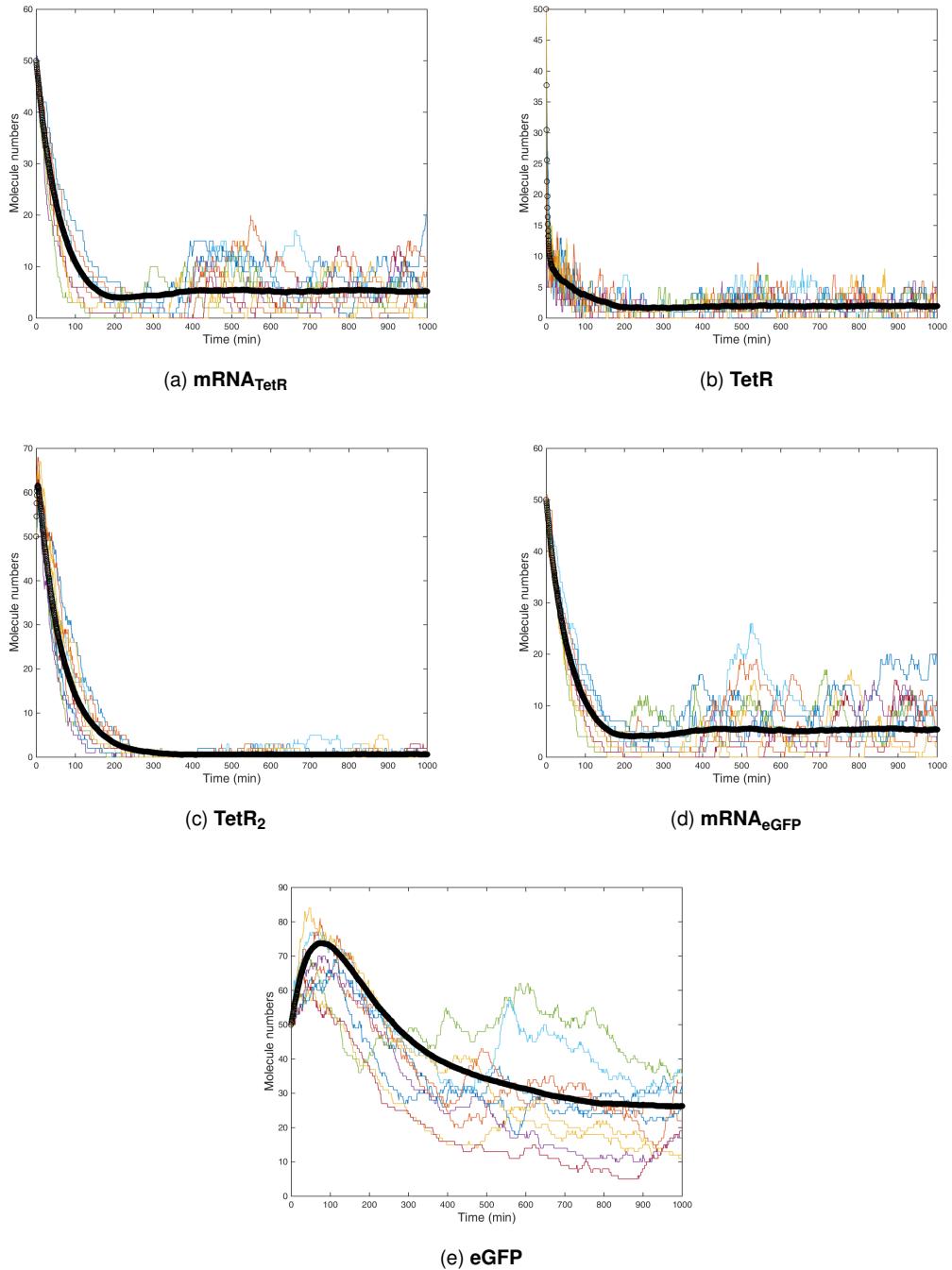


Figure 7.15: Starting molecule numbers set to 50 for all species except $\text{Op}=1$, $\text{TetR}_2, \text{Op}=0$:
 Time series trajectories for 10 randomly selected Gillespie simulations over a time-span of 1000 minutes for the species: a) $\text{mRNA}_{\text{TetR}}$, b) TetR , c) TetR_2 , d) $\text{mRNA}_{\text{eGFP}}$ and e) eGFP . The axis show Number of molecules vs Time in minutes. The band in black represents the average number of molecules over time.

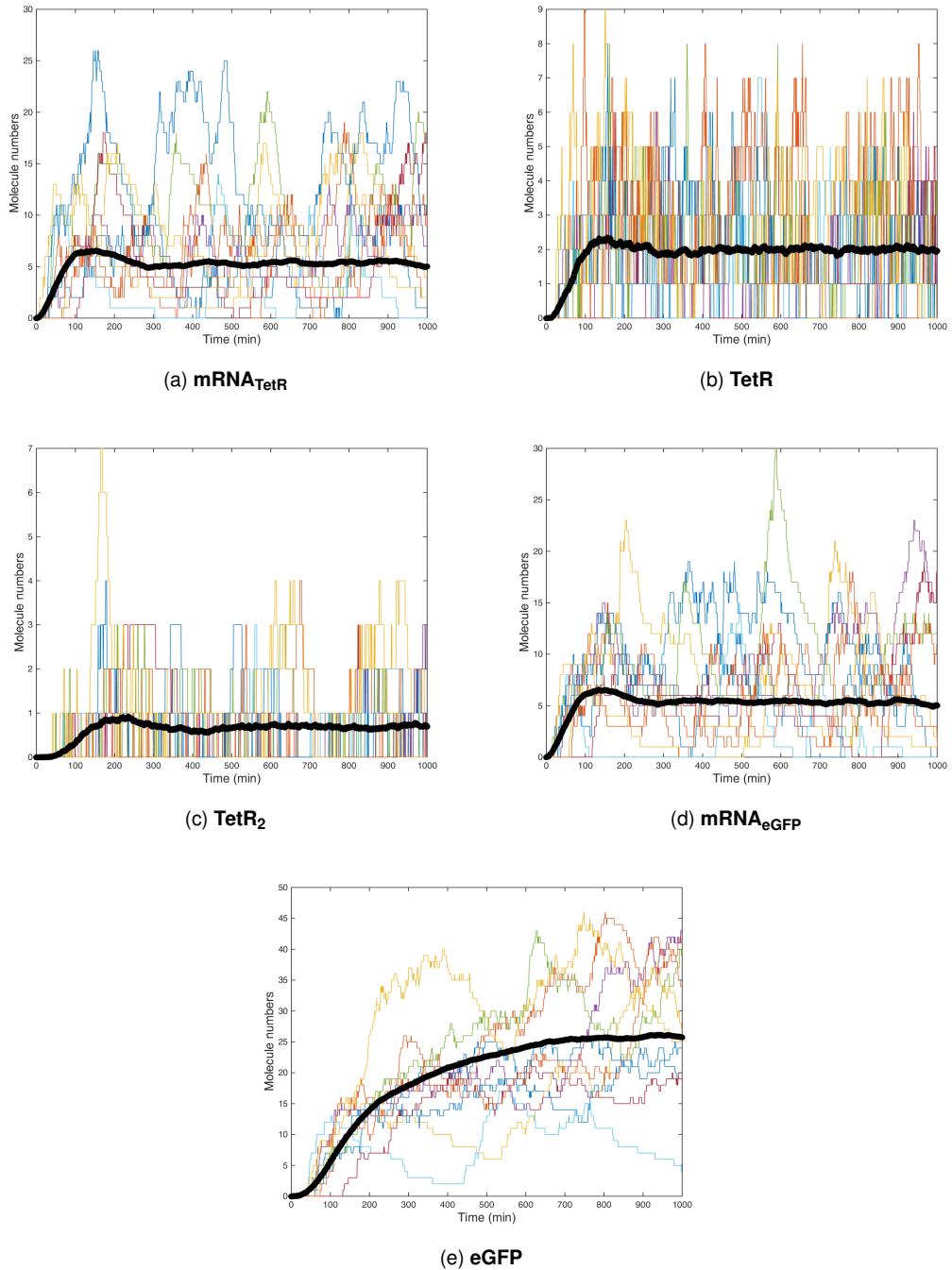


Figure 7.16: **Starting molecule numbers set to 0 for all species except TetR_2 . Op=1**: Time series trajectories for 10 randomly selected Gillespie simulations over a time-span of 1000 minutes for the species: a) $\text{mRNA}_{\text{TetR}}$, b) TetR , c) TetR_2 , d) $\text{mRNA}_{\text{eGFP}}$ and e) eGFP . The axis show Number of molecules vs Time in minutes. The band in black represents the average number of molecules over time.

Analysis via MATLAB confirmed that oscillations were present in all 500 simulations of both TetR₂ and eGFP protein. Analysis was based on a custom-written algorithm which detected decreasing inclines and troughs within each simulation and flagged up those simulations wherein protein quantities remained constant or almost constant or were monotonic.

Figures 7.17 and 7.18 show twelve individual randomly selected trajectories from simulations of the TetR₂ and eGFP protein respectively, using different parameters in each case. The individual trajectories reveal a broad range of oscillation characteristics, such as frequency and amplitude, with a linear growth trend observed in a proportion of the eGFP trajectories. The source of this variety and whether it is of a stochastic or deterministic nature is unknown. Moreover, it is unclear whether the fluctuations themselves are genuine or of a stochastic nature. In order to eliminate noise sources, the protein trajectories can be simulated using deterministic methods and compared to simulations derived from an SSA.

7.2.2.4 Parameter analysis based on TetR₂ dynamics

Given the variety of oscillation characteristics seen in Figure 7.17 and Figure 7.18 for TetR₂ and eGFP protein quantities, analysis was carried out to better understand the source of this variation and analyse whether specific parameter combinations led to specific characteristics. The maximum coordinate was selected as the characteristic of interest in order to help inform microscopy results and observed fluorescence patterns. Note that the maximum coordinate is not equivalent to the amplitude. A large maximum coordinate can be correlated to observed fluorescence, unlike the amplitude which does not necessarily have the same correlation due to linear growth trends which can mask the true amplitude. The maximum coordinate (in terms of protein molecule number) for each of the 500 trajectories of the TetR₂ and eGFP components was recorded and plotted as a histogram, as seen in Figure 7.19. The plot indicates that the maximum coordinate data can be fitted to a Poisson distribution, although this was not the scope of the analysis.

Based on the information represented in Figure 7.19, maximum coordinates of TetR₂ were divided into low (proteins < 3), medium (3 ≤ proteins < 5) and high (5 ≤ proteins) as based on first and third quartile values. Due to the Poisson distribution of the maximum coordinates, the track sizes are unequal. The reaction rate parameter combinations which led to each set of TetR₂ kinetics were then extracted (Figure 7.20, 7.21 and 7.22), in order to assess whether specific parameter combinations favoured certain trajectories.

Data is represented as a matrix plot. Along the diagonal is a histogram for each of the parameters (reaction rates) being analysed. The histograms correspond to the parameter labels and range of values along the x-axis. The y-axis for the histograms, the frequency, is not labelled due to space constraints. The remainder of the plot consists of scatter plots which showcase relations between each pair of parameters. Each data point relates to a maximum coordinate which is being analysed. The rate of degradation of eGFP is not included since it has a fixed value.

A bias can be observed between the rates of transcription, translation and degradation which lead to the three sets of maximum TetR₂ coordinates, as seen in the histograms along the diagonal of each matrix plot (Figure 7.20 - 7.22). Combinations of low transcription and translation rates plus high mRNA, TetR and TetR₂ degradation rates result in small quantities of protein with a high

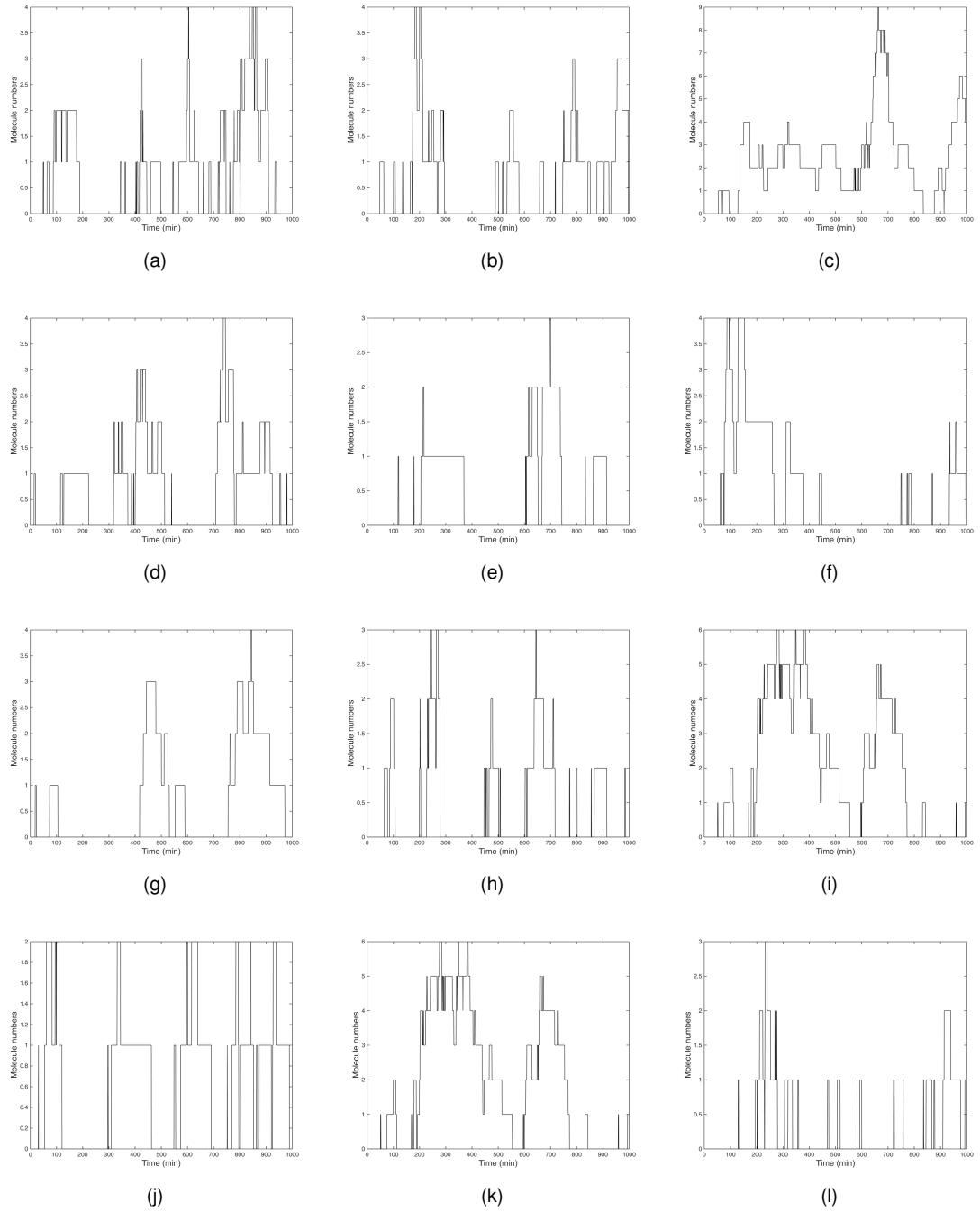


Figure 7.17: TetR_2 dynamics: Individual plots of 12 randomly selected TetR_2 protein quantity trajectories simulated with the Gillespie SSA using different parameter sets. The axis show Number of molecules vs Time in minutes. Starting conditions were set to 0 for all species except $\text{Op}=1$.

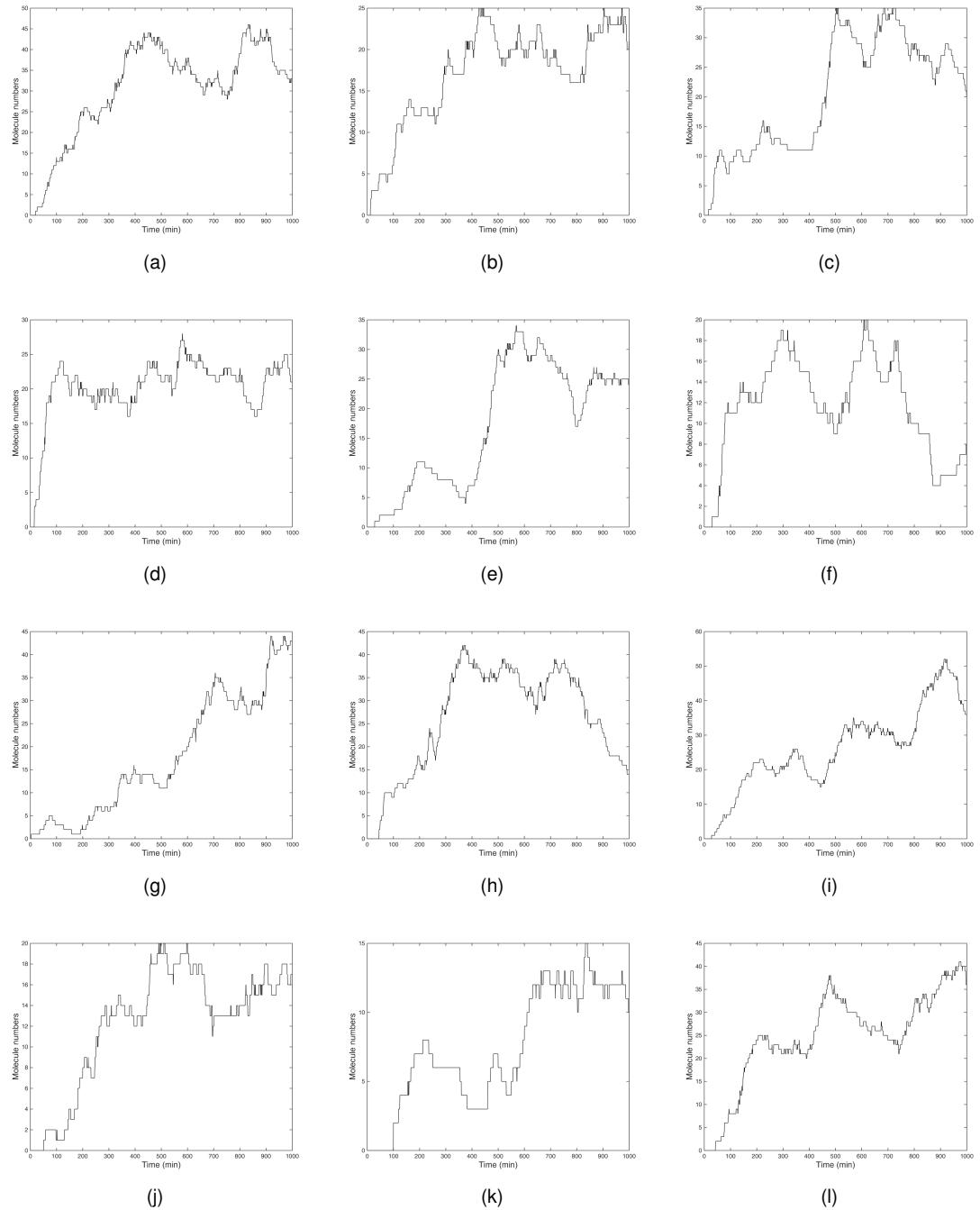


Figure 7.18: eGFP dynamics: Individual plots of 12 randomly selected eGFP protein quantity trajectories simulated with the Gillespie SSA using different parameter sets. The axis show Number of molecules vs Time in minutes. Starting conditions were set to 0 for all species except $Op=1$.

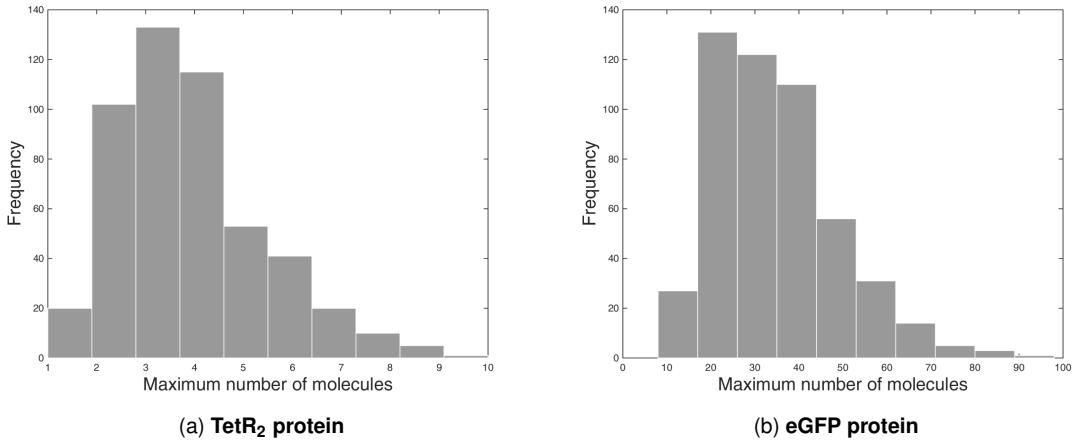


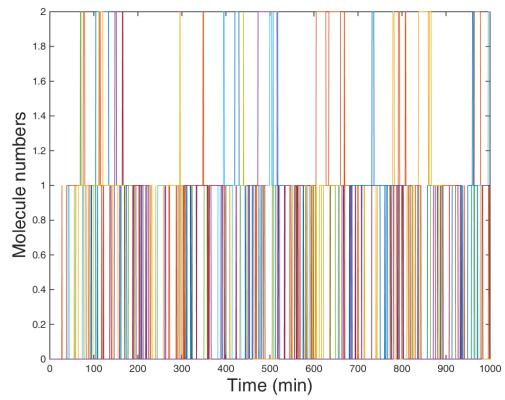
Figure 7.19: **Maximum number of molecules:** Histogram plots of the maximum number of molecules attained during the 500 simulations for the protein trajectories of: a) TetR₂ and b) eGFP.

turnover and favour low value TetR₂ oscillations (Figure 7.20). There are no clear trends in the parameter values and combinations which lead to medium value oscillations (Figure 7.21). A more pronounced trend can be seen in parameter combinations which led to high value oscillations (Figure 7.22), with high transcription and translation rates plus low RNA and TetR degradation rates. This combination results in larger protein quantities with a low protein turnover, leading to a build-up of protein. It should be noted that the bias in the synthesis and degradation rates is not pronounced. In fact, in each of the rates, over 50% of the simulated values were scattered across the remaining range of possible values. Coupled with the fact that the other parameters analysed did not demonstrate a clear trend, this further highlights the importance of distinguishing between genuine and stochastic fluctuations in protein dynamics.

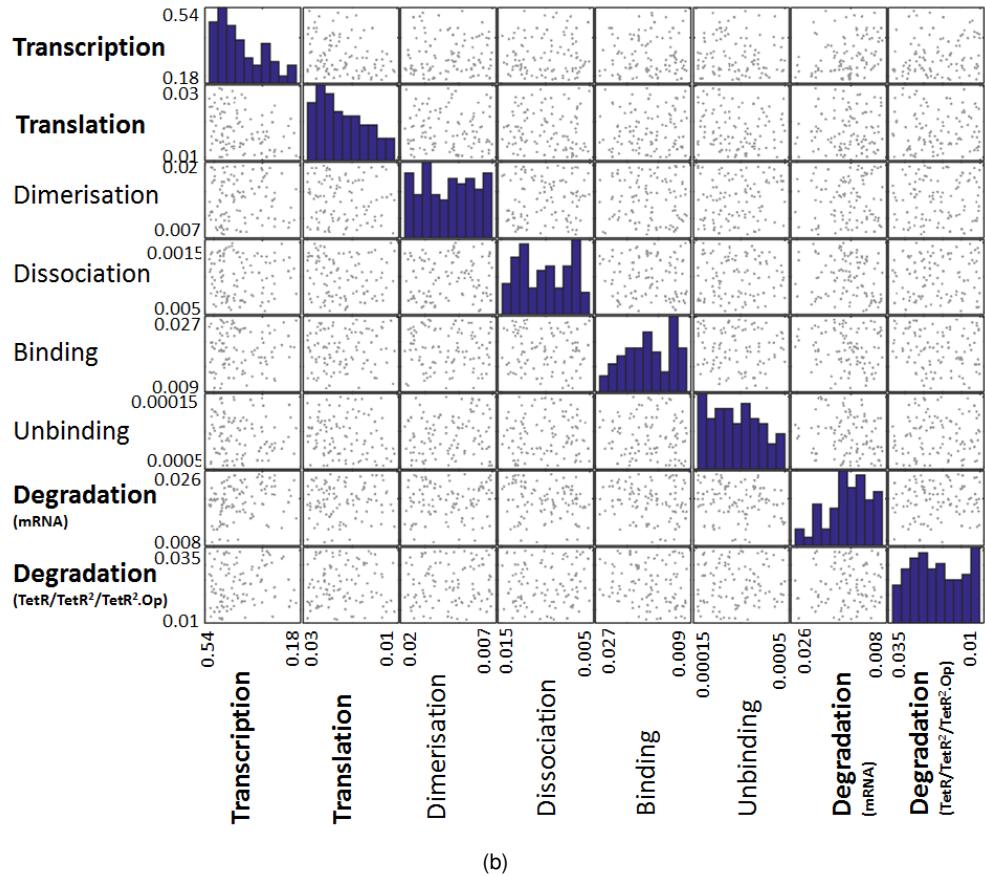
7.2.2.5 Parameter analysis based on eGFP dynamics

Similarly to the above analysis with TetR₂, parameter combinations which led to low, medium or high maximum coordinate eGFP oscillations were analysed in order to extract parameter trends. The analysis was carried out on both TetR₂ and eGFP related biochemical parameters within the network rather than just eGFP-specific rates. This is because TetR₂ has a regulatory influence over eGFP and hence, TetR₂ patterns have the potential to affect eGFP oscillations. Based on first and third quartile data from Figure 7.19b, low oscillations were considered to be those where the maximum eGFP coordinate is at less than 23 proteins (Figure 7.23a). Analysis of the histograms and scatter plots of the rate parameters which correspond to this outcome (Figure 7.23b) revealed a trend for low transcription and translation rates in addition to high mRNA degradation rates. This is similar to results observed for low amplitude TetR₂ oscillations. TetR degradation rates tended to be low, implying that the tetracycline operator is less likely to be unbound and hence eGFP expression is less likely to take place.

For medium level oscillations, having a maximum coordinate between 24 and 42 proteins (Figure 7.24a), parameter analysis (Figure 7.24b) was inconclusive. For high value oscillations having a maximum coordinate of over 42 proteins (Figure 7.25a), parameter trends were the opposite of those observed for low level oscillations, with high transcription and translation rates favoured

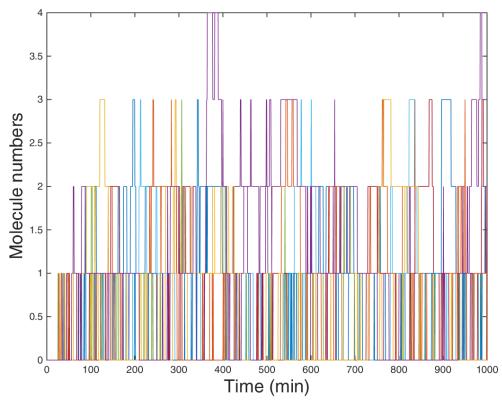


(a)

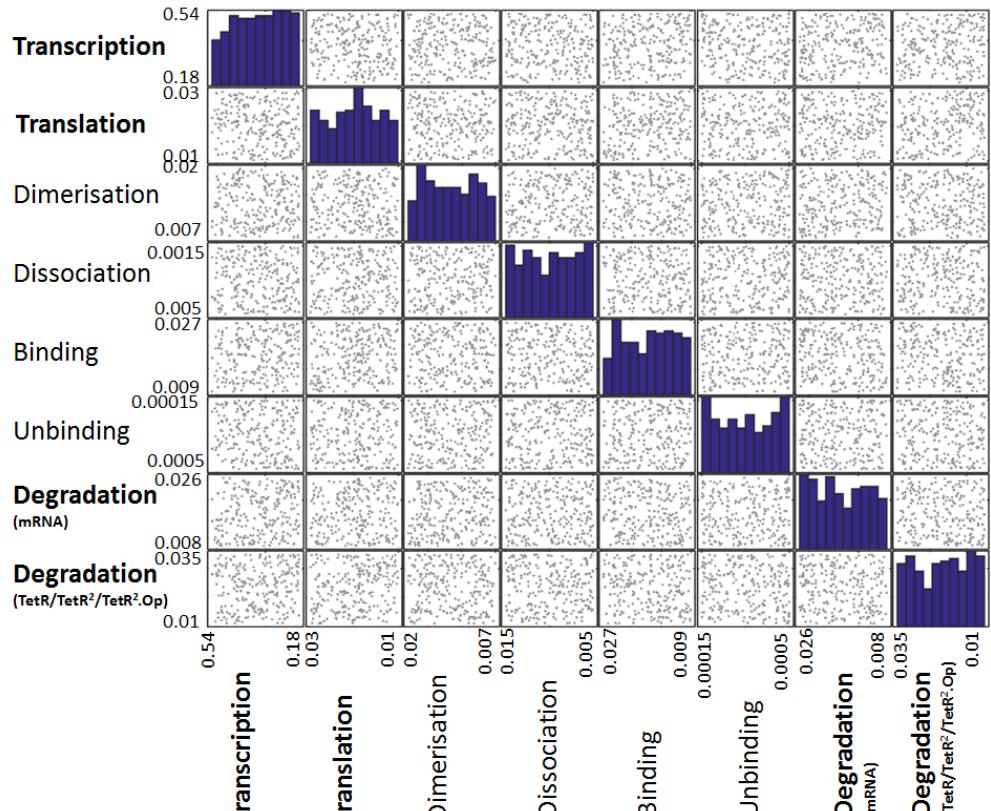


(b)

Figure 7.20: Analysing the parameter space which results in low value (<3 molecules) TetR₂ oscillations: a) TetR₂ trajectories when the maximum coordinate is less than 3 proteins. These were selected from the original 500 simulations carried out. The axis show Number of molecules vs Time in minutes. b) The corresponding matrix plot showing the values of and the relations between the eight distinguishable and non-fixed biochemical parameters which lead to the trajectories in (a): the rate of transcription, rate of translation, rate of dimerisation, rate of dissociation, rate of binding, rate of unbinding and degradation rate for the mRNA and protein components. The histograms along the diagonal highlight the distribution of values of each parameter. The scatter plots show the relation between each pair of parameters. The x- and y- axis show the range of possible values for each of the parameters.



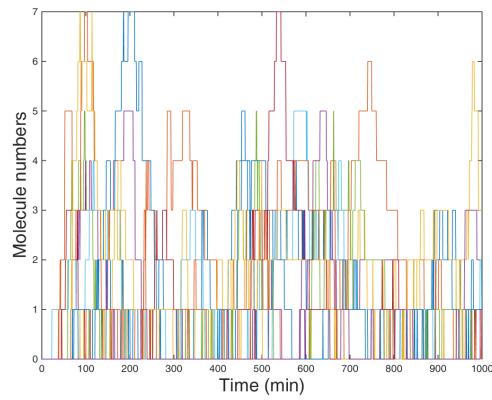
(a)



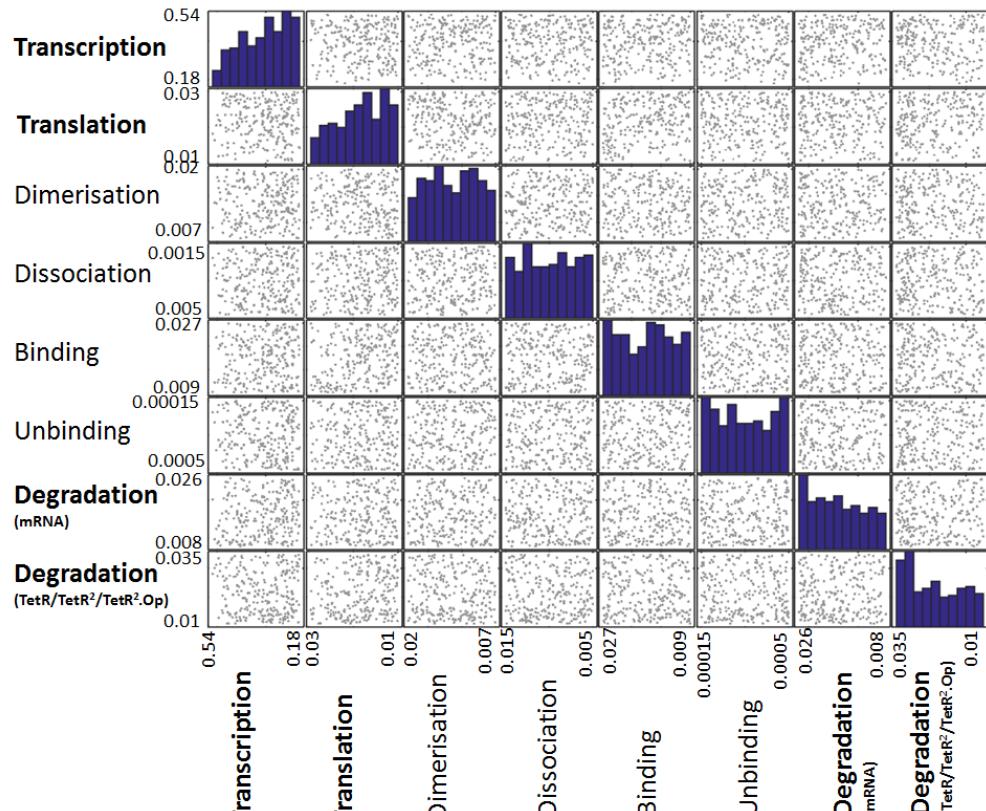
(b)

Figure 7.21: **Analysing the parameter space which results in medium value (3-4 molecules) TetR₂ oscillations:**

TetR₂ oscillations: a) TetR₂ trajectories when the maximum coordinate is 3-4 proteins. These were selected from the original 500 simulations carried out. The axis show Number of molecules vs Time in minutes. b) The corresponding matrix plot showing the values of and the relations between the eight distinguishable and non-fixed biochemical rate parameters which lead to the trajectories in (a): the rate of transcription, rate of translation, rate of dimerisation, rate of dissociation, rate of binding, rate of unbinding and degradation rate for the mRNA and protein components. The histograms along the diagonal highlight the distribution of values of each parameter. The scatter plots show the relation between each pair of parameters. The x- and y- axis show the range of possible values for each of the parameters.



(a)



(b)

Figure 7.22: **Analysing the parameter space which results in high value (>4 molecules) TetR₂ oscillations:**

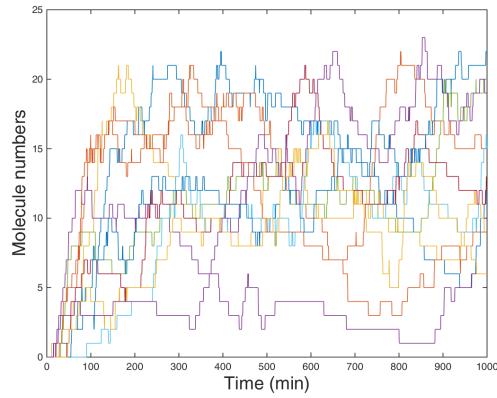
a) TetR₂ trajectories when the maximum coordinate is more than 4 proteins. These were selected from the original 500 simulations carried out. The axis show Number of molecules vs Time in minutes. b) The corresponding matrix plot showing the values of and the relations between the eight distinguishable and non-fixed biochemical rate parameters which lead to the trajectories in (a): the rate of transcription, rate of translation, rate of dimerisation, rate of dissociation, rate of binding, rate of unbinding and degradation rate for the mRNA and protein components. The histograms along the diagonal highlight the distribution of values of each parameter. The scatter plots show the relation between each pair of parameters. The x- and y- axis show the range of possible values for each of the parameters.

along with low mRNA degradation rates and high TetR degradation rates. The latter decreases the probability of having a TetR₂.Op complex, leading to expression and increasing eGFP numbers. Thus, similarly to conclusions made for TetR₂ oscillation analysis, there is a small output bias based on the parameter combinations which requires further analysis, especially of the noise component.

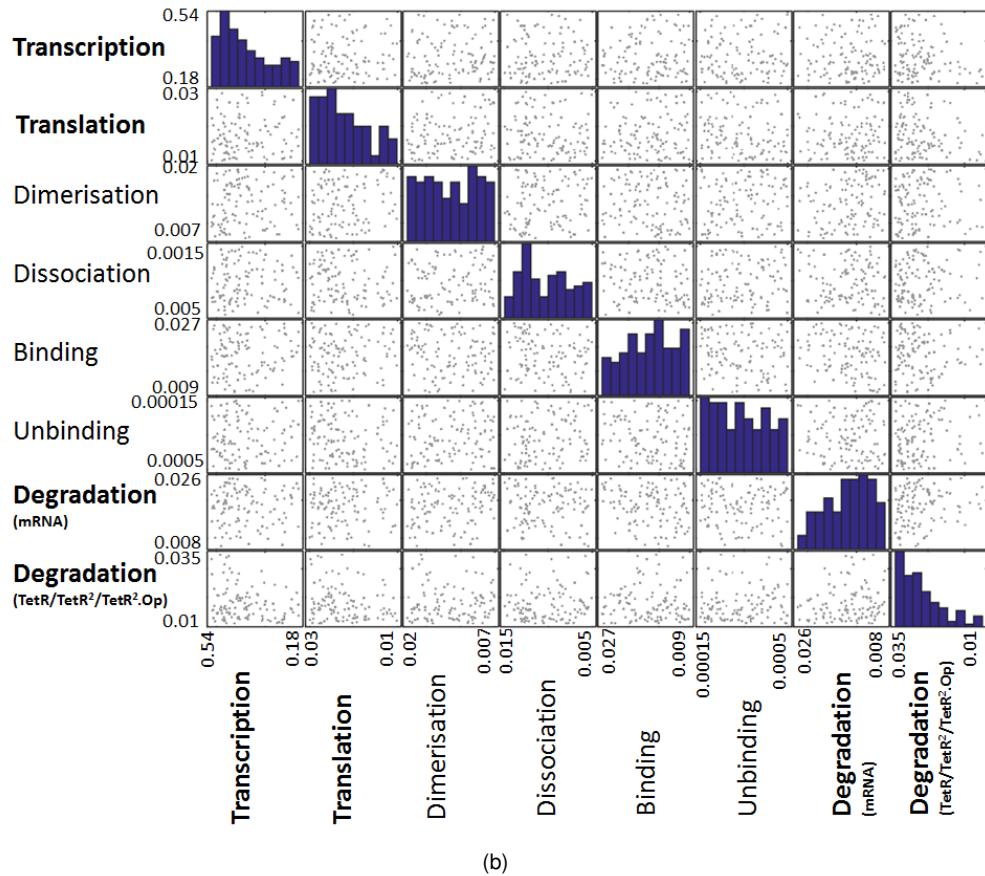
7.2.3 Evaluation of the Goodwin oscillator parameter space following screening analysis

Following preliminary stochastic and quantitative mathematical modelling of protein expression dynamics within the trypanosomal Goodwin oscillator, the following findings and observations were made:

1. A mathematical model was developed to represent the pStable and pUnstable Goodwin oscillators as a set of components, such as mRNA and proteins, which interact with each other via biochemical reactions, such as synthesis and degradation. These biochemical reactions govern how component quantities vary over time. Based on the available literature, this is the first mathematical model of a synthetic GRN used to inform outcome within a trypanosomal context.
2. A screening assay of the 11-dimensional parameter space of the model was carried out via Gillespie stochastic simulations which used different parameter combinations at each iteration in order to scan which combinations led to the onset of oscillations of protein quantities. Results repeatedly showed that parameters from across the whole spectrum led to low-amplitude fluctuations in TetR₂ and eGFP mRNA and protein quantities.
3. Results indicated that dynamics were independent of starting conditions. Simulations showed that in the scenarios which were tested wherein components had different numbers of molecules at time 0, dynamics resumed their original profile following a brief calibration period.
4. The components exhibited different types of sinusoidal oscillations; oscillations for mRNA_{TetR} and mRNA_{eGFP} had a similar profile which had a higher amplitude and longer period than the oscillatory profile of TetR and TetR₂. eGFP oscillations were sinusoidal with an increasing trend, attributed to the degradation rate of eGFP which does not allow for a sufficiently quick turnover of eGFP protein in line with the oscillation period but instead builds on previous eGFP levels.
5. Oscillations were initially in synchrony, as indicated by the average protein trajectory calculated for each component. After one oscillation, the cells (each of which was using a different combination of parameters) became desynchronised.
6. Oscillation characteristics, such as maximum coordinate, amplitude and period, varied between simulation runs. Analysis showed that certain parameter rate values, specifically synthesis and degradation rates, show a bias towards high or low maximum coordinates. The remainder of the parameters do not show a clear bias. In light of this and the wide range of trajectories which were exhibited, it becomes necessary to distinguish between genuine and stochastic fluctuations. This can be done by developing a mathematical model of control plasmid pBC188 to observe whether fluctuations within the system lead to similar

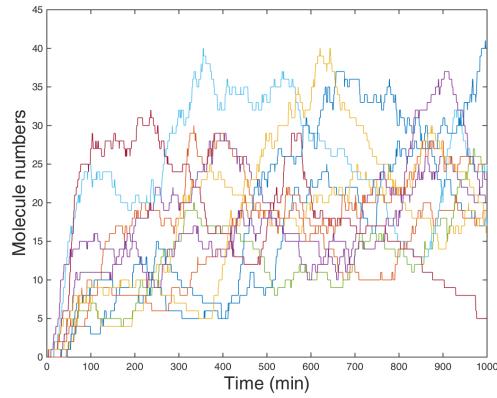


(a)

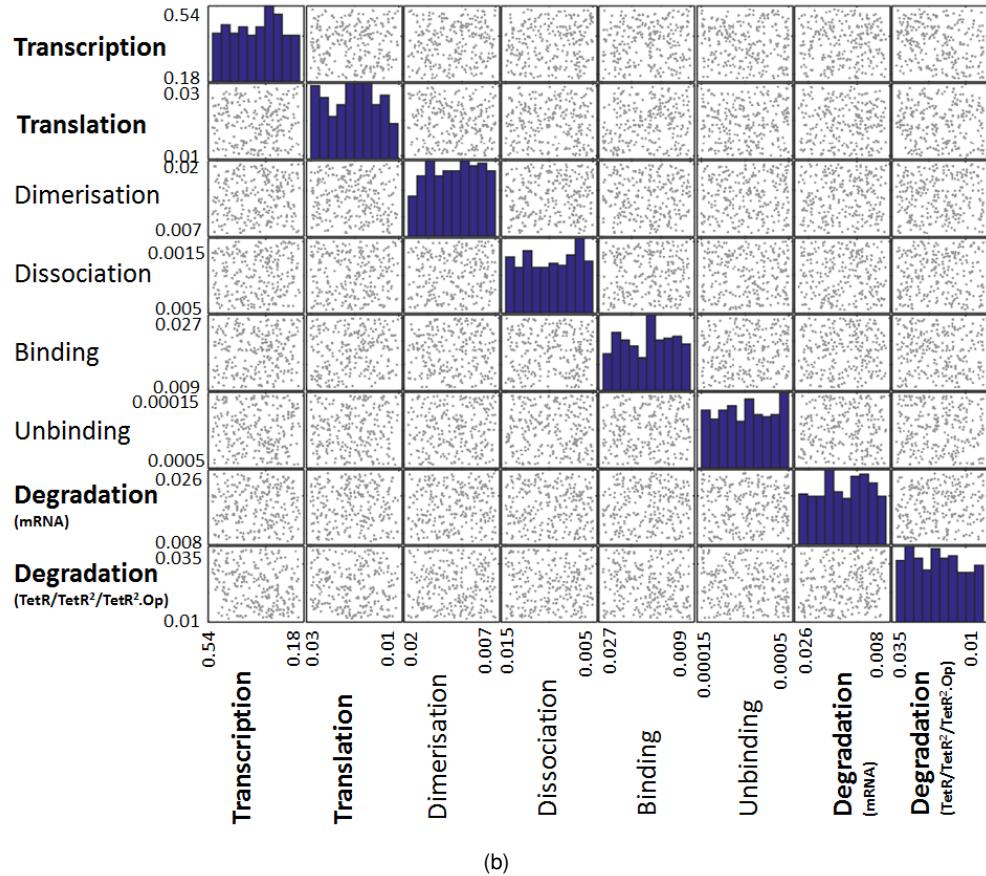


(b)

Figure 7.23: Analysing the parameter space which results in low value (<24 molecules) eGFP oscillations: a) eGFP trajectories when the maximum coordinate is less than 24 proteins. These were selected from the original 500 simulations carried out. The axis show Number of molecules vs Time in minutes. b) The corresponding matrix plot showing the values of and the relations between the eight distinguishable and non-fixed biochemical rate parameters which lead to the trajectories in (a): the rate of transcription, rate of translation, rate of dimerisation, rate of dissociation, rate of binding, rate of unbinding and degradation rate for the mRNA and protein components. The histograms along the diagonal highlight the distribution of values of each parameter. The scatter plots show the relation between each pair of parameters. The x- and y- axis show the range of possible values for each of the parameters.

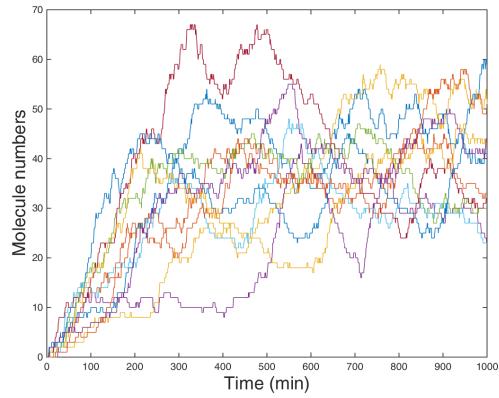


(a)

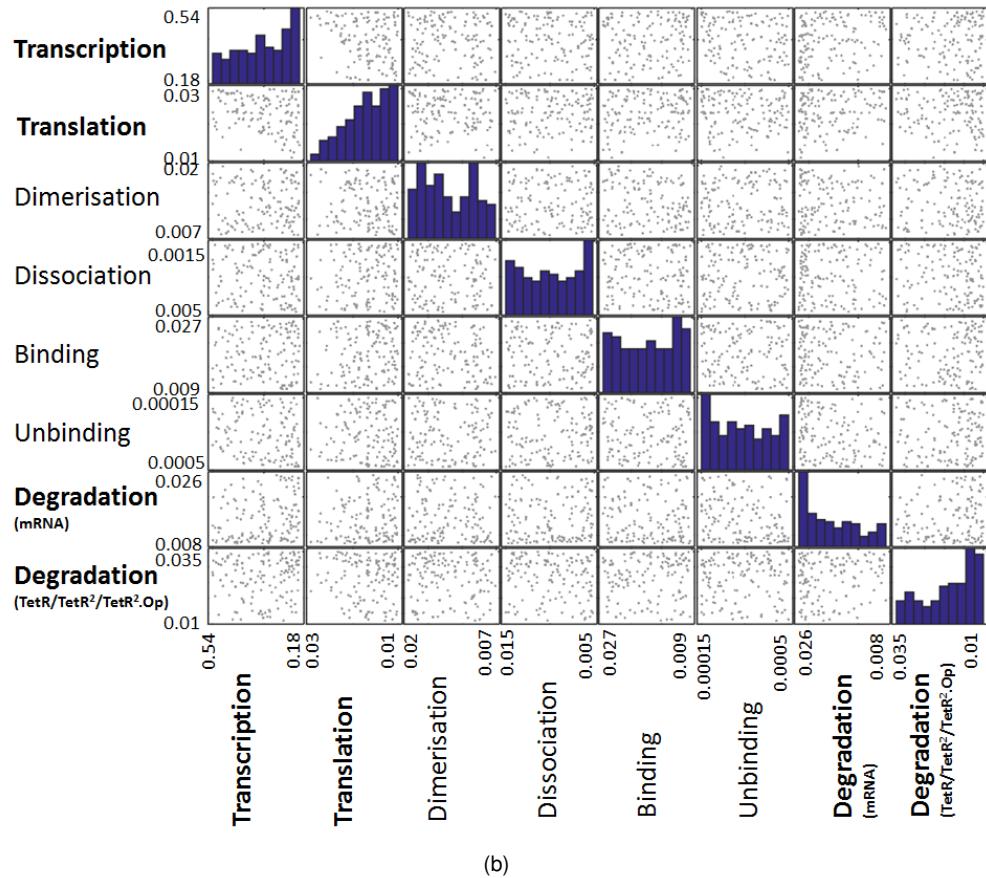


(b)

Figure 7.24: Analysing the parameter space which results in medium value (24-42 molecules) eGFP oscillations: a) eGFP trajectories when the maximum coordinate is 24-42 proteins. These were selected from the original 500 simulations carried out. The axis show Number of molecules vs Time in minutes. b) The corresponding matrix plot showing the values of and the relations between the eight distinguishable and non-fixed biochemical rate parameters which lead to the trajectories in (a): the rate of transcription, rate of translation, rate of dimerisation, rate of dissociation, rate of binding, rate of unbinding and degradation rate for the mRNA and protein components. The histograms along the diagonal highlight the distribution of values of each parameter. The scatter plots show the relation between each pair of parameters. The x- and y- axis show the range of possible values for each of the parameters.



(a)



(b)

Figure 7.25: Analysing the parameter space which results in high value (>42 molecules) eGFP oscillations: a) eGFP trajectories when the maximum coordinate is more than 42 proteins. These were selected from the original 500 simulations carried out. The axis show Number of molecules vs Time in minutes. b) The corresponding matrix plot showing the values of and the relations between the eight distinguishable and non-fixed biochemical rate parameters which lead to the trajectories in (a): the rate of transcription, rate of translation, rate of dimerisation, rate of dissociation, rate of binding, rate of unbinding and degradation rate for the mRNA and protein components. The histograms along the diagonal highlight the distribution of values of each parameter. The scatter plots show the relation between each pair of parameters. The x- and y- axis show the range of possible values for each of the parameters.

Strain	Plasmid introduced	Conditions	Section
1	BSF	pStable	Uninduced
2	BSF	pStable	Induced with tc_{thres}
3	PCF	pStable	Uninduced
4	PCF	pStable	Induced with tc_{thres}

Table 7.7: The pStable-transfected BSF and PCF *T. brucei* clones which were imaged using live-cell fluorescent and brightfield microscopy to observe fluorescent patterns and characterise the phenotype of pStable. ‘ tc_{thres} ’ is the tetracycline concentration above which protein expression across the cell population is induced.

fluctuations. Moreover, a deterministic model wherein all sources of stochasticity are eliminated should be simulated in order to observe whether a similar output as for the stochastic model is achieved.

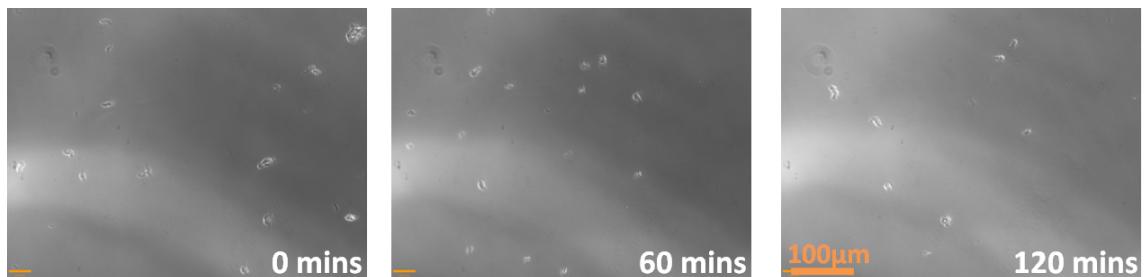
7. In conclusion, the trypanosomal Goodwin oscillator has been shown to lead to a periodic TetR₂ and eGFP output. Based on these findings, along with characterisation results, a decision was taken to go ahead with microscopy fluorescent analysis.

7.3 Characterisation of the phenotype of the stable Goodwin oscillator via live cell imaging and qualitative mathematical pattern analysis

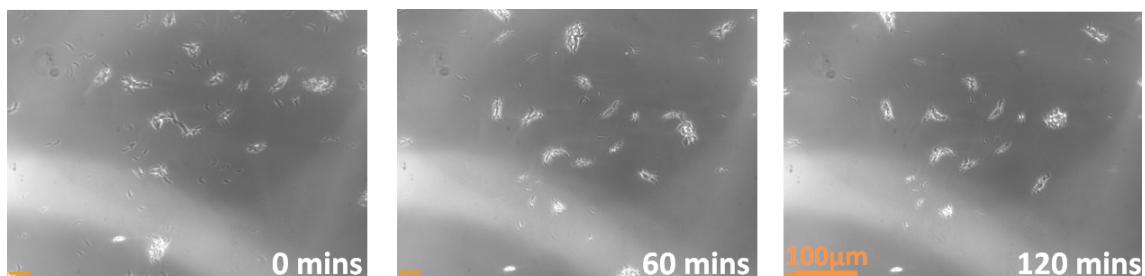
Following assessment of the main pStable components and mathematical screening analysis, live-cell fluorescent microscopy imaging of pStable-transfected BSF and PCF clones was carried out. Similarly to pUnstable-transfected *T. brucei*, this was done in order to quantify and analyse fluorescence pattern profiles of live trypanosome cells. Images of pStable transfectants were acquired at two minute intervals over a two-hour period using fluorescent microscopy with a 20X objective. Cells at a density of 5×10^7 cells/ml were embedded in 40% Cygel as per Section 5.7. Throughout, brightfield microscopy was used in tandem to fluorescent microscopy as a control. This was done to confirm that observed fluorescence is not being emitted by debris and secondly so as to facilitate cell tracking. Once completed, the images were collated, corrected for background fluorescence and analysed as a video in order to detect and quantify changes in fluorescence levels over time in individual cells. Quantification consisted of both automated detection by Volocity and manual tracking. The latter was necessary due to an element of cell motility which can lead to incorrect cell tracking by Volocity. Table 7.7 lists the strains which were imaged.

7.3.1 Phenotype of the stable Goodwin oscillator in live BSF *T. brucei*

Fluorescent microscopy imaging was carried out on BSF *T. brucei* bearing the stable Goodwin oscillator. No fluorescence was observed throughout, as confirmed by quantitation (results not shown). Throughout, brightfield microscopy confirmed that cells were viable and healthy. This does not disqualify pStable as an oscillator but raises the same challenges as seen in the live cell imaging of pUnstable-transfected cells and may be an indication of the limits of sensitivity of the



(a) Images of BSF *T. brucei* bearing pStable highlighting BSF cell motility over 120 minutes.



(b) Images of PCF *T. brucei* bearing pStable highlighting PCF cell motility over 120 minutes.

Figure 7.26: Images of: a) uninduced pStable-transfected BSF *T. brucei* and b) uninduced pStable-transfected PCF *T. brucei* highlighting cell motility over 120 minutes. Images are taken using a 20X objective and brightfield microscopy. Scale bar = 100 μ m.

imaging technology being used (Section 6.2.2).

Following this, a second imaging study was carried out on pStable-transfected BSF clones induced with tetracycline at the tc_{thres} level, based on the hypothesis of increasing eGFP protein numbers while still allowing for oscillations to occur and thus controlling for the probability that oscillations are present but too low to be detected, as discussed in Figure 6.7. The effect of tc_{thres} on the cell population could not be properly gauged due to cell motility. Figure 7.26a shows three stills taken using brightfield microscopy of the induced transfected BSF cell population at time 0, 1 hour and 2 hours, which serves to highlight how the landscape changes over time. This highlights the challenge in tracing live motile cells over multiple time-points. Thus, no conclusion was made regarding the phenotype of pStable.

7.3.2 Phenotype of the stable Goodwin oscillator in uninduced live PCF *T. brucei* cells

Following the imaging of pStable-transfected BSF clones, the same analysis was carried out on uninduced PCF *T. brucei* bearing the pStable network. No fluorescent signals were detected. However, Figure 7.26b shows stills of the PCF cell population taken throughout the two hour imaging procedure, which serve to highlight two main differences from its BSF counterpart. Firstly, cell density is higher, in spite of using the same cell preparation techniques. Secondly, cell motility has been greatly reduced such that a proportion of the cells could be traced over the two hour period.

7.3.3 Phenotype of the stable Goodwin oscillator in live PCF *T. brucei* cells induced with tc_{thres}

7.3.3.1 Live cell imaging of the stable Goodwin oscillator in PCF *T. brucei* induced with tc_{thres}

Since no fluorescence was observed for uninduced PCF clones, a similar imaging procedure was carried out using PCF *T. brucei* bearing the pStable network and induced with tetracycline at a concentration of tc_{thres} as per Section 6.2.2.2. This was done in order to control for the probability that changes in fluorescent levels were not being detected, as seen per Figure 6.7. Automated and manual tracking showed that the tetracycline successfully led to a proportion of the cell population expressing eGFP protein at detectable levels, in line with the hypothesis.

Following imaging, due to improved tractability conditions of PCF *T. brucei*, three cells were selected which: 1) did not clump with other cells for the duration of imaging, 2) did not swim out of the plane of vision for the duration of imaging and 3) could be tracked over the majority of the stills. These cells, labelled Cell A, B and C, were taken forward for background correction, tracking and quantification and are highlighted in Figure 7.27 which shows both fluorescent background corrected and brightfield stills from the two hour imaging procedure. Figure 7.28 shows close-up stills from fluorescent and brightfield microscopy of Cell A over 120 minutes, as extracted from Figure 7.27. A change in fluorescence levels over time was observed in these cells, as will be discussed next. The remainder of the cells did not satisfy the three criteria simultaneously.

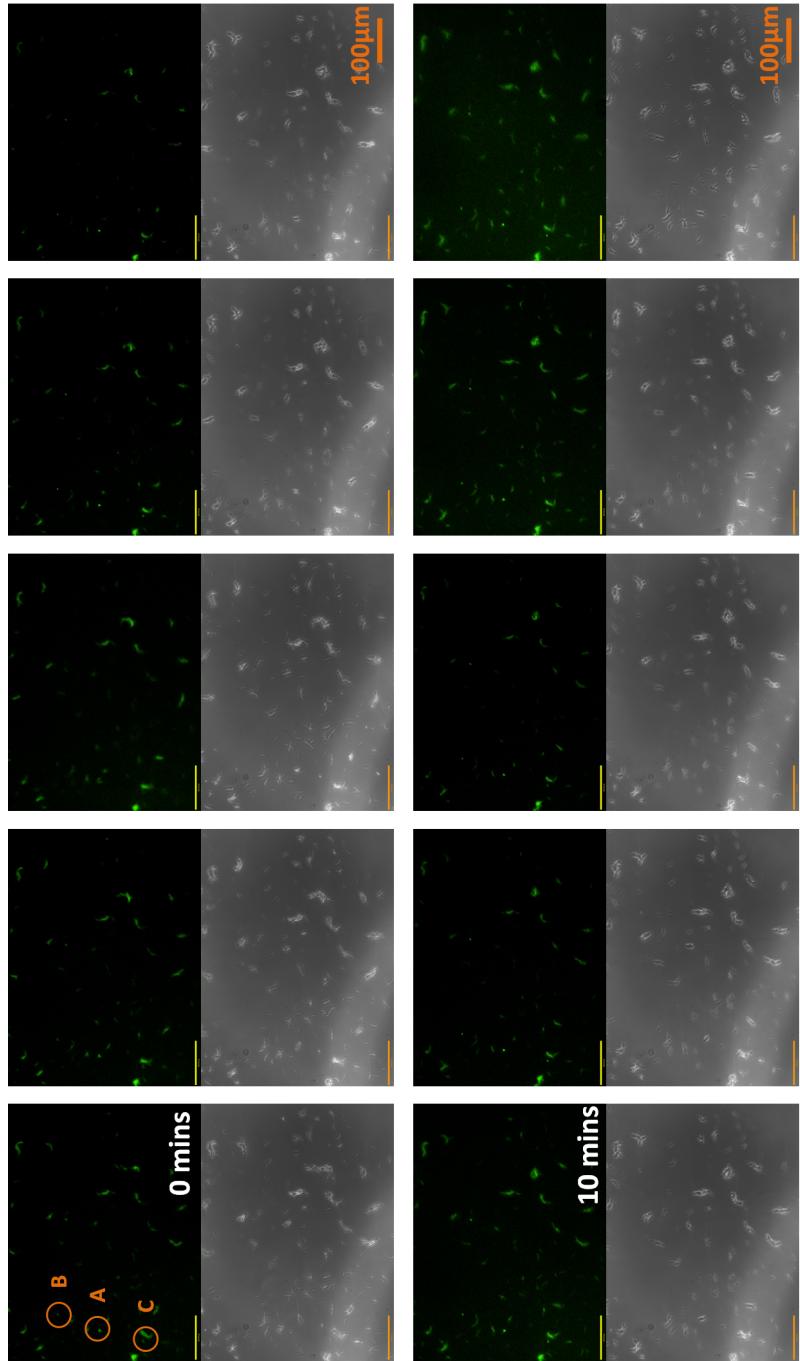


Figure 7.27: **Part 1/6** - Images of PCF *T. brucei* bearing the stable Goodwin oscillator and induced with tCires tetracycline, taken at two minute intervals and a 20X objective for a two hour period showing fluorescent background corrected (top level) and brightfield (bottom level) microscopy stills. The circled cells were taken forward for quantification analysis. Scale bar = $100\mu\text{m}$.

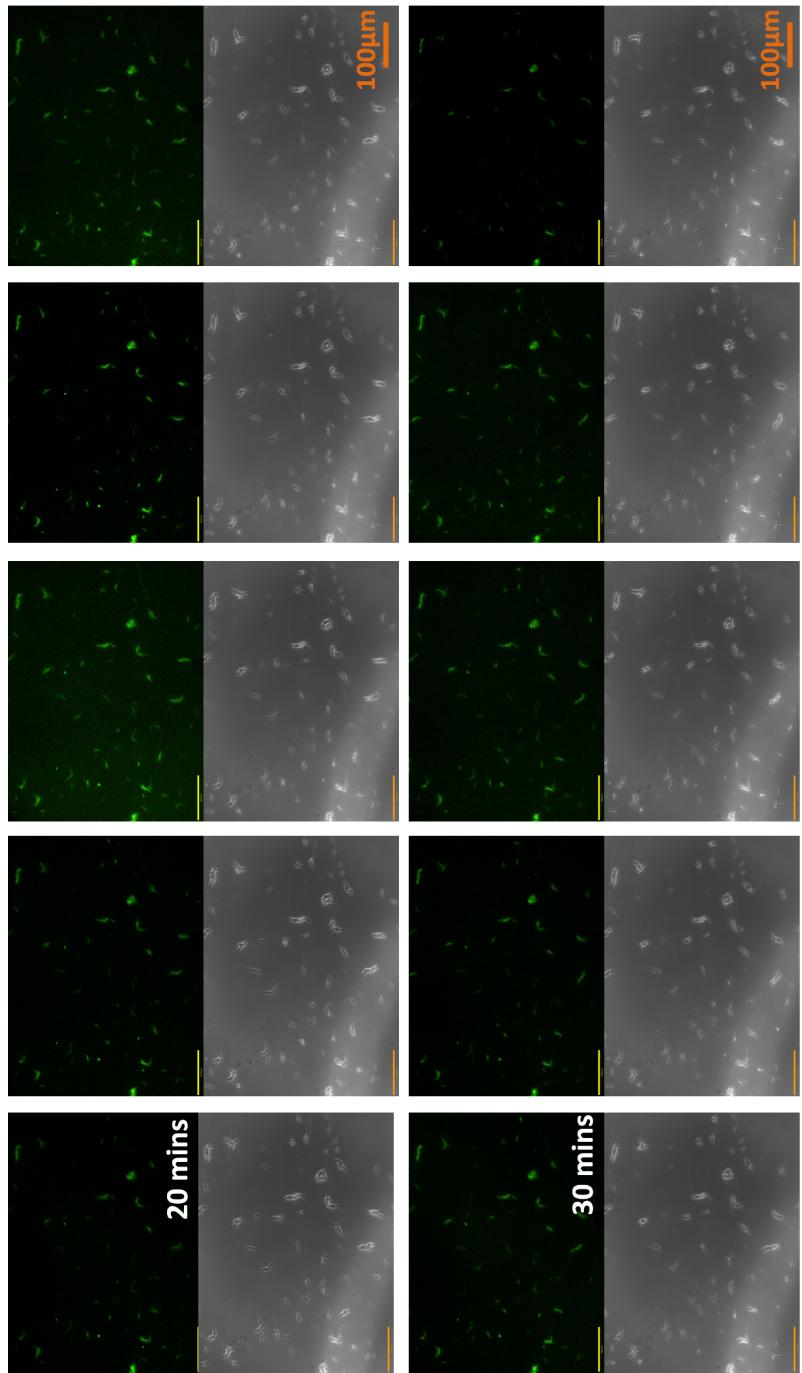


Figure 7.27: Part 2/6 - Images of PCF *T. brucei* bearing the stable Goodwin oscillator and induced with tCires tetracycline, taken at two minute intervals and a 20X objective for a two hour period showing fluorescent background corrected (top level) and brightfield (bottom level) microscopy stills. The circled cells were taken forward for quantification analysis. Scale bar = $100\mu\text{m}$.

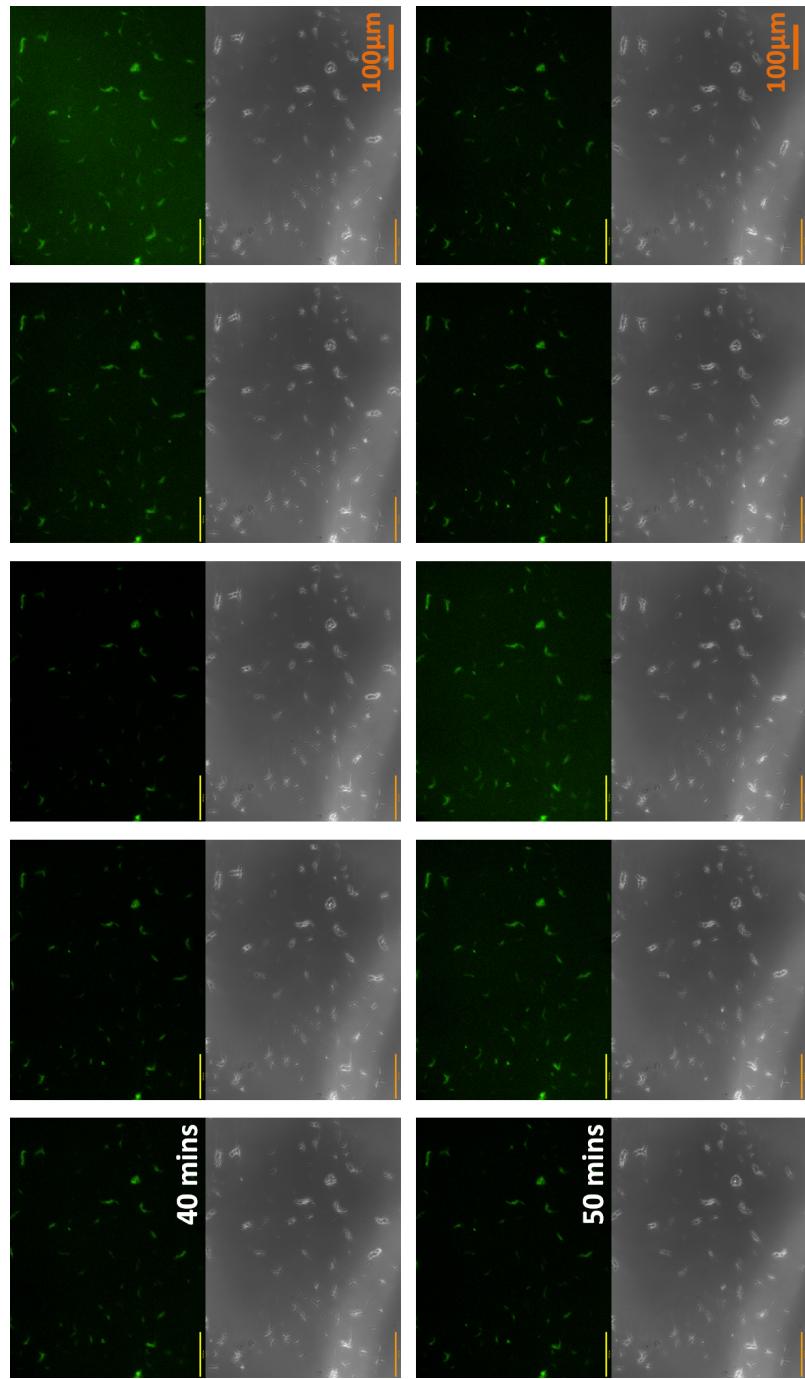


Figure 7.27: **Part 3/6** - Images of PCF *T. brucei* bearing the stable Goodwin oscillator and induced with tC_{10res} tetracycline, taken at two minute intervals and a 20X objective for a two hour period showing fluorescent background corrected (top level) and brightfield (bottom level) microscopy stills. The circled cells were taken forward for quantification analysis. Scale bar = 100 μ m.

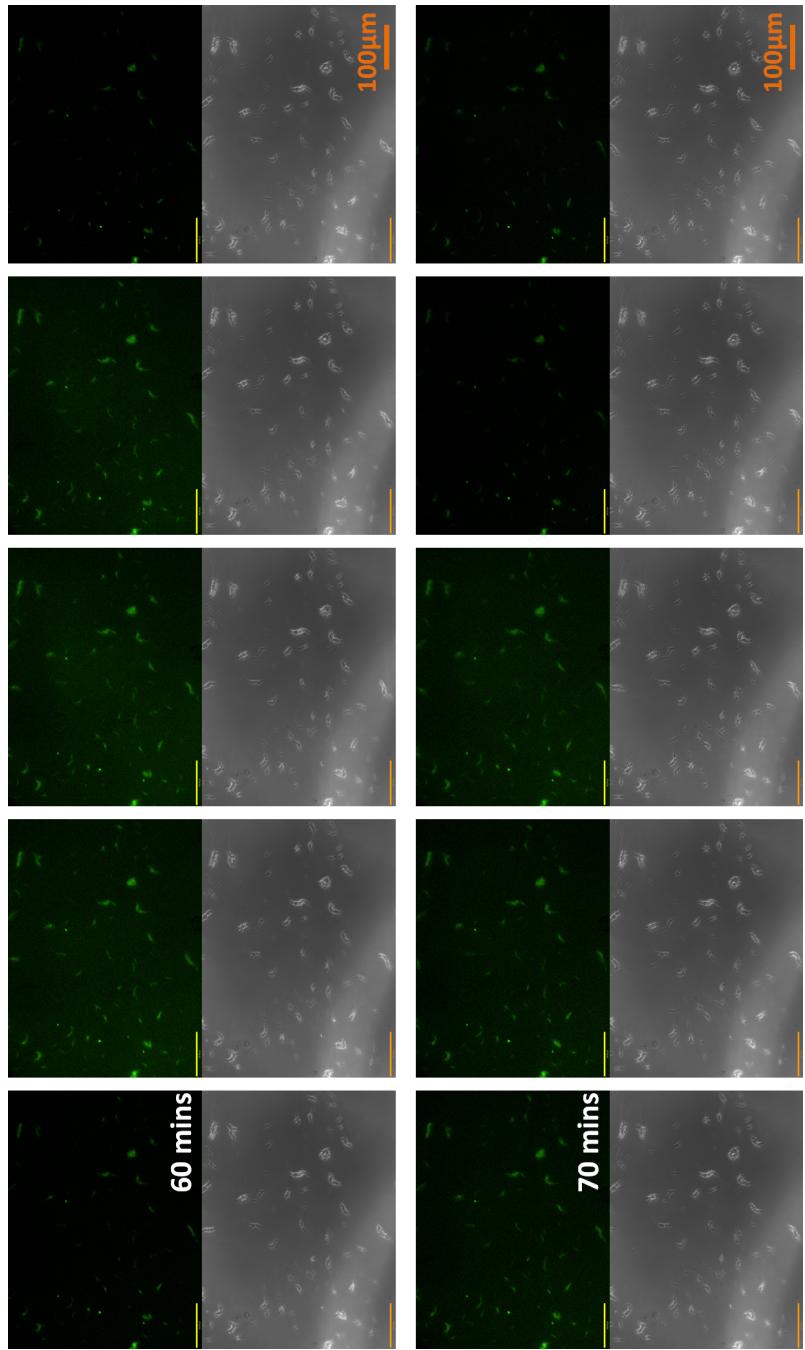


Figure 7.27: **Part 4/6** - Images of PCF *T. brucei* bearing the stable Goodwin oscillator and induced with tCires tetracycline, taken at two minute intervals and a 20X objective for a two hour period showing fluorescent background corrected (top level) and brightfield (bottom level) microscopy stills. The circled cells were taken forward for quantification analysis. Scale bar = $100\mu\text{m}$.

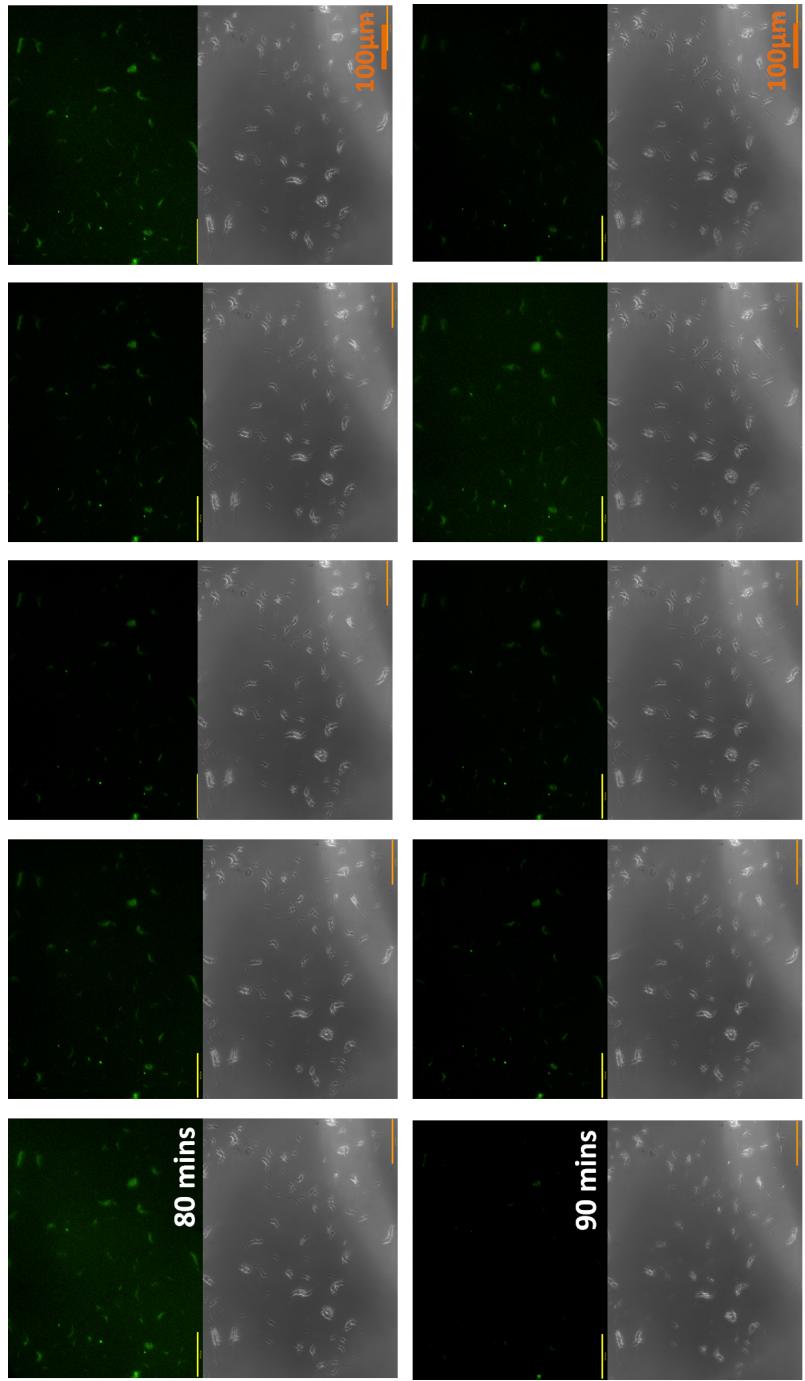


Figure 7.27: Part 5/6 - Images of PCF *T. brucei* bearing the stable Goodwin oscillator and induced with tCires tetracycline, taken at two minute intervals and a 20X objective for a two hour period showing fluorescent background corrected (top level) and brightfield (bottom level) microscopy stills. The circled cells were taken forward for quantification analysis. Scale bar = $100\mu\text{m}$.

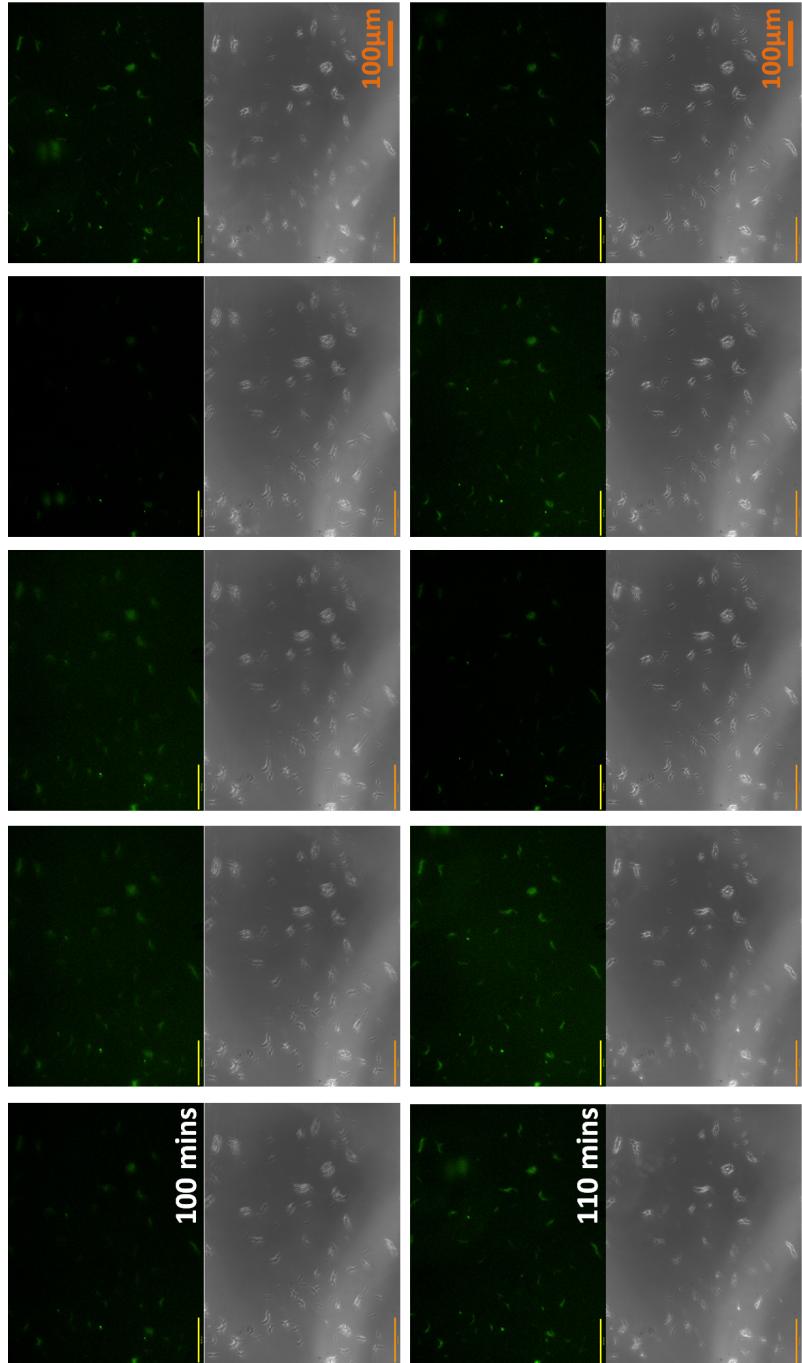


Figure 7.27: Part 6/6 - Images of PCF *T. brucei* bearing the stable Goodwin oscillator and induced with tCires tetracycline, taken at two minute intervals and a 20X objective for a two hour period showing fluorescent background corrected (top level) and brightfield (bottom level) microscopy stills. The circled cells were taken forward for quantification analysis. Scale bar = $100\mu\text{m}$.

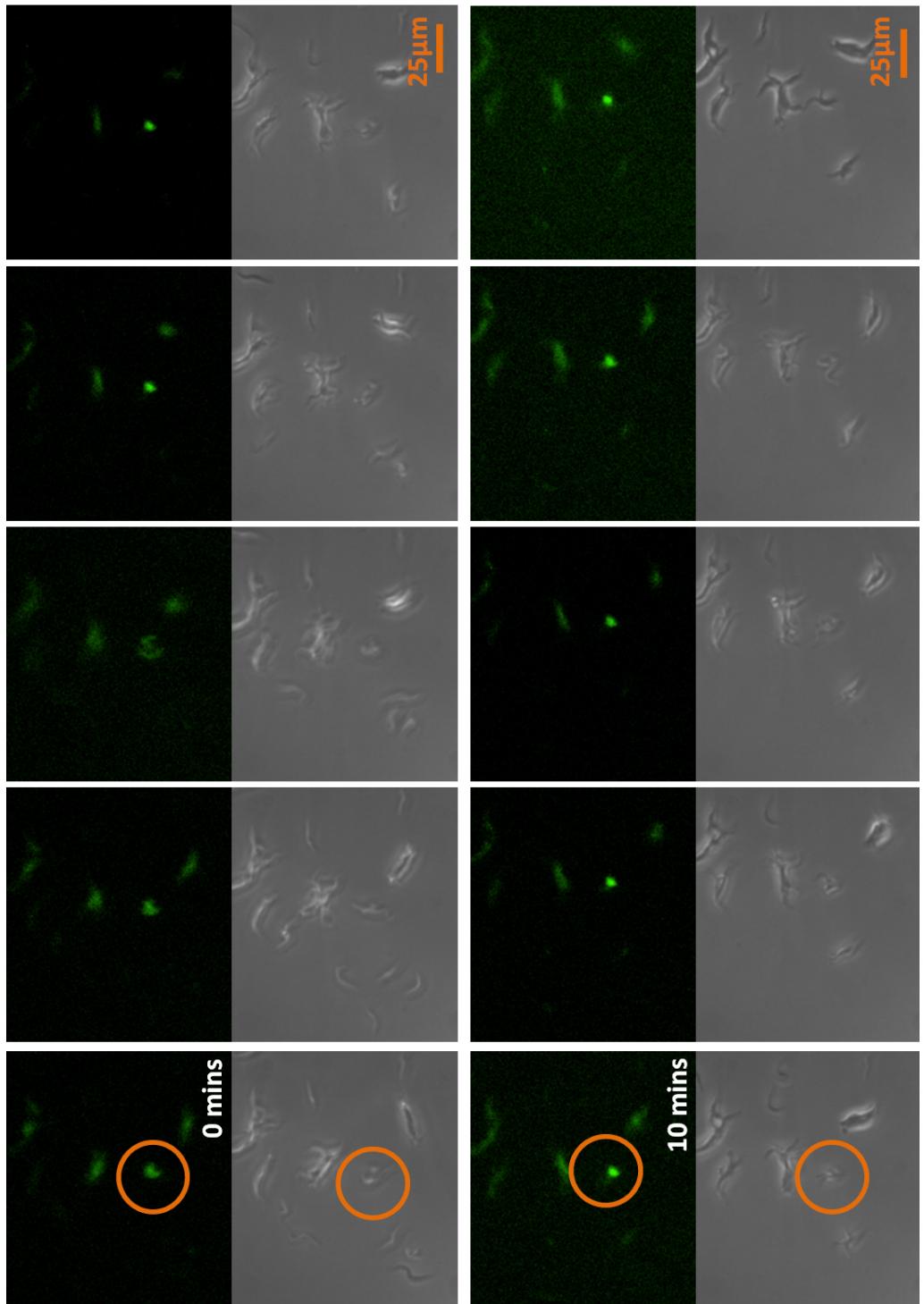


Figure 7.28: **Cell A, Part 1/6** - Images showing close-up of Cell A selected from the t_{Cthres} -induced PCF *T. brucei* bearing the stable Goodwin oscillator. Top level of each row shows images from fluorescent background-corrected microscopy while bottom level shows brightfield microscopy images. Images were taken at two minute intervals over a two hour period. Scale bar = $25\mu\text{m}$.

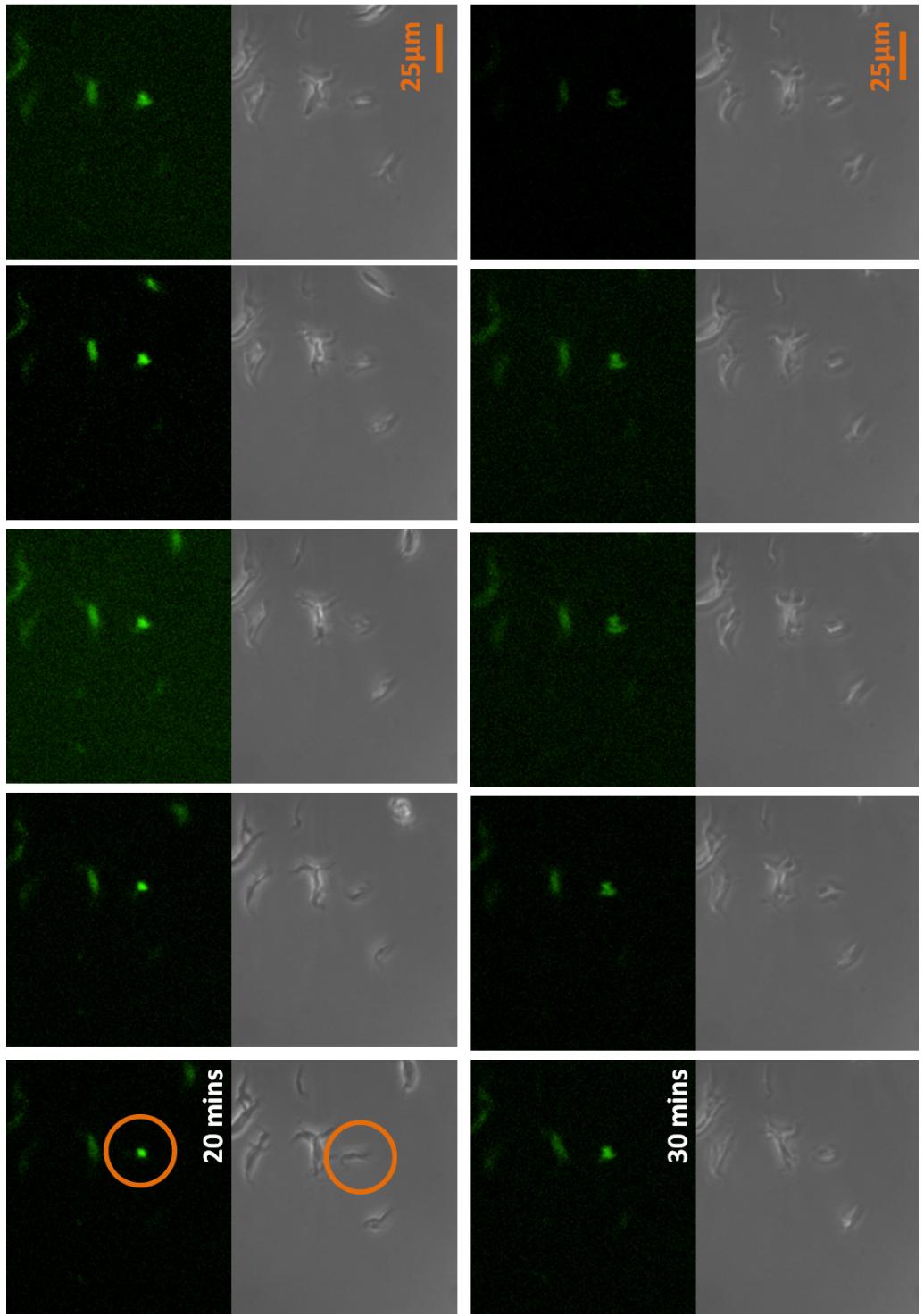


Figure 7.28: **Cell A, Part 2/6** - Images showing close-up of Cell A selected from the t_{Cthres} -induced PCF *T. brucei* bearing the stable Goodwin oscillator. Top level of each row shows images from fluorescent background-corrected microscopy while bottom level shows brightfield microscopy images. Images were taken at two minute intervals over a two hour period. Scale bar = $25\mu\text{m}$.

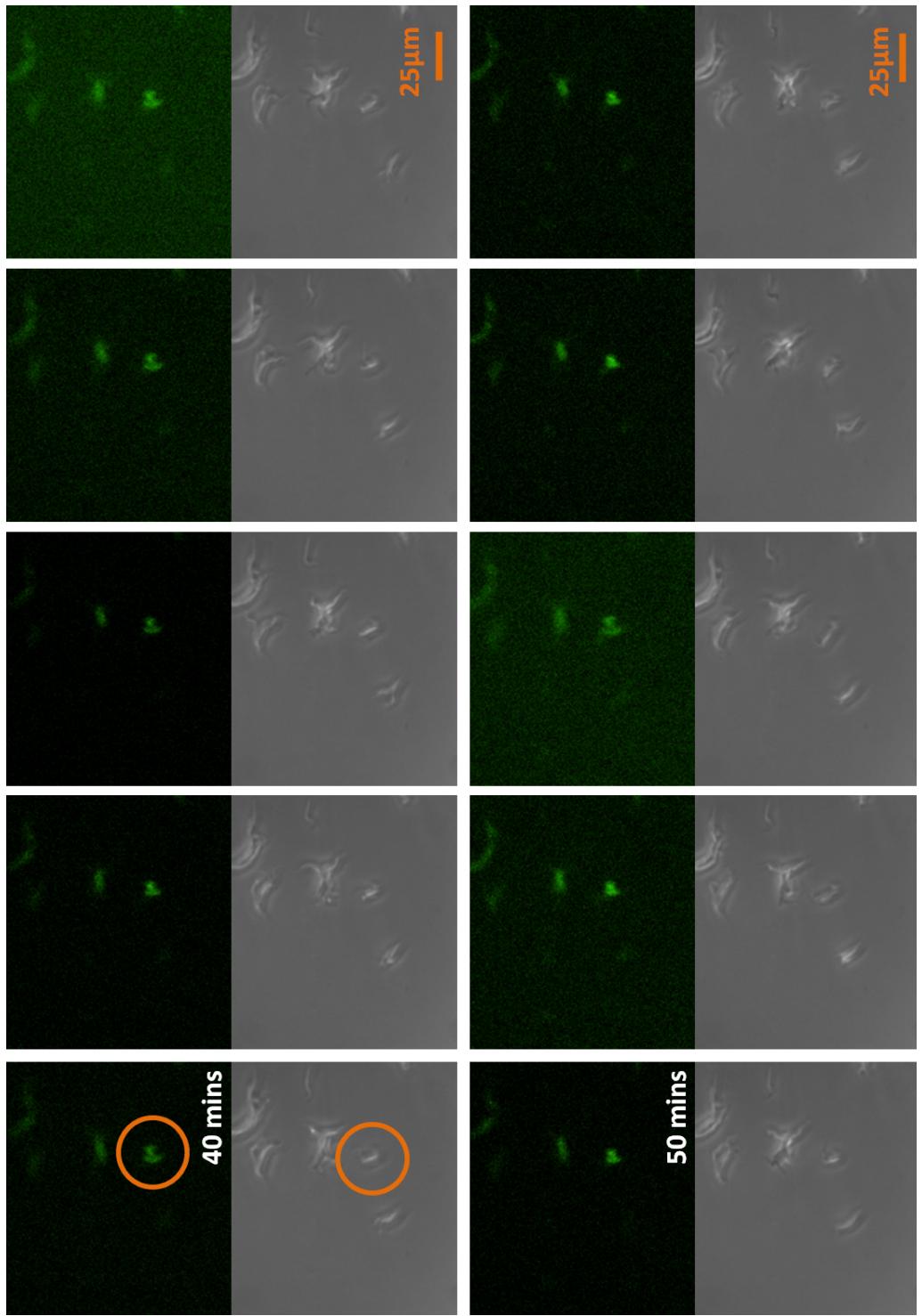


Figure 7.28: **Cell A, Part 3/6** - Images showing close-up of Cell A selected from the t_{Cthres} -induced PCF *T. brucei* bearing the stable Goodwin oscillator. Top level of each row shows images from fluorescent background-corrected microscopy while bottom level shows brightfield microscopy images. Images were taken at two minute intervals over a two hour period. Scale bar = $25\mu\text{m}$.

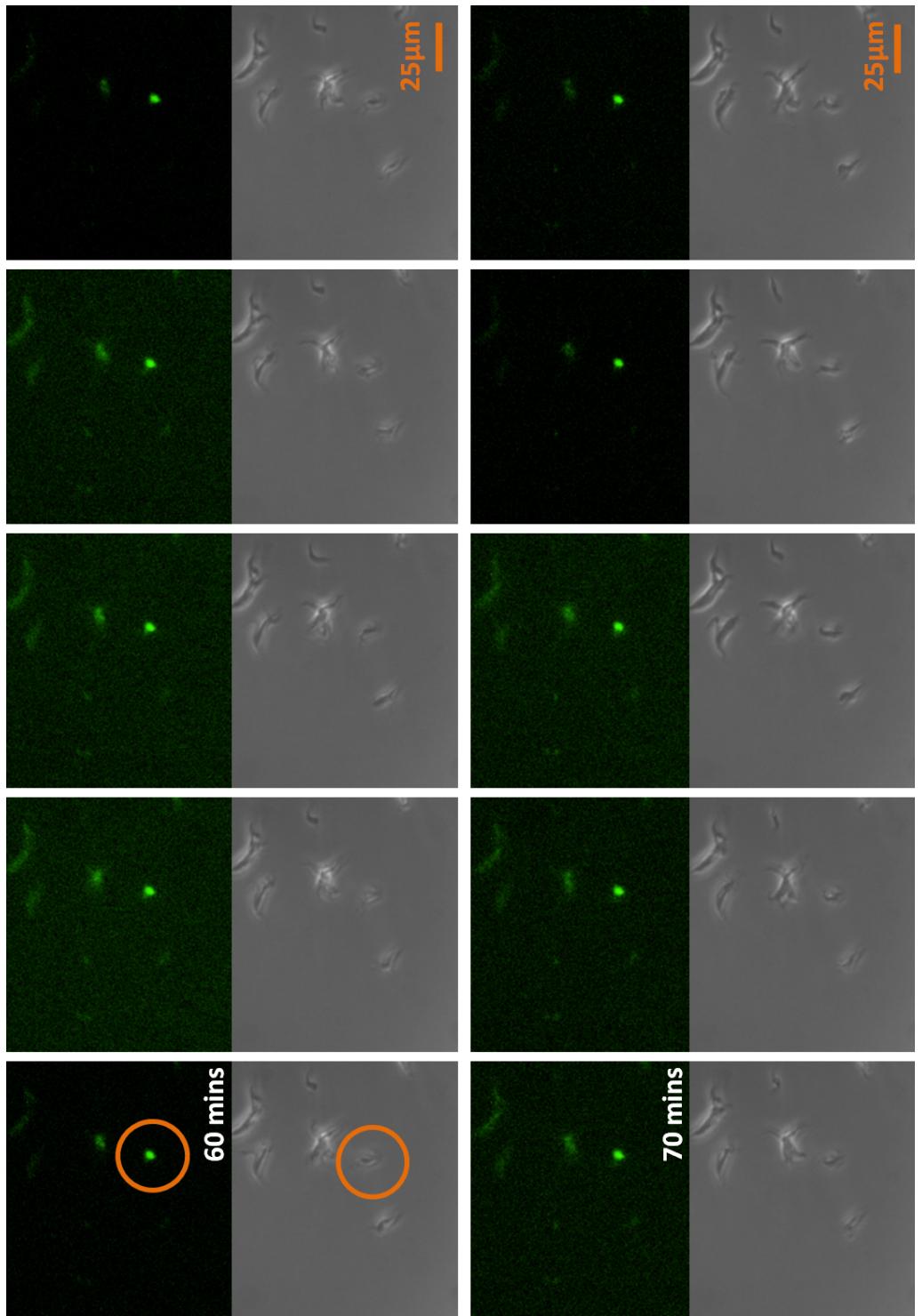


Figure 7.28: **Cell A, Part 4/6** - Images showing close-up of Cell A selected from the t_{Cthres} -induced PCF *T. brucei* bearing the stable Goodwin oscillator. Top level of each row shows images from fluorescent background-corrected microscopy while bottom level shows brightfield microscopy images. Images were taken at two minute intervals over a two hour period. Scale bar = $25 \mu\text{m}$.

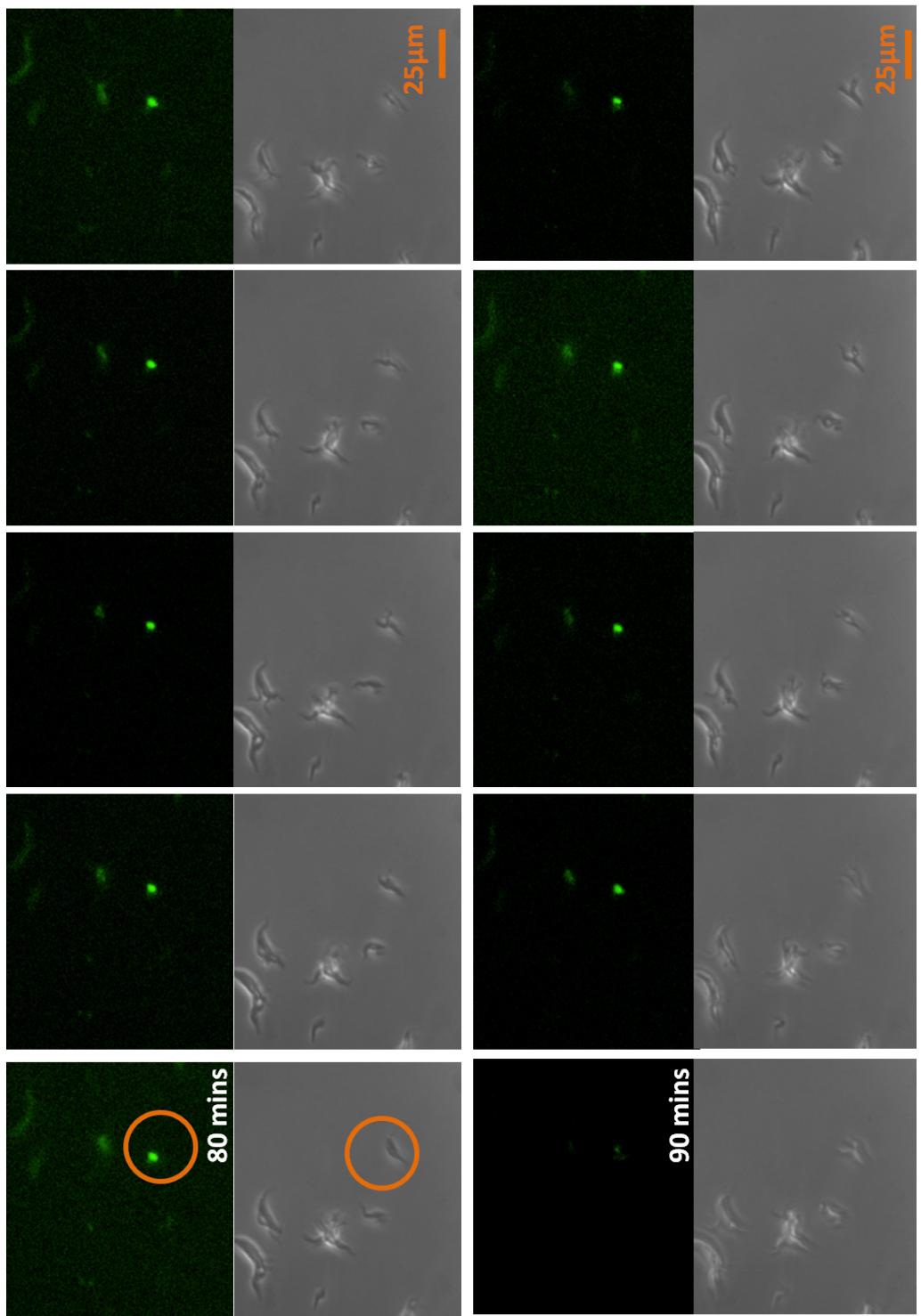


Figure 7.28: **Cell A, Part 5/6** - Images showing close-up of Cell A selected from the tc_{thres} -induced PCF *T. brucei* bearing the stable Goodwin oscillator. Top level of each row shows images from fluorescent background-corrected microscopy while bottom level shows brightfield microscopy images. Images were taken at two minute intervals over a two hour period. Scale bar = $25\mu\text{m}$.

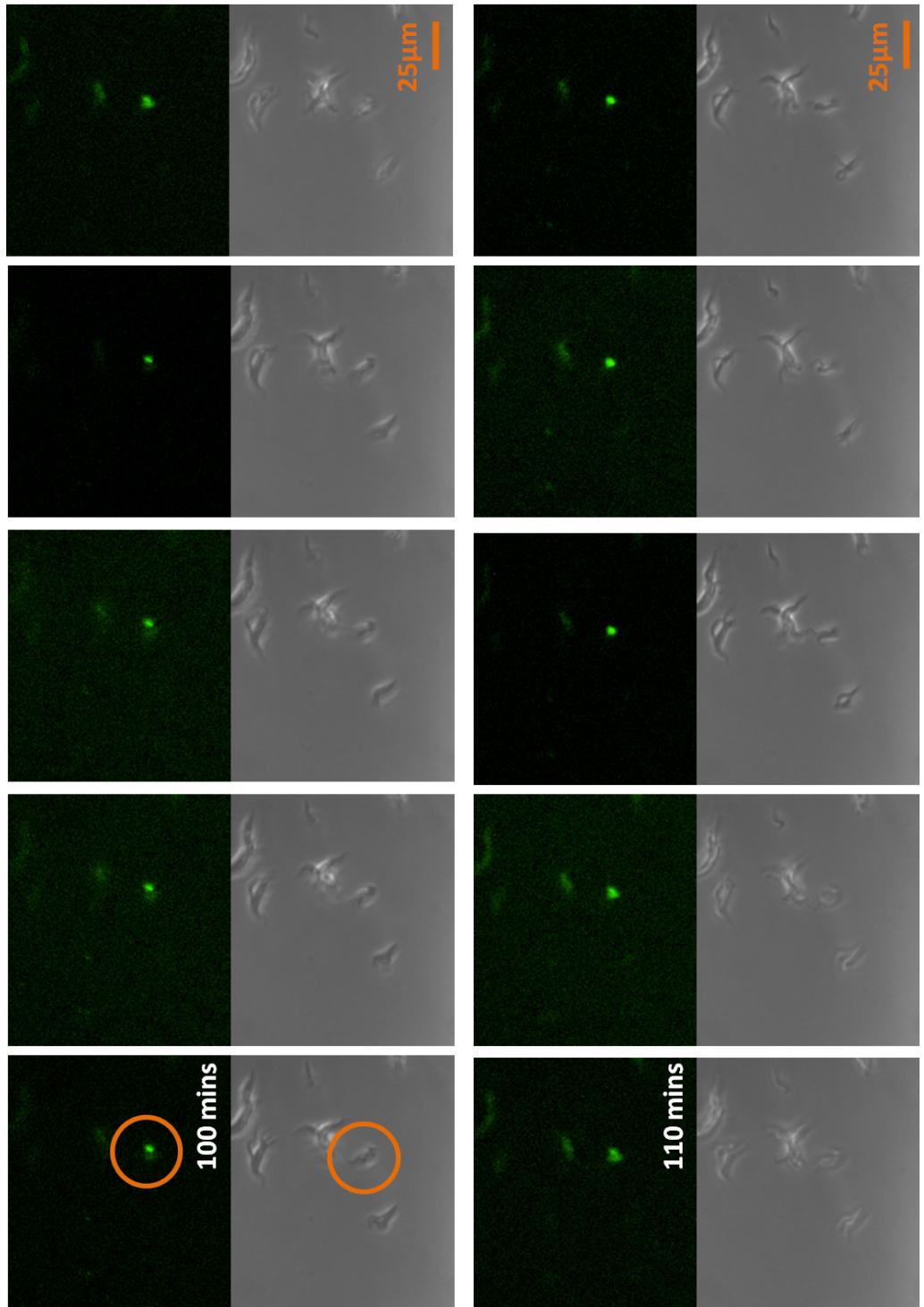


Figure 7.28: **Cell A, Part 6/6** - Images showing close-up of Cell A selected from the tc_{thres} -induced PCF *T. brucei* bearing the stable Goodwin oscillator. Top level of each row shows images from fluorescent background-corrected microscopy while bottom level shows brightfield microscopy images. Images were taken at two minute intervals over a two hour period. Scale bar = $25\mu\text{m}$.

7.3.3.2 Quantification and analysis of fluorescence oscillatory trajectories at the single cell level in *T. brucei* cells bearing pStable and induced with tc_{thres}

The time series of the quantified fluorescence in each of the three cells is plotted in Figure 7.29. A fluctuating trend, in line with oscillations, can be seen in each of the profiles. The standard deviation for each cell was 22%-35% of the mean signal, indicating a trend in the variations which goes beyond the stochasticity resulting from focusing variability and stochastic intracellular processes, quantified at c. 17% of the mean signal in a plasmid expressing GFP constitutively (Section 6.2.1). It needs to be taken into account that the latter calculation was based on readings over a 30 minute period, while oscillations within pStable were taken over two hours. Therefore, fluorescence signals of the constitutive plasmid should be measured over an equivalent two hour time-span prior to conclusions being made. This will be enabled by the development of an imaging device which allows for better control of the organisms.

Prior to further data manipulation, missing data points were imputed using a window of size 3 i.e. the missing data point was estimated by averaging the value of the three previous and three subsequent data points. There was one missing data point per cell, attributed to cell undulations and loss of microscope focus. Following this, the data was noise-filtered so as to extract the underlying trend from the stochastic environment. Two different smoothing techniques were effected: an 8-point moving average filter (Oppenheim et al., 1999) and a cubic spline interpolation. This allows for the visualisation of the data as a smooth, continuous curve.

There were 60 data points in total, representing the fluorescence signal of the cell at two minute intervals over a two hour period. In the moving average filter using a window of size eight, the average value of each data point was averaged out over the time point t_i of interest and eight time points prior and subsequent to it $t_{i-8}, \dots, t_{i-1}, t_i, t_{i+1}, \dots, t_{i+8}$ given that $i - j > 0$ and $i + j < 61$ for $j = 1, \dots, 8$. Otherwise, the time-point out of bounds was ignored. Moving average filters have been used to smooth oscillations and eliminate the effects of noise in synthetic biology (Tigges et al., 2009), biochemistry (Jiang et al., 2013), engineering (Johnson & Principe, 2005), astronomy (Bloomfield, 2004) and neurology (Azami et al., 2011). Although other filtering algorithms are available (as seen in, for example, Stricker et al. (2008) and Tigges et al. (2010)) this technique was selected because it allows for a tailored filtering process via the user-defined window size.

In the cubic spline interpolation, polynomials were fitted to sections of the noise-filtered data such that connection points are smooth, to form a piecewise function known as a cubic spline. Splines have been used in biochemistry and synthetic biology (Yang et al., 2013b; Chen et al., 2009; Morrissey et al., 2011) to fit piecewise functions to irregular data.

The imputed, noise-filtered and smoothed functions for the three cells, as seen in Figure 7.29, show a low-amplitude sinusoidal trajectory, having a period of 50 minutes and a small linear growth trend. It is not yet clear how this compares to the observed half-life of destabilised eGFP (c. 3 hours). This also highlights the need to deduce the half-life of TetR, since it will help inform whether the fluctuations can be traced back to changes in TetR protein levels or otherwise.

The cells appear to be oscillating in synchrony. The low proportion of oscillating cells, in comparison to the cell population being imaged does not automatically pose a challenge to the nature of the GRN. For example, only 12% of the cells having the oscillator developed by Tigges et al.

(2010) were observed to fluctuate. Further investigative work is expected to inform whether the observed oscillations are recurrent or not, whether the synchrony holds over generations and more importantly, whether the fluctuations are TetR-generated or stochastic.

Moreover, the oscillatory characteristics (low-amplitude, 50 minute period) need to be fitted against a mathematical model in order to extract the parameter profile. The preliminary stochastic simulations carried out yielded a wide array of fluctuations which strengthened the case for carrying out live cell imaging but had little information to extract otherwise due to the variability. However, the mathematical model of the trypanosomal Goodwin oscillator can be used to explain why oscillations or changes in fluorescence patterns were not observed in uninduced *T. brucei* bearing the stable or unstable oscillator and in induced BSF *T. brucei* bearing the stable Goodwin oscillator; based on the consistently fluctuation output, it is possible that the oscillations were present but were too low for detection by the available technology.

7.3.3.3 Phase space reconstruction of the stable Goodwin oscillator fluorescent profile in the *T. brucei* host

In order to visualise recurring patterns and establish whether the oscillations observed held under mathematical scrutiny, following data quantification and smoothing, qualitative analysis of the non-linear system was carried out by reconstructing the phase space from the smoothed cell trajectories. The phase space represents all the possible states(values) of a selected variable within the network. In this case, the phase space of the fluorescence signal was reconstructed. The states attained by the variable were traced on the phase space. This trajectory then allowed for underlying structures to be elicited. The phase space was reconstructed as per Taken's time-delay embedding method (Takens, 1981). Phase space analysis has been widely used in the study of genetic networks for similar purposes (Gao et al., 2010b; Li et al., 2012; Fung et al., 2005).

Since the oscillatory time series profile for each cell showed a growing average linear trend (Figure 7.29), the first step prior to plotting the phase space of each cell was to de-trend the data. Initially, the mean value of the trajectory was subtracted from each data point of the smoothed dataset. Secondly, a linear growth line (the trend) was fitted to this smoothed dataset with mean 0. The trend was then removed by subtracting the linear growth to get de-trended oscillations. Figure 7.30 shows the smoothed data oscillations and the fitted growth trend for each of Cells A, B and C, along with the de-trended oscillation and the de-trended mean, which has value 0. The linear growth trend for the eGFP signal is attributed to an accumulation of eGFP due to its relatively low degradation rate which does not allow for all eGFP to be degraded prior to the next wave of protein expression. This was also reflected to an extent in mathematical simulations.

The phase space was then generated using the de-trended dataset seen in Figure 7.30 and using a delay of 8, which was chosen empirically. Although the delay could be estimated using the False Nearest Neighbours' algorithm (Kennel et al., 1992) or the Mutual Information scheme (Fraser & Swinney, 1986) the phase space generated using this data was suboptimal (not shown). The generated phase planes (Figure 7.31) are indicative of a periodic limit cycle i.e. a closed trajectory, although it does not clarify whether the observed fluorescent dynamics are oscillations resulting from changes in eGFP quantity and hence, the interplay between expression and inhibitory regulation. Figure 7.31d, showing the reconstructed phase space with the three trajectories su-

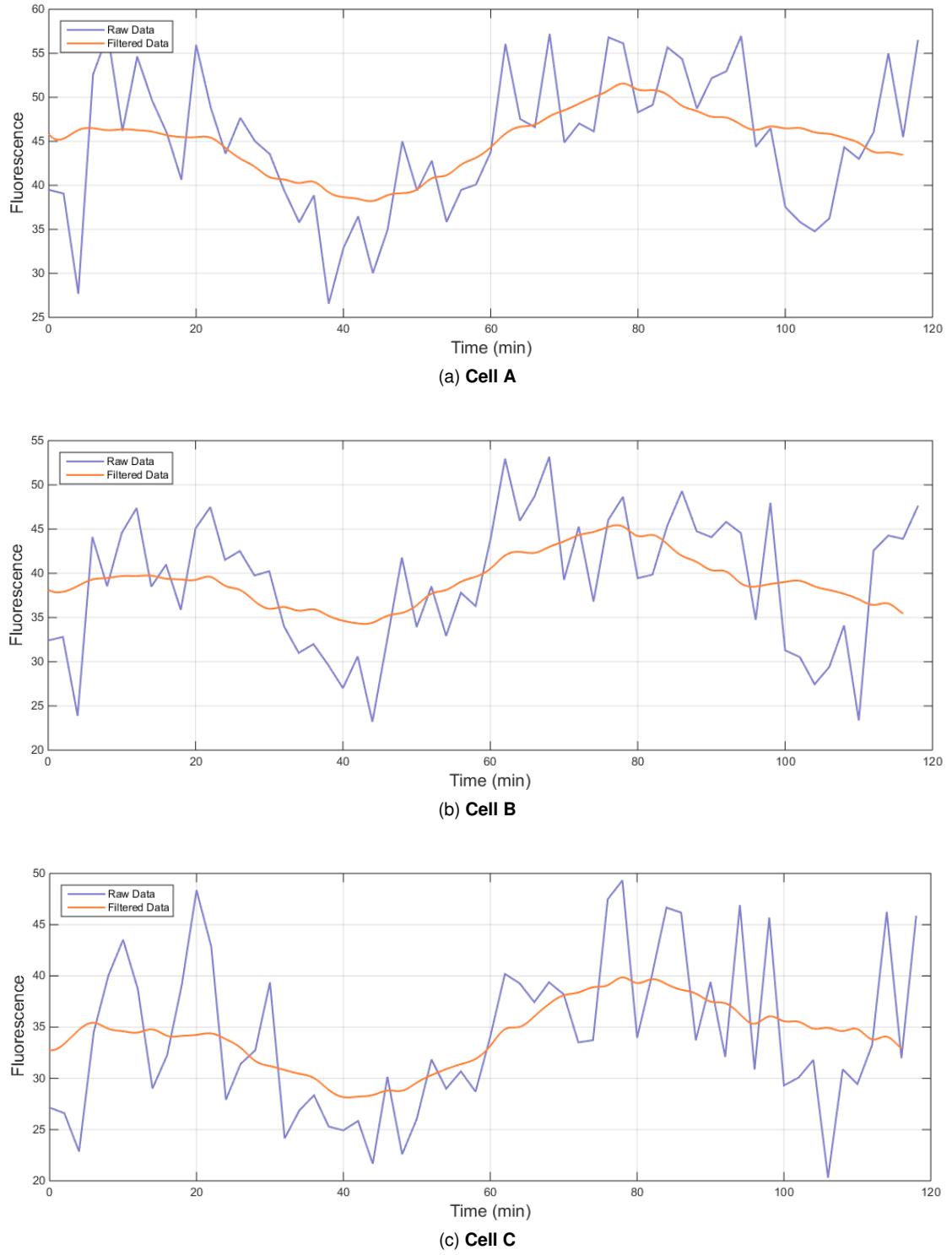


Figure 7.29: Quantified fluorescent trajectories in tc_{thres} induced PCF cells bearing the pStable network: Plot showing raw quantified fluorescence data (in blue) and imputed, noise-filtered and smoothed fluorescence (in orange) time series for three PCF tc_{thres} -induced cells bearing the pStable network, which did not clump or move out of the plane of vision during the two hour experiment and had detectable fluorescence levels. The axis show Fluorescence vs Time in minutes. Fluorescence intensity has no units.

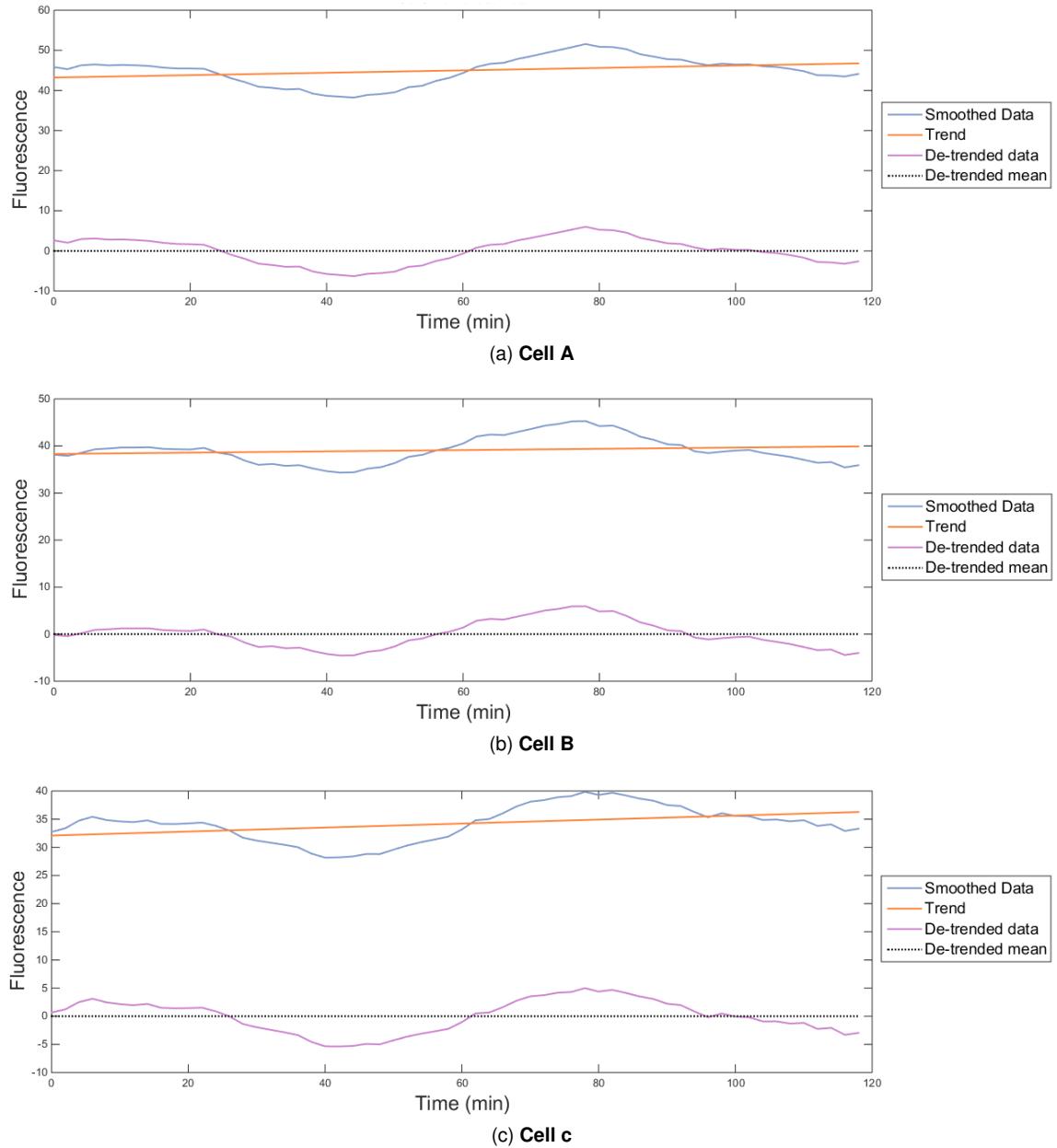


Figure 7.30: Smoothed and De-trended data: Diagrams showing smoothed de-trended time series profile (purple) for the fluorescent intensity in a) Cell A, b) Cell B and c) Cell C. The axis shows Fluorescence (no units) vs Time in minutes. Each plot also shows the imputed, noise-filtered, smoothed data which was then detrended (blue), the linear trend which was removed (orange) and the detrended mean which has value zero (black).

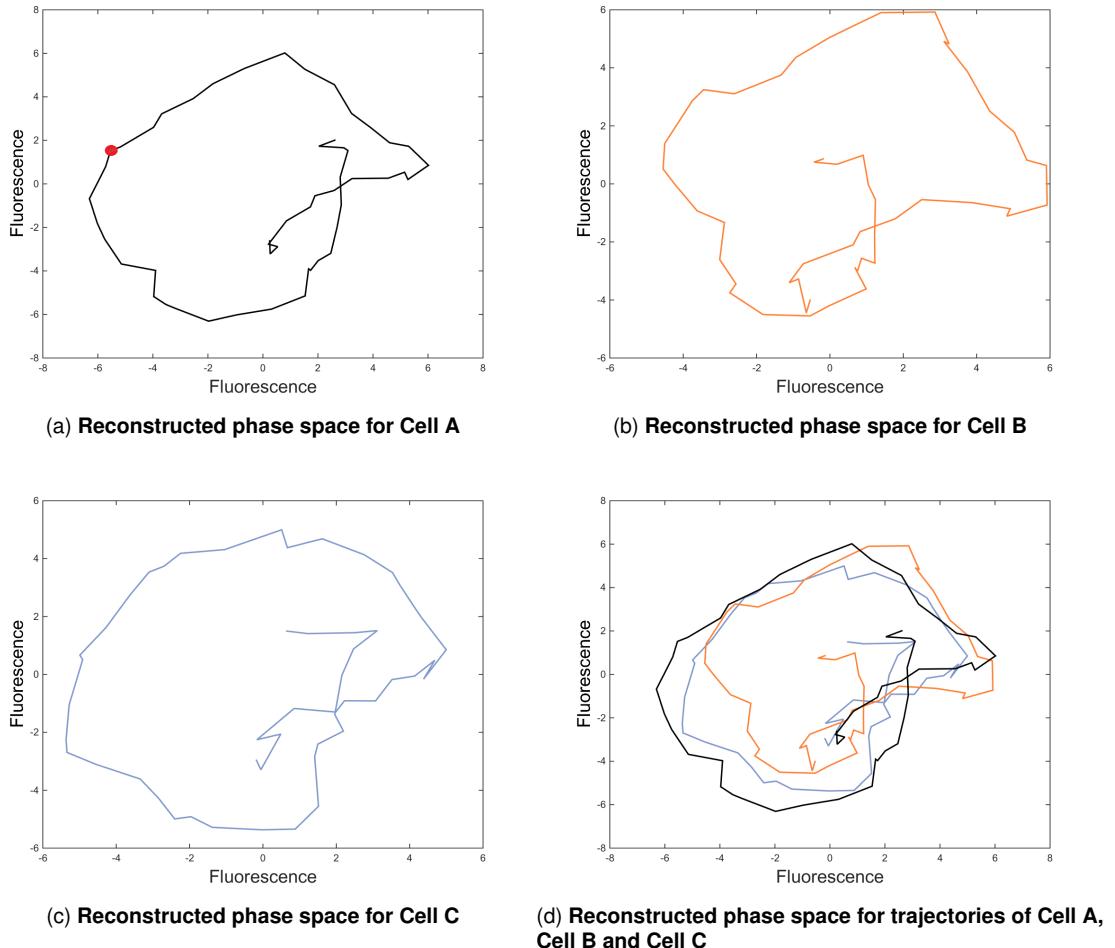


Figure 7.31: Reconstructed phase space of de-trended data: Diagrams showing reconstructed phase space of smoothed de-trended fluorescence signal trajectory of a) Cell A, b) Cell B and c) Cell C. d) shows the reconstructed phase space using all three trajectories. A delay of 8 was used. The axis shows Fluorescence (no units) vs Fluorescence (no units).

perimposed on each other, indicates that the trajectories are similar to each other and that there is consistency among the oscillations. Outliers are attributed to the data points on the edge of the data set (the first or last data points), which can not be manipulated the same as the remaining points. Further investigations, using larger data sets and a second generation mathematical model, are expected to complement these findings.

7.3.4 Evaluation of the stable oscillator phenotype following live cell microscopy imaging and qualitative mathematical analysis

Following characterisation of the phenotype of the stable Goodwin oscillator via live cell microscopy imaging, data quantification and qualitative mathematical analysis, the following results and observations were made:

1. PCF cells bearing the stable Goodwin oscillator allow for more straightforward analysis than their BSF counterparts, in that microscopy images showed decreased cell motility and denser cell populations. Thus, quantification of single cell fluorescent trajectories and the

screening of populations of cells was facilitated.

2. The preliminary mathematical model which was developed to screen biochemical parameter rate combinations which led to an oscillatory output can be used to explain why no oscillations were observed in BSF *T. brucei* bearing the stable Goodwin oscillator and their uninduced PCF counterparts. Simulations repeatedly showed an oscillatory output with a low amplitude and low maximum coordinate, indicating that oscillations may be occurring but are not being detected by the available microscopy technology.
3. A varying fluorescent pattern was observed over a two hour imaging segment in a small proportion of PCF cells bearing the stable Goodwin oscillator and induced with tetracycline at a threshold concentration of $t_{c\text{thres}}=10\text{fg/ml}$. As hypothesised, this acted as a tuning measure, increasing the overall quantity of eGFP while still allowing for oscillations to take place and controlling for the probability that oscillations were present but not being detected by the available technology.
4. The fluorescence signal profile of three induced PCF *T. brucei* bearing the pStable network which retained viability and tractability over the two hour imaging procedure was quantified using a combination of thresholding, edge detection and watershed segmentation. Following this, results were confirmed visually. The time series trajectory was smoothed and filtered using a moving-average filter and spline fitting algorithm to dampen the effect of noise and allow for the dominant trend to be extracted. This revealed a low-amplitude oscillatory pattern in the fluorescence trajectories of all three cells.
5. In order to better assess whether the observed and quantified oscillatory trajectories hold under qualitative mathematical analysis, the phase space for the fluorescence intensity of each oscillation was reconstructed. Results showed the presence of a limit cycle within each phase space with outlier edges which correspond to the first and last smoothed data points which are generated from smaller data points than the remainder.
6. Following characterisation and mathematical analysis, results yielded data consistent with oscillations, the nature of which is unclear. Further investigations using larger data sets, deterministic models and control measurements should be carried out in order to yield conclusive data in line with this or otherwise.

7.4 Discussion

This chapter presented results following the characterisation of the stable Goodwin oscillator phenotype in *T. brucei* cells via protein analysis, microscopy imaging and quantitative and qualitative mathematical modelling. The following conclusions are made:

1. In line with synthetic biology principles, mathematical and wet lab analysis was used to inform the other and build on the available knowledge in a cyclical pattern.
2. Characterised the novel pStable synthetic GRN within a *T. brucei* chassis. This is the first synthetic plasmid, alongside pUnstable, to be characterised within a trypanosomal platform.
3. Assessed function of the pStable network components. Promoters, operator and antibiotic resistance mechanism were found to be functional. Moreover, it was confirmed that TetR and eGFP protein was expressed and that expression was regulated via TetR inhibition, the effect of which could be reversed via tetracycline. When added at a concentration of $1\mu\text{g/ml}$, tetracycline induced expression for over 24 hours, after which inhibition was resumed.

4. Deduced that when induced with tetracycline at a concentration of 10fg/ml, pStable exhibited partial fluorescence in that a proportion of cells were observed to fluoresce while the remainder did not. This was used throughout imaging in order to increase the probability of detecting changing fluorescence patterns.
5. Established that the degradation rate of Ubiquitin tagged eGFP within a *T. brucei* platform is 0.004min^{-1} . From the available data, this appears to be the first such exercise to determine reaction rates of GRN components within a *T. brucei* chassis.
6. Deduced that growth rates of *T. brucei* bearing the pStable network are not statistically different from wildtype *T. brucei*. This is similar to results for the pUnstable network.
7. Developed a mathematical model for the trypanosomal Goodwin oscillation based on the expression dynamics and interactions of the network components via biochemical rates. This is one of the first applications of a mathematical model to a synthetic GRN as part of an analytical package to investigate GRNs within trypanosomes. Experimental and literature data was used to inform parameter rates within the model. In order to screen the parameter space for combinations which lead to oscillatory protein expression dynamics, Gillespie simulations were run using different parameter combinations to generate time series profiles for the regulatory and reporter proteins, TetR and eGFP, and their associated components. Results indicated that low-level sinusoidal oscillations occurred over a wide variety of frequencies, irrespective of the parameter values selected. This brings into question whether the oscillations are stochastic or genuine.
8. Elicited information on the influence of different parameter combinations on the oscillatory output. Analysis of parameter rates with respect to the maximum coordinate of TetR₂ oscillations showed that low protein synthesis coupled with high degradation rates led to low-level maximum coordinates in oscillations, while high protein synthesis rates coupled with low degradation rates led to oscillations which had a higher maximum coordinate. Analysis of eGFP simulations showed that low protein synthesis, combined with a low degradation rate for the TetR components and a high degradation of mRNA favoured a low level maximum coordinate in oscillations. The reverse is true for an eGFP oscillatory output with higher maximum coordinates. Therefore, oscillations appear to be dependant on the protein turnover rate. However, the relation is noisy and again brings into questions whether this is reflective of genuine or stochastic fluctuations.
9. Characterised the phenotype of the pStable network in live *T. brucei* via microscopy imaging and quantitation. PCF cells bearing the stable Goodwin oscillator and induced with tc_{thres} tetracycline had detectable fluorescence levels. Three cells which retained healthy viability and remained in focus were tracked and their fluorescence levels found to correspond to an oscillatory pattern having a period of c. 50 minutes and a varying amplitude with a linear growth trend. This was corroborated by mathematical analysis via phase space reconstruction.
10. Data is consistent with the pStable plasmid functioning as a stable Goodwin oscillator within a *T. brucei* chassis, although results are based on a small dataset. Further experiments are expected to yield conclusive results.
11. Demonstrated the potential of *T. brucei* as a synthetic biology chassis, having developed, inserted and partially characterised a novel synthetic gene network within this chassis. This opens the door for further research regarding the use of *T. brucei* as a bioprocessing platform.

Chapter 8

Conclusion

Synthetic biology currently depends on the use of first generation chassis such as *E. coli* and mammalian CHO cells. This may limit its ability to address challenges in areas as diverse as bioprocessing, biotherapeutics, bioremediation and energy generation. In order to address this gap, this multi-faceted project incorporating biochemistry, mathematics and engineering initiated an investigative process to establish *Roseobacter* marine bacteria and unicellular *T. brucei* protozoa as synthetic biology chassis.

1. Achievements regarding the establishment of *Roseobacter* clade marine bacteria as a synthetic biology chassis by identifying reliable molecular biology procedures for the development of a synthetic gene network.

The clade of *Roseobacter* marine bacteria was selected for the purpose of expanding the set of available synthetic biology chassis because strains are naturally found in various diverse oceanic habitats and are implicated in balancing global carbon levels (Brinkhoff et al., 2008; Buchan et al., 2005). Moreover, *Roseobacter* utilise mixotrophic metabolism (Sato-Takabe et al., 2014; Swingley et al., 2007), employ quorum signalling (Moran et al., 2007) and are easily cultivated (Brinkhoff et al., 2008; Buchan et al., 2005). Therefore, they hold potential as chassis for the application of synthetic biology in addressing marine bioremediation and geo-engineering challenges.

The following achievements towards the establishment of a standardised and reliable synthetic biology practice in *Roseobacter* were made during the investigation on the *Roseobacter* strains *O. indolifex*, *D. shibae* and *R. denitrificans*, drawn from across the *Roseobacter* phylogenetic spectrum:

- Established that cell banks can be developed for *Roseobacter* strains. Glycerol stocks of the strains stored at -80 °C were shown to revive well in both shaken and static cultures for a period of 12 weeks.
- Demonstrated that all three marine bacterial strains could be successfully cultured at 37 °C. This enabled the adoption of a more standard practice, bypassing the requirement for different culturing temperatures for the different strains.
- Deduced that antibiotic selectivity, which is fundamental to cloning procedures, can be used to inhibit cell growth of the three wildtype *Roseobacter* strains. Cell growth was inhibited

by Chloramphenicol concentrations as low as $17\mu\text{g}/\text{ml}$ instead of the previously reported $30\mu\text{g}/\text{ml}$ (Piekarski et al., 2009).

- Mapped the region of experimental design space within which the three *Roseobacter* strains are resistant to plasmid uptake with respect to plasmid DNA concentration, origin of replication, voltage and transformation method used. This information will be of paramount importance in future research on the development of a transformation procedure for *Roseobacter* clade bacteria.
- **Examined the potential of *Roseobacter* clade bacteria as a synthetic biology chassis. This was the first project of its nature. The achievements discussed above have contributed to the development of reliable molecular biology procedures for use with *Roseobacter* platforms. This study has also demonstrated that synthetic biology within *Roseobacter* is non-trivial and is not sufficiently de-skilled as to allow for the participation of citizen scientists.**

Future work. In order to establish *Roseobacter* clade bacteria as synthetic biology host cells, the first recommendation is to develop a reliable, simple and rapid transformation procedure which allows *Roseobacter* cells to stably uptake foreign plasmid DNA. Moreover, alternative methods such as colony PCR should be incorporated when assessing success of transformation. These developments will allow for the participation of both professional and citizen scientists. Moreover, this will enable the characterisation and modelling of synthetic GRNs within *Roseobacter*.

2. Achievements regarding the establishment of *T. brucei* as a synthetic biology chassis by designing, modelling, constructing and measuring a Goodwin oscillator encoded by a novel synthetic gene network.

The eukaryote *T. brucei* protozoa were selected as potential synthetic biology host cells because they combine mechanisms found in both prokaryote and eukaryote species (Palenchar & Bellofatto, 2006; Bouvier et al., 2013), such as fast doubling times and mammalian-like post-translational glycosylation of host proteins. *T. brucei* make use of sophisticated antigen variation and genome modification mechanisms (Rudenko, 2011) to evade host immune systems. Moreover, *T. brucei* have been studied as a model organism for the research of protozoan biological processes (Serricchio & Butikofer, 2011), leading to well established culturing and cloning protocols via standard genetic engineering cloning techniques (Clayton & Shapira, 2007; Clayton, 1999). Therefore, there is scope for establishing the use of *T. brucei* in the application of synthetic biology within bioprocessing and to further probe its pathobiology.

The following achievements in the development of a novel synthetic GRN encoding a Goodwin oscillator to establish *T. brucei* as a chassis were made:

- Designed a GRN which functions as an autonomous oscillator via the use of a TetR inhibiting feedback loop and which made use of a green fluorescent reporter protein. Two versions of the oscillator were constructed in order to tune protein expression dynamics; pStable and pUnstable, one using a standard *tetR* gene and the other using a Ub-L destabilised *tetR* gene. The GRN was specifically designed for use in *T. brucei*, having features such as an rRNA sequence at which the plasmid was linearised for insertion into rRNA spacers in the genome and UTRs to enable protein synthesis. The network architecture incorporated

non-linear, time-delayed negative feedback which enabled oscillatory dynamics.

- Assembled the pStable oscillator plasmid using a combination of ligation techniques; standard, three-fragment, non-directional and Gibson. Throughout, the Gibson ligation method (Gibson & Bruck, 2000) proved to be the more efficient approach for the construction of the synthetic GRN. The pUnstable plasmid was synthesised commercially in six weeks (compared to eighteen months it took for the construction of pStable). Following this development, it is recommended that plasmid DNA is assembled and synthesised commercially to save both on time and costs. The assembly also led to the development and construction of three other plasmids having the same pRP backbone as the Goodwin oscillators but different combinations of the *tetR*, Ub-L and eGFP genes. Thus, the construction process resulted in five novel synthetic GRNs for use within trypanosomes. These can be provided to future researchers on request and can therefore be used in new lines of *T. brucei* research within synthetic biology, bioprocessing and pathobiology.
- Established an experimental framework for the imaging of live motile *T. brucei* cells. The set-up was required to constrict cell movement and allow for single cell tracking over a span of hours without compromising cell viability. This enables the imaging of cell populations and the tracking and identification of single cells, making for a time and resource efficient characterisation process. The developed set-up uses 40% Cygel™ to form an encasing gel matrix around *T. brucei* on a 5mm well slide base. The device restricted cell movement, thus allowing for cells to be tracked. *T. brucei* maintained viability for over two hours. The establishment of this framework from scratch has highlighted the importance of having the appropriate technology and equipment for the purpose of the study. No other similar device is available on the market. Therefore, this set-up can be used for the immobilisation of other motile organisms than *T. brucei*.
- Developed from scratch a quantitative mathematical model of the pStable and pUnstable Goodwin oscillator plasmids to complement characterisation studies. In line with synthetic biology principles, mathematical and wet lab analysis were used to inform each other and to build on the available knowledge in a cyclical pattern. The model was partly built using data from component characterisation in the pStable network (see below).
- Carried out preliminary quantitative analysis on the mathematical model via Gillespie simulations to screen the model's parameter space. Since *T. brucei*-specific biochemical rates were not available, it was necessary to examine whether oscillations occurred under very specific conditions or otherwise. Simulated trajectories repeatedly indicated that the GRN functioned as a low-level amplitude oscillator, irrespective of the parameters selected. Moreover, parameter analysis showed that protein synthesis and degradation rate combinations may be used to tune characteristics of the oscillatory output, since they show a small bias towards a higher or a lower maximum coordinate based on their values.
- Characterised the phenotype of the pStable and pUnstable Goodwin oscillators in both BSF and PCF *T. brucei*. Protein analysis was initially carried out in order to assess whether network components were functional. This was carried out via growth profiling assays, western blot assays and half-life assays which also served to inform mathematical modelling. Throughout, established molecular biology methods were used, highlighting the practicality of using *T. brucei* as host cells. Afterwards, fluorescent patterns in *T. brucei* were imaged via live cell population microscopy imaging and analysed at the single cell level. This allowed for deductions on functionality and protein expression dynamics within the networks.
- Demonstrated the both BSF and PCF cells bearing pStable and pUnstable do not have

impaired cell growth when compared to their wildtype counterparts. This implies that *T. brucei* host cells retain stability following the addition of <10kb extracellular plasmid DNA.

- Demonstrated that the components of pUnstable are functional, with regulation of expression governed by TetR inhibition which could be reversed via tetracycline induction. When tetracycline was added at a concentration of $1\mu\text{g/ml}$, the induction effect was noticed for over 6 hours when analysing protein sample. Throughout, protein was expressed at a constitutive basal level, indicating that the pUnstable network can be used as a tuned version of the pStable network.
- Demonstrated that the components of pStable are functional, with regulation of expression governed by TetR inhibition. Expression could be induced via tetracycline at a concentration of above 10fg/ml . When added at a concentration of $1\mu\text{g/ml}$, the induction effect was noticed for over 24 hours when analysing protein sample.
- Imaged BSF and PCF, induced and uninduced cell populations. No changes in fluorescent patterns were observed in BSF and uninduced PCF transfectants. This does not disprove previous characterisation carried out, but may reflect findings from quantitative mathematical analysis which indicate that the plasmid functions as a low-level oscillator. It is also an indication of the limits of the set-up and sensitivity of the imaging technology being used. This highlights how the production of oscillations and their detection are separate issues. However, in the images of induced PCF *T. brucei* bearing the pStable network, three cells were identified as having changes in fluorescent patterns. These were successfully recorded and quantified using a combination of computational techniques and visual confirmation. In order to extract the dominant trajectory, the noise components were removed by applying a smoothing filter and a cubic spline, resulting in an oscillatory eGFP trajectory for each cell with a period of c. 50 minutes and a varying amplitude. This is in agreement with the function of the network as an oscillator.
- Mathematically analysed the observed oscillations in fluorescence signals by reconstructing their phase space. Results showed that there was a linear growth trend within each trajectory, attributed to the low eGFP degradation rate, which results in a proportion of the eGFP not being degraded prior to the re-initiation of expression at the beginning of each cycle. Protein expression dynamics were again consistent with oscillations and the pStable network functioning as an oscillator.
- In conclusion, a novel plasmid encoding a Goodwin oscillator for implementation in a *T. brucei* chassis was designed, assembled, characterised and modelled. Results, albeit based on a small dataset, were consistent with the network functioning as an oscillator but also highlight the need for further conclusive data in order to establish the source of oscillations. The assembly process also led to the development of five novel trypanosomal synthetic GRNs, the first mathematical model of a synthetic GRN used within a *T. brucei* based study and the development of a device which allows for the imaging of live and motile cells over two hours while being able to track a proportion of the cells individually.
- **The investigation has shown that *T. brucei* hold potential as a synthetic biology platform and has contributed towards the diversification of the synthetic biology chassis pool by initiating the study of *T. brucei* as host cells. This was the first investigation of its kind and provides a solid foundation for future research on the use of *T. brucei* within the application of synthetic biology for bioprocessing and pathobiology.**

Future work. Based on the above findings, the following main recommendations are set out for

future work and the continuation of the investigation of *T. brucei* as synthetic biology chassis. The current set-up for the imaging of motile *T. brucei* cells should be optimised so as to allow for more limited cell motility and longer imaging times. The current device highlighted the challenge in maintaining cells viable for over two hours while immobilising them. The set-up should draw from microfluidics in order to allow for the constant supply of nutrients to embedded cells while constricting their motility. In light of previous results, wherein findings made when using an inverted microscope did not translate fully to when using a fluorescent microscope, all research should be carried out within the same conditions as when imaging.

This development is important not just for the analysis of live *T. brucei* but for the analysis of live, motile unicellular organisms in general. Once this is in place, it is recommended that further microscopy imaging is carried out on pStable and pUnstable in order to generate a larger dataset of recorded and quantified fluorescence trajectories from a representative sample of the cell population. This will allow for further characterisation of the network phenotype and the development of a more detailed mathematical model based on parameter estimation.

The need for a second-generation mathematical model is coupled with the need for deterministic mathematical modelling. This will help to differentiate between genuine or stochastic fluctuations and help inform whether the low level oscillations observed in both simulations and imaging were a result of noise or not. Moreover, a second generation model is required in order to be able to analyse the relation between the frequencies of the observed and simulated trajectories.

Protein analysis highlighted the need for easily characterised parts within synthetic biology. During protein assays, for example, the TetR protein was not detected. When a Myc tag was added upstream of the TetR gene in the pUnstable network, no protein was detected either. This was not for lack of protein being expressed. Moreover, in adding the Myc tag upstream of the *tetR* gene, the ability of the TetR protein to bind to the operator appeared to be affected. This needs to be investigated further in order to develop systems which can be characterised more efficiently or to develop mathematical models which allow for the calculation of an output as a function of, for example, TetR protein.

This work has highlighted a novel challenge which synthetic biologists should seek to address immediately. Based on the available data, there is no clearly defined set of guidelines on what constitutes a host cell within synthetic biology. A document needs to be developed in collaboration with leaders within the field of synthetic biology, defining the requirements for an organism to be classified as a synthetic biology chassis. This will clarify the studies which need to be carried out when investigating species as novel host cells in the future.

Conclusive statement.

In conclusion, this work contributed towards addressing the current limitations of first generation synthetic biology host cells by investigating two novel potential platforms; *Roseobacter* marine clade bacteria and lower eukaryote unicellular *T. brucei* as synthetic biology chassis. This has opened up a novel and interesting area of investigation within synthetic biology with the potential to influence the application of synthetic biology in fields as diverse as bioprocessing, biotherapeutics and bioremediation.

Appendices

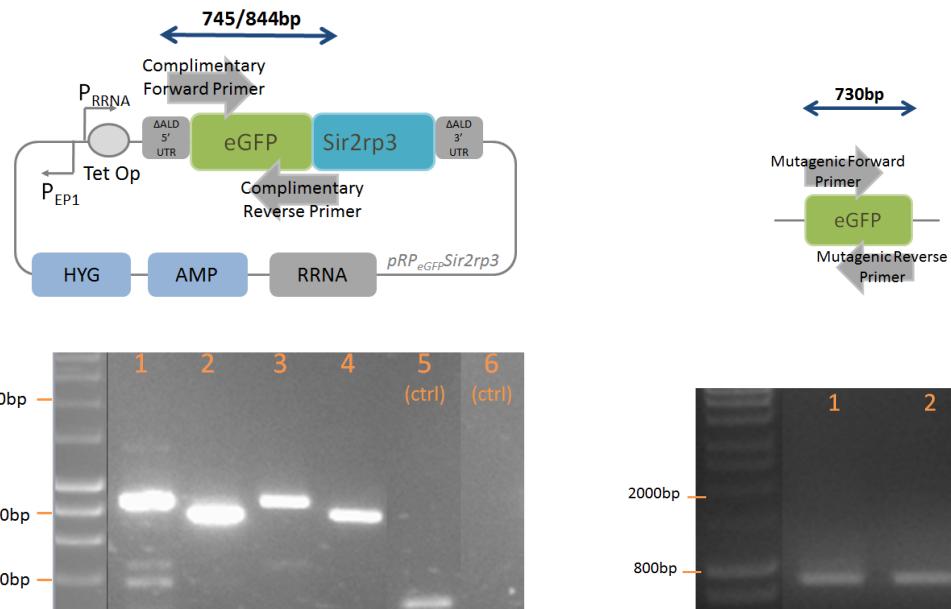
Appendix A

Ligation recipes and gel electrophoresis results from cloning during construction of the stable Goodwin oscillator

This appendix lists the ligation recipes which were developed and used during each of the four assembly strategies in the construction of the stable Goodwin oscillator. This appendix also contains images of gel electrophoresis runs using amplified, digested or ligated DNA cassettes and plasmids generated during the cloning process.

A.1 Oscillator assembly strategy using non-directional ligation

The following recipes and gel images were generated during the first ligation strategy using non-directional ligation to construct the stable Goodwin oscillator.



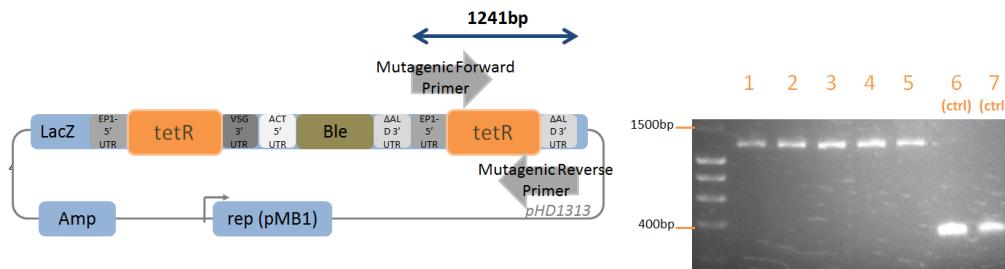
(a) Diagram on top shows the segment on pRP_{eGFP}SIR2rp3 to be amplified using **complementary primers**. The bottom panel shows gel results for PCR. Lanes: 5 μ l

- 1) PCR using first forward primer and reverse primer with use of DMSO and HF buffer,
- 2) PCR using second forward primer and reverse primer with use of DMSO and HF buffer,
- 3) PCR using first forward primer and reverse primer with DMSO and GC buffer,
- 4) PCR using second forward primer and reverse primer with DMSO and GC buffer,
- 5) Positive control using M2M3 plasmid,
- 6) Negative control using no DNA template.

(b) Diagram on top shows the segment on the previously amplified segment to be PCR'd using **mutagenic primers**. The bottom panel shows gel results for the reaction. Lanes:

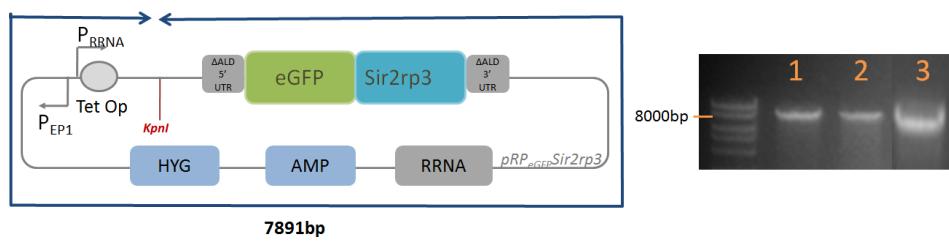
- 1-2) 5 μ l eGFP, amplified and purified. (Positive and negative controls not shown).

Figure A.1: Assembly strategy using non-directional ligation - Plasmid preparation: Two-step process to amplify eGFP from pRP_{eGFP}SIR2rp3 and add overhangs using: a) complementary primers on pRP_{eGFP}SIR2rp3 in the first PCR step and b) mutagenic primers on previously amplified DNA segment in the second PCR step.



(a) Diagram on the left shows region of pHD1313 plasmid to be amplified using mutagenic primers in order to extract *tetR* cassette and add *Kpn*I sites. The right panel shows PCR results. Lanes:

- 1) 5 μ l *tetR* cassette, amplified using 10ng/ μ l template with DMSO,
- 2) 5 μ l *tetR* cassette, amplified using 50ng/ μ l template with DMSO,
- 3) 5 μ l *tetR* cassette, amplified using 50ng/ μ l template without DMSO,
- 4) 5 μ l *tetR* cassette, amplified using 100ng/ μ l template with DMSO,
- 5) 5 μ l *tetR* cassette, amplified using 100ng/ μ l template without DMSO,
- 6) 5 μ l Positive control PCR using M2M3 plasmid with DMSO,
- 7) 5 μ l Positive control PCR using M2M3 plasmid without DMSO.



(b) Diagram on the left shows pRP_{eGFP}SIR2rp3 cut with *Kpn*I for non-directional ligation of the *tetR* cassette into it. The right panel shows digest results. Lanes:

- 1-2) 5 μ l pRP_{eGFP}SIR2rp3, *Kpn*I cut & dephosphorylated & purified
- 3) 5 μ l pRP_{eGFP}SIR2rp3, *Kpn*I cut & purified.

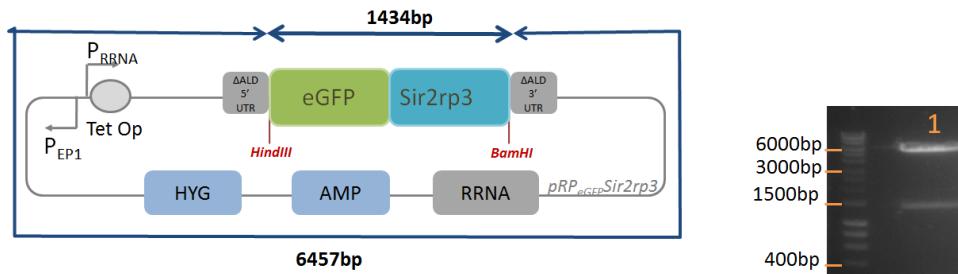
Figure A.2: Assembly strategy using non-directional ligation - Plasmid preparation: Gel results showing the: a) *tetR* cassette and b) pRP_{eGFP}SIR2rp3 during preparation steps for a *Kpn*I non-directional ligation to insert the *tetR* cassette into pRP_{eGFP}SIR2rp3.

		Ligation								
		1	2	3	4	5	6	7	8	9
Insert:Backbone molar ratio		3:1	2:1	3:1	2:1	-	-	-	-	-
Ingredient		Amount	Amount	Amount	Amount	Amount	Amount	Amount	Amount	Amount
pRP _{egFP} SIR2rp3 backbone		3 μ l	3 μ l	3 μ l	3 μ l					
1	Cut, dephos., pur. (29ng/ μ l)	(11mols) (87ng)	(11mols) (87ng)	(11mols) (87ng)	(11mols) (87ng)	(11mols) (87ng)	(11mols) (87ng)	-	-	-
pRP _{egFP} SIR2rp3 backbone		-	-	-	-	-	-	-	-	-
2	Cut, pur. (68ng/ μ l)	-	-	-	-	-	-	-	-	-
pRP _{egFP} SIR2rp3 backbone		-	-	-	-	-	-	-	-	-
3	Untcut (55ng/ μ l)	-	-	-	-	-	-	-	-	-
<i>tetR</i> insert		3.3 μ l	2.2 μ l	-	-	-	-	-	-	-
4	Cut, dephos., pur. (12 ng/ μ l)	(33mols)	(22mols)	-	-	-	-	-	-	-
<i>tetR</i> insert		-	-	6.6 μ l	4.4 μ l	-	-	-	-	-
5	Cut, pur. (6 ng/ μ l)	-	-	(33mols)	(22mols)	-	-	-	-	-
6	Quick T4 DNA Ligase	1 μ l	1 μ l	1 μ l	-					
7	2X Ligase Buffer	10 μ l	10 μ l	10 μ l	10 μ l					
8	dH ₂ O	to 21 μ l	to 21 μ l	to 21 μ l	to 21 μ l					
Total		21 μ l	21 μ l	21 μ l	21 μ l					
Colony Count (W3110)	3	16	25	9	lawn	lawn	lawn	lawn	lawn	lawn
Colony Count (TOP10)	lawn	lawn	lawn	lawn	lawn	lawn	lawn	lawn	lawn	lawn

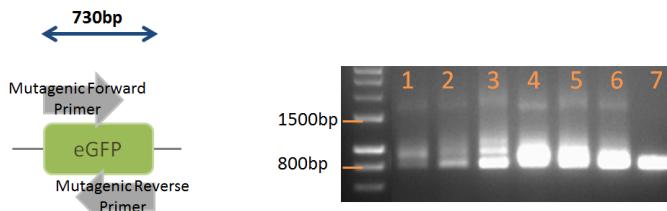
Table A.1: Assembly strategy using non-directional ligation - Ligation: Set-up of ligations and controls for the KpnI non-directional ligation of the *tetR* cassette into pRP_{egFP}SIR2rp3, using Quick T4 DNA ligase. Each ligation lists the cut dephosphorylated purified backbone or uncut backbone in volume, mols and mass, the cut dephosphorylated purified insert or cut purified insert in volume, mols and mass (optional), the volume of Quick T4 DNA ligase (optional), 2X Ligase Buffer and dH₂O (optional) used for ligation. Colony counts are from transformation plates following incubation times. 'dephos.' stands for dephosphorylated plasmid, 'pur.' stands for purified plasmids, 'Lawn' refers to +250 small, non-isolated colonies.

A.2 Oscillator assembly strategy using three-fragment ligation

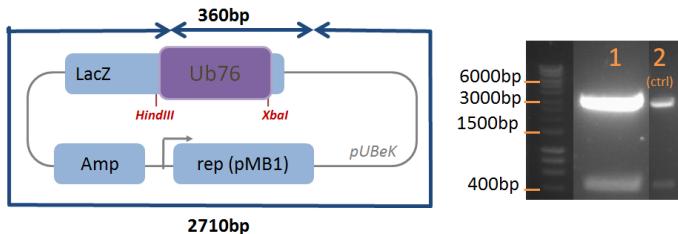
The following recipes and gel images were generated during the second ligation strategy using a three-fragment ligation to construct the stable Goodwin oscillator.



(a) Diagram on the left shows pRP_{eGFP}Sir2rp3 cut with HindIII/BamHI for three-fragment ligation of Ub-L and eGFP into it. The right panel shows gel results of 50 μl pRP_{eGFP}Sir2rp3 backbone digested with HindIII and BamHI.



(b) Diagram on the left shows eGFP region amplified using mutagenic primers in preparation for three-fragment ligation of Ub-L and eGFP into pRP_{eGFP}Sir2rp3. The right panel shows PCR results: 1-7) 5 μl eGFP, amplified using complementary and mutagenic primers.



(c) Diagram on the left shows pUBeK cut with HindIII/XbaI for three-fragment ligation of Ub-L and eGFP into pRP_{eGFP}Sir2rp3. The right panel shows digest results. Lanes:
1) 50 μl pUBeK, HindIII/XbaI cut,
2) 5 μl pUBeK, HindIII cut.

Figure A.3: Assembly strategy using three-fragment ligation, Plasmid preparation: Gel results showing: a) pRP_{eGFP}Sir2rp3 backbone, b) eGFP insert and c) Ub-L insert during the preparation process for the HindIII/XbaI/BamHI three-fragment ligation of Ub-L and eGFP into pRP_{eGFP}Sir2rp3.

	Insert:Backbone molar ratio	Ligation				
		1 3:1	2 1:1	3	4	5
	Ingredient	Amount	Amount	Amount	Amount	Amount
1	pRP _{eGFP} SIR2rp3 backbone Cut, purified (240ng/μl)	4μl (148mols) (960ng)	4μl (148mols) (960ng)	4μl (148mols) (960ng)	4μl (148mol) (960ng)	-
2	pRP _{eGFP} SIR2rp3 backbone Uncut (325ng/μl)	-	-	-	-	3.6μl (148mols) (1167ng)
3	Ub-L insert Cut, purified (39 ng/μl)	4.1μl (443mols)	1.4μl (148mols)	-	-	-
4	eGFP insert Amplified, cut, purified (160 ng/μl)	2μl (443mols)	0.7μl (148mols)	-	-	-
5	Quick T4 DNA Ligase	1μl	1μl	1μl	-	-
6	2X Ligase Buffer	10μl	10μl	10μl	10μl	10μl
7	dH ₂ O	to 21μl	to 21μl	to 21μl	to 21μl	to 21μl
	Total	21μl	21μl	21μl	21μl	21μl
	Colony Count	27	42	32	0	Lawn

Table A.2: Assembly strategy using three-fragment ligation - Ligation: Set-up of ligations and controls for the HindIII/XbaI/BamHI three-fragment ligation using increased backbone and insert mass to insert Ub-L and eGFP into pRP_{eGFP}SIR2rp3 to get pConstitutive, using Quick T4 DNA ligase. Each ligation lists the cut purified backbone or uncut backbone in volume, mols and mass, the insert in volume, mols and mass (optional), the volume of Quick T4 DNA ligase (optional), 2X Ligase Buffer and dH₂O (optional) used for ligation. Colony counts are from transformation plates following incubation times. 'Lawn' refers to +250 small, non-isolated colonies.

	Insert:Backbone molar ratio	Ligation				
		1 3:1	2 2:1	3 -	4 -	5 -
1	Ingredient	Amount	Amount	Amount	Amount	Amount
1	pRP _{eGFP} SIR2rp3 backbone Cut, purified (212ng/μl)	6μl (178mols) (1272ng)	6μl (178mols) (1272ng)	6μl (178mols) (1272ng)	6μl (178mols) (1272ng)	-
2	pRP _{eGFP} SIR2rp3 backbone Uncut textcolor{white}a (325ng/μl)	-	-	-	-	4.3μl (178mols) (1396ng)
3	eGFP insert Amplified, cut, purified (160ng/μl)	2.4μl (530mols)	1.6μl (353mols)	-	-	-
4	Quick T4 DNA Ligase	1μl	1μl	1μl	-	-
5	2X Ligase Buffer	10μl	10μl	10μl	10μl	10μl
6	dH ₂ O	to 21μl				
	Total	21μl	21μl	21μl	21μl	21μl
	Colony Count	14	18	26	0	+250

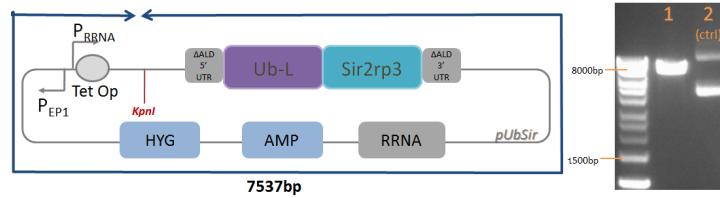
Table A.3: Assembly strategy using three-fragment ligation - Ligation: Set-up of ligations and controls for an XbaI/BamHI directional ligation to insert eGFP into pRP_{eGFP}SIR2rp3 as an alternative to the three-fragment ligation. Each ligation lists the cut purified backbone or uncut backbone in volume, mols and mass, the insert in volume, mols and mass (optional), the volume of Quick T4 DNA ligase (optional), 2X Ligase Buffer and dH₂O (optional) used for ligation. Colony counts are from transformation plates following incubation times.

		Ligation								
		1	2	3	4	5	6	7	8	9
Insert:Backbone molar ratio		3:1	2:1	3:1	2:1	-	-	-	-	-
Ingredient		Amount	Amount	Amount	Amount	Amount	Amount	Amount	Amount	Amount
pRP _{egFP} SIR2rp3 backbone		1 μ l	1 μ l	1 μ l	1 μ l	1 μ l	1 μ l	1 μ l	1 μ l	1 μ l
1	Cut, dephos., pur. (89ng/ μ l)	(11mols) (89ng)	(11mols) (89ng)	(11mols) (89ng)	(11mols) (89ng)	(11mols) (89ng)	(11mols) (89ng)	-	-	-
pRP _{egFP} SIR2rp3 backbone		-	-	-	-	-	-	-	-	-
2	Cut, pur. (68ng/ μ l)	-	-	-	-	-	-	-	-	-
pRP _{egFP} SIR2rp3 backbone		-	-	-	-	-	-	-	-	-
3	Uncut (50ng/ μ l)	-	-	-	-	-	-	-	-	-
<i>tetR</i> insert		3.4 μ l	2.3 μ l	-	-	-	-	-	-	-
4	Cut, dephos., pur. (12 ng/ μ l)	(34mols)	(23mols)	-	-	-	-	-	-	-
<i>tetR</i> insert		-	-	4.1 μ l (34mols)	2.7 μ l (23mols)	-	-	-	-	-
5	Cut, pur. (10 ng/ μ l)	-	-	-	-	-	-	-	-	-
6	Quick T4 DNA Ligase	1 μ l	1 μ l	1 μ l	1 μ l	1 μ l	1 μ l	-	1 μ l	-
7	2X Ligase Buffer	10 μ l	10 μ l	10 μ l	10 μ l	10 μ l	10 μ l	10 μ l	10 μ l	10 μ l
8	dH ₂ O	to 21 μ l	to 21 μ l	to 21 μ l	to 21 μ l	to 21 μ l	to 21 μ l	to 21 μ l	to 21 μ l	to 21 μ l
Total		21 μ l	21 μ l	21 μ l	21 μ l	21 μ l	21 μ l	21 μ l	21 μ l	21 μ l
Colony Count		11	37	57	48	15	15	50	40	30

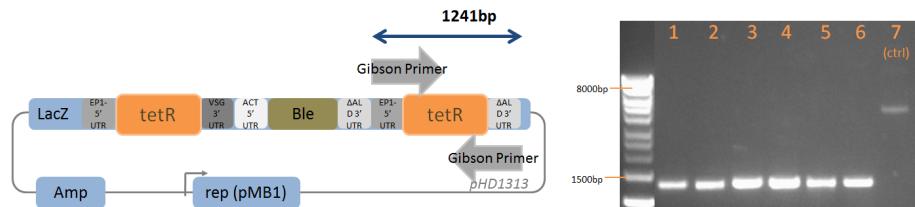
Table A.4: Assembly strategy using three-fragment ligation - Ligation: Set-up of ligations and controls for the KpnI non-directional ligation of the *tetR* cassette into pRP_{egFP}SIR2rp3, using Quick T4 DNA ligase. Each ligation lists the cut dephosphorylate purified backbone or uncut backbone in volume, mols and mass, the insert in volume, mols and mass (optional), the volume of Quick T4 DNA ligase (optional), 2X Ligase Buffer and dH₂O (optional) used for ligation. Colony counts are from transformation plates following incubation times. 'dephos.' stands for dephosphorylated plasmid, 'pur.' stands for purified plasmids.

A.3 Oscillator assembly strategy inserting the *tetR* cassette followed by the eGFP gene via Gibson assembly

The following recipes and gel images were generated during the third ligation strategy using a Gibson ligation to insert the *tetR* cassette followed by the eGFP gene to construct the stable Goodwin oscillator.



(a) Diagram on the left shows the pUbSir plasmid digested with KpnI in preparation of Gibson ligations to insert *tetR* cassette. The right panel shows digest results. Lanes:
 1) 5µl pUbSir, KpnI cut,
 2) 5µl pUbSir, uncut.

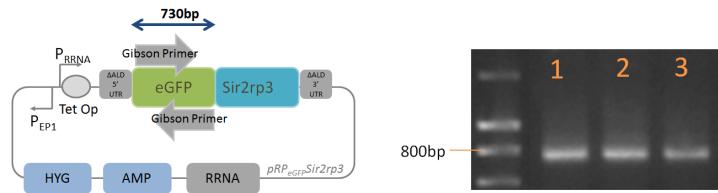


(b) Diagram on the left shows region of pHd1313 plasmid to be amplified using Gibson primers in order to extract *tetR* cassette. The right panel shows PCR results. Lanes:
 1-2) 10µl *tetR* cassette, amplified from pHd1313 with Gibson primers using annealing temperature of 55°C
 3-4) 10µl *tetR* cassette, amplified from pHd1313 with Gibson primers using annealing temperature of 60°C
 5-6) 10µl *tetR* cassette, amplified from pHd1313 with Gibson primers using annealing temperature of 72°C
 7) 5µl *tetR*, uncut.

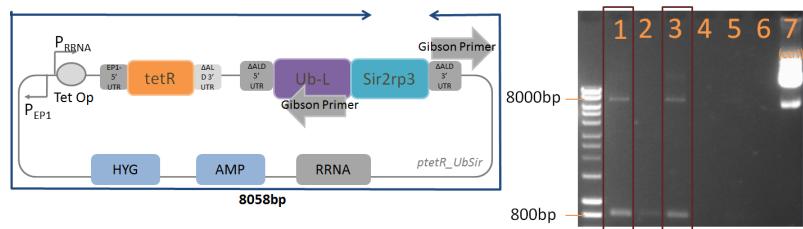
Figure A.4: TetR-first assembly strategy - Plasmid preparation: Gel results showing the: a) pUbSir backbone and b) *tetR* cassette during preparation steps for a Gibson ligation to insert the *tetR* cassette into pUbSir to get ptetR_UbSir.

		Ligation				
		1	2	3	4	5
Insert:Backbone molar ratio		3:1	2:1	-	-	-
Ingredient		Amount	Amount	Amount	Amount	Amount
1	pUbSir backbone Cut, purified (80ng/ μ l)	3 μ l (0.048pmol) (240ng)	3 μ l (0.048pmol) (240ng)	3 μ l (0.048pmol) (240ng)	3 μ l (0.048pmol) (240ng)	-
2	Positive Control (NEB) (325ng/ μ l)	-	-	-	-	10 μ l
3	<i>tetR</i> insert Amplified, purified (20ng/ μ l)	8.7 μ l (0.145pmol)	5.8 μ l (0.097pmol)	-	-	-
4	Gibson Master Mix	10 μ l	10 μ l	10 μ l	-	-
5	dH ₂ O	to 20 μ l	to 20 μ l			
Total		20 μ l	20 μ l	20 μ l	20 μ l	20 μ l
Colony Count		130	100	30	110	0

Table A.5: TetR-first assembly strategy - Ligation: Recipes for reactions and controls used in Gibson ligation to insert the *tetR* cassette into pUbSir to get ptetR_UbSir. Each ligation lists the cut purified backbone or control backbone in volume, mols and mass, the insert in volume, mols and mass (optional), the volume of Gibson Master Mix (optional) and dH₂O (optional) used for ligation. Colony counts are from transformation plates following incubation times.



(a) Diagram on the left shows region of *pRP_{eGFP}SIR2rp3* plasmid to be amplified using Gibson primers in order to extract eGFP. The right panel shows gel results of the amplification of eGFP from *pRP_{eGFP}SIR2rp3*. Lanes:
 1-2) 10 μ l eGFP, amplified using 2 μ l template, 58°C annealing temperature
 3) 10 μ l eGFP, amplified using 1 μ l template, 58°C annealing temperature.



(b) Diagram on the left shows region of *ptetR_UbSir* plasmid to be amplified as preparation for a Gibson ligation to insert eGFP into *ptetR_UbSir*. The right panel shows PCR results using different recipes and conditions. Lanes:
 1) 5 μ l *ptetR_UbSir*, amplified using HF buffer, 58°C annealing temp,
 2) 5 μ l *ptetR_UbSir*, amplified using GC buffer, 58°C annealing temp,
 3) 5 μ l *ptetR_UbSir*, amplified using HF buffer, 65°C annealing temp,
 4) 5 μ l *ptetR_UbSir*, amplified using GC buffer, 65°C annealing temp,
 5) 5 μ l *ptetR_UbSir*, amplified using HF buffer, 72°C annealing temp,
 6) 5 μ l *ptetR_UbSir*, amplified using GC buffer, 72°C annealing temp,
 7) 10 μ l *ptetR_UbSir*, uncut.

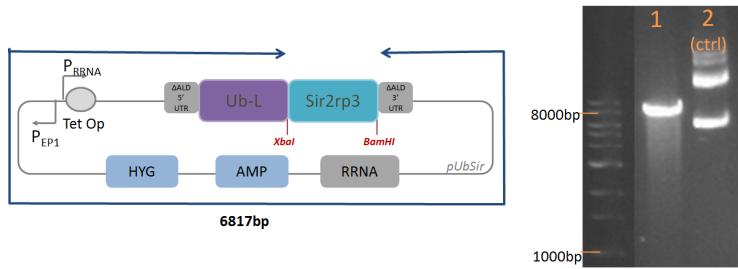
Figure A.5: TetR-first assembly strategy - Plasmid preparation: Gel results showing the: a) eGFP insert and b) *ptetR_UbSir* backbone during the preparation process for a Gibson ligation to insert eGFP into *ptetR_UbSir* to get *pStable*.

	Insert:Backbone molar ratio	Ligation			
		1	2	3	4
	Ingredient	Amount	Amount	Amount	Amount
1	ptetR_UbSir backbone Amplified, purified (16ng/ μ l)	8.5 μ l (0.026pmol) (136ng)	8.5 μ l (0.026pmol) (136ng)	8.5 μ l (0.026pmol) (136ng)	-
2	ptetR_UbSir backbone Uncut (100ng/ μ l)	-	-	-	1.4 μ l (0.026pmol) (136ng)
3	eGFP insert Amplified, purified (30ng/ μ l)	1.8 μ l (0.077pmol)	1.2 μ l (0.052pmol)	-	-
4	Gibson Master Mix	10 μ l	10 μ l	10 μ l	-
5	dH ₂ O	to 20 μ l			
	Total	20 μ l	20 μ l	20 μ l	20 μ l
	Colony Count	8	0	0	lawn

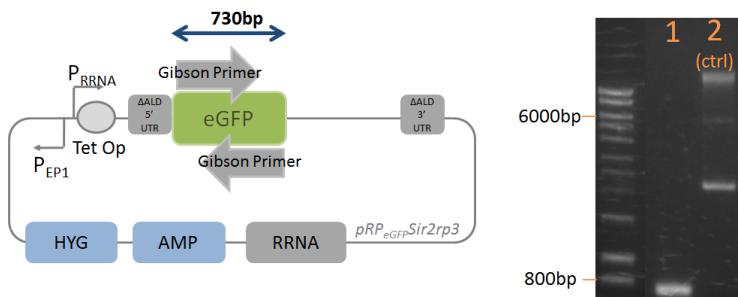
Table A.6: TetR-first assembly strategy - Ligation: Recipes for reactions and controls used in Gibson ligation to insert eGFP into ptetR_UbSir to get pStable. Each ligation lists the cut purified backbone or uncut backbone in volume, mols and mass, the insert in volume, mols and mass (optional), the volume of Gibson Master Mix (optional) and dH₂O (optional) used for ligation. Colony counts are from transformation plates following incubation times. 'Lawn' refers to +250 small, non-isolated colonies.

A.4 Oscillator assembly strategy inserting eGFP gene followed by the *tetR* cassette via Gibson assembly

The following recipes and gel images were generated during the fourth ligation strategy using a Gibson ligation to insert the eGFP gene followed by the *tetR* cassette to construct the stable Goodwin oscillator.



(a) Diagram on the left shows pUbSir amplified without the Sir2rp3 gene for Gibson ligation of the eGFP into it. The right panel shows PCR results. Lanes:
1) 10 μ l pUbSir, amplified with Gibson primers, purified,
2) 5 μ l pUbSir, uncut.

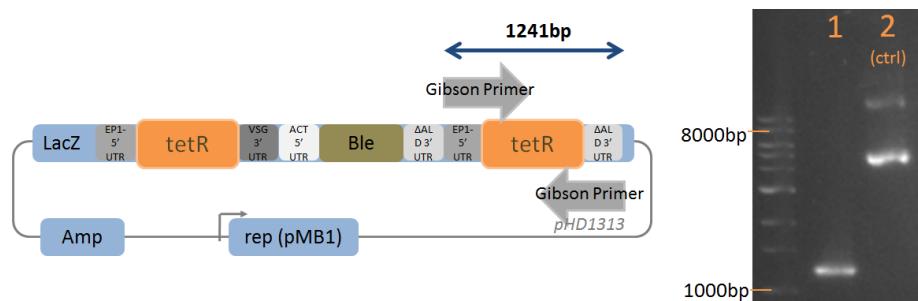


(b) Diagram on the left shows region of pRP_{eGFP}SIR2rp3 to be amplified using Gibson primers in order to extract eGFP for ligation into pUbSir. The right panel shows PCR results. Lanes:
1) 10 μ l eGFP, amplified with Gibson primers, purified
2) 5 μ l pRP_{eGFP}SIR2rp3, uncut.

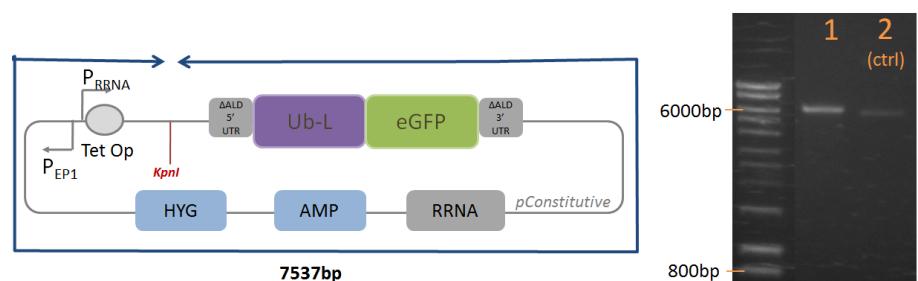
Figure A.6: eGFP-first assembly strategy - Plasmid preparation: a) Purified pUbSir backbone and b) Purified eGFP insert in preparation for the Gibson ligation to insert eGFP into pUbSir to get pConstitutive.

		Ligation			
Insert:Backbone molar ratio		1	2	3	4
1	Ingredient	Amount	Amount	Amount	Amount
1	pUbSir backbone Amplified, purified (15ng/ μ l)	5 μ l (0.017pmol) (75ng)	5 μ l (0.017pmol) (75ng)	5 μ l (0.017pmol) (75ng)	-
2	pUbSir backbone Uncut (50ng/ μ l)	-	-	-	1.5 μ l (0.017pmol) (75ng)
3	eGFP insert Amplified, purified (24ng/ μ l)	5 μ l (0.175pmol)	1 μ l (0.033pmol)	-	-
4	Gibson Master Mix	10 μ l	10 μ l	10 μ l	-
5	dH ₂ O	to 20 μ l			
	Total	20 μ l	20 μ l	20 μ l	20 μ l
	Colony Count	1	10	2	lawn

Table A.7: eGFP-first assembly strategy - Ligation: Recipes for reactions and controls used in Gibson ligation to insert eGFP into pUbSir to get pConstitutive. Each ligation lists the cut purified backbone or uncut backbone in volume, mols and mass, the insert in volume, mols and mass (optional), the volume of Gibson Master Mix (optional) and dH₂O (optional) used for ligation. Colony counts are from transformation plates following incubation times. 'Lawn' refers to +250 small, non-isolated colonies.



(a) Diagram on the left shows the region on pHD1313 to be amplified to extract the *tetR* cassette in preparation for a Gibson ligation to insert the cassette into pConstitutive. The right panel shows PCR results. Lanes:
1) 10 μ l *tetR* cassette, amplified with Gibson primers & purified,
2) 5 μ l pHD1313, uncut.



(b) Diagram on the left shows pConstitutive digested with KpnI digest in preparation for a Gibson ligation to insert the *tetR* cassette into it. The right panel shows digest results. Lanes:
1) 10 μ l pConstitutive, KpnI cut & purified,
2) 5 μ l pConstitutive, uncut.

Figure A.7: eGFP-first assembly strategy, Plasmid preparation: Gel results showing: a) *tetR* and b) pConstitutive DNA amplified, digested and purified in preparation for a Gibson ligation to insert *tetR* into pConstitutive to get pStable.

	Insert:Backbone molar ratio	Ligation					
		1 3:1	2 2:1	3 3:1	4 2:1	5	6
	Ingredient	Amount	Amount	Amount	Amount	Amount	Amount
1	pConstitutive backbone Cut (KpnI), purified (117ng/ μ l)	1 μ l (0.024pmol) (117ng)	1 μ l (0.024pmol) (117ng)	-	-	1 μ l (0.024pmol) (117ng)	-
2	pConstitutive backbone Cut (KpnI & DpnI), purified (13ng/ μ l)	-	-	9 μ l (0.024pmol) (117ng)	9 μ l (0.024pmol) (117ng)	-	9 μ l (0.024pmol) (117ng)
3	pConstitutive backbone Uncut (220ng/ μ l)	-	-	-	-	-	0.5 μ l (0.022pmol) (110ng)
4	<i>tetR</i> insert Amplified, purified (43ng/ μ l)	2 μ l (0.071pmol) (85ng)	1.3 μ l (0.047pmol) (57ng)	2 μ l (0.071pmol) (85ng)	1.3 μ l (0.047pmol) (57ng)	-	-
5	Gibson Master Mix	10 μ l to 20 μ l	10 μ l to 20 μ l	10 μ l to 20 μ l	10 μ l to 20 μ l	10 μ l to 20 μ l	10 μ l to 20 μ l
6	dH ₂ O	20 μ l	20 μ l	20 μ l	20 μ l	20 μ l	20 μ l
	Colony Count	53	82	29	0	0	1 +250

Table A.8: eGFP-first assembly strategy - Ligation: Recipes for reactions and controls used in Gibson ligation to insert *tetR* into pConstitutive to get pStable. Each ligation lists the cut purified backbone or uncut backbone in volume, mols and mass, the insert in volume, mols and mass (optional), the volume of Gibson Master Mix (optional) and dH₂O (optional) used for ligation. Colony counts are from transformation plates following incubation times.

Appendix B

List of primers used during PCR amplification

This appendix lists all primers used in the PCR amplification processes in preparation for standard, three-fragment, non-directional and Gibson ligations, with mutagenic nucleotides in bold and overhangs underlined. In addition, the primer length, direction and a brief description of the primers is included. These have been referred to in Chapter 4. All primers were ordered from Eurofins MWG Operon (Ebersberg, Germany).

1	Length: 37bp	Direction: Forward	Amplifies: <i>eGFP</i> gene
Sequence(5' - 3'): <u>TCT</u> AGA GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG G			
Notes: Replace ATG start codon with AGA codon and HindIII site with in-frame XbaI site.			
2	Length: 32bp	Direction: Reverse	Amplifies: <i>eGFP</i> gene
Sequence(5' - 3'): <u>GGA</u> TCC GCC TTC AAG ACT TGT ACA GCT CGT CC			
Notes: Replace AGA with TGA stop codon and XbaI site with BamHI site			
3	Length: 21bp	Direction: Forward	Amplifies: <i>eGFP</i> gene
Sequence(5' - 3'): GAG TGG TAC CCT GCA CGC CGC			
4	Length: 21bp	Direction: Forward	Amplifies: <i>eGFP</i> gene
Sequence(5' - 3'): ATG GTG AGC AAG GGC GAG GAG			
5	Length: 21bp	Direction: Reverse	Amplifies: <i>eGFP</i> gene
Sequence(5' - 3'): AAT CAT ACG ATT GGG CCG CC			
6	Length: 31bp	Direction: Forward	Amplifies: <i>tetR</i> cassette
Sequence(5' - 3'): GGT ACC GTC TTG GTG TGT CGA CCT TGC AGG C			
Notes: Introduces KpnI site upstream of the <i>tetR</i> cassette			
7	Length: 33bp	Direction: Reverse	Amplifies: <i>tetR</i> cassette
Sequence(5' - 3'): <u>GGT</u> ACC TTG TAC ATA TTG TCG TTA GAA CGC GGC			
Notes: Introduces KpnI site downstream of the <i>tetR</i> cassette			
8	Length: 38bp	Direction: Forward	Amplifies: <i>tetR</i> cassette
Sequence(5' - 3'): <u>CAA</u> TGA TAG AGT GGT ACC GTC TTG GTG TGT CGA CCT TG			
Notes: Adds pRP compatible overhangs, Used for Gibson ligation.			
9	Length: 40bp	Direction: Reverse	Amplifies: <i>tetR</i> cassette
Sequence(5' - 3'): <u>GCG</u> CGT GCA GGG TAC CTT GTA CAT ATT GTC GTT AGA ACG C			
Notes: Adds pRP compatible overhangs, Used for Gibson ligation.			
10	Length: 36bp	Direction: Forward	Amplifies: <i>eGFP</i> gene

Sequence(5' - 3'): <u>CTA GAC AAG TTT CTA GAG</u> AAG CTT ATG GTG AGC AAG			
Notes: Adds pRP compatible overhangs, Used for Gibson ligation.			
11	Length: 38bp	Direction: Reverse	Amplifies: eGFP gene
Sequence(5' - 3'): <u>GCC AAC TAA ATG GGC ATT</u> CTA GAC TTG TAC AGC TCG TC			
Notes: Adds pRP compatible overhangs, Used for Gibson ligation.			
12	Length: 18bp	Direction: Forward	Amplifies: pRP backbone
Sequence(5' - 3'): <u>TGC CCA TTT AGT TGG CTT</u>			
Notes: Used for Gibson ligation.			
13	Length: 25bp	Direction: Reverse	Amplifies: pRP backbone
Sequence(5' - 3'): <u>CTC TAG AAA CTT GTC TAG CCA ATT G</u>			
Notes: Used for Gibson ligation.			
14	Length: 22bp	Direction: Forward	Amplifies: pRP backbone
Sequence(5' - 3'): <u>GGA TCC TGC CCA TTT AGT TGG C</u>			
Notes: Used for Gibson ligation.			
15	Length: 19bp	Direction: Reverse	Amplifies: pRP backbone
Sequence(5' - 3'): <u>TCT AGA AAC TTG TCT AGC C</u>			
Notes: Used for Gibson ligation.			
16	Length: 35bp	Direction: Forward	Amplifies: eGFP gene
Sequence(5' - 3'): <u>CTA GAC AAG TTT CTA GAG</u> TGA GCA AGG GCG AGG AG			
Notes: Adds pRP compatible overhangs, Used for Gibson ligation.			
17	Length: 36bp	Direction: Reverse	Amplifies: eGFP gene
Sequence(5' - 3'): <u>GCC AAC TAA ATG GGC AGG ATC CGC CTT CAA GAC TTG</u>			
Notes: Adds pRP compatible overhangs, Used for Gibson ligation.			
18	Length: 38bp	Direction: Forward	Amplifies: Ub-L gene
Sequence(5' - 3'): <u>ACC AAA AAG TAA AAT TCA CAA GCT TAT GCA GAT CTT CG</u>			
Notes: Adds pRP compatible overhangs. Not used.			
19	Length: 34bp	Direction: Reverse	Amplifies: Ub-L gene
Sequence(5' - 3'): <u>CTT GCT CAC TCT AGA AAC TTG TCT AGC CAA TTG C</u>			
Notes: Adds pRP compatible overhangs. Not used.			
20	Length: 27bp	Direction: Forward	Amplifies: eGFP gene
Sequence(5' - 3'): <u>AGA CAA GTT TCT AGA GTG AGC AAG GGC</u>			
Notes: Adds pRP compatible overhangs. Not used.			
21	Length: 32bp	Direction: Forward	Amplifies: eGFP gene
Sequence(5' - 3'): <u>AGA TCT GCC TTC AAG ACT TGT ACA GCT CGT CC</u>			
Notes: Adds pRP compatible overhangs. Not used.			
22	Length: 36bp	Direction: Reverse	Amplifies: eGFP gene
Sequence(5' - 3'): <u>GCC AAC TAA ATG GGC AAG ATC TGC CTT CAA GAC TTG</u>			
Notes: Adds BgIII site, stop codon, pRP compatible overhangs. Not used.			

Table B.1: List of primers used throughout the construction process of the Goodwin oscillator. The nucleotides underlined represent overhangs. The nucleotides in bold are introduced by mutagenic PCR.

Appendix C

List of primers used during plasmid sequencing

This appendix lists the primers used in sequencing reactions following cloning in order to confirm success of the procedures. The list includes primer length, direction and which DNA segment the primer sequences. All primers were ordered from Eurofins MWG Operon (Ebersberg, Germany).

1	Length: 21bp	Direction: Forward	Sequences: pUBeK
Sequence(5' - 3'): AAG ATA AAG GGT ATC CCT CCG			
2	Length: 21bp	Direction: Forward	Sequences: pUBeK
Sequence(5' - 3'): ATA AGG AGG GTA TCC CTC CGG			
3	Length: 21bp	Direction: Reverse	Sequences: pUBeK
Sequence(5' - 3'): AAT CAG TCG CTG CTG ATC CGG			
4	Length: 20bp	Direction: Forward	Sequences: pRP backbone
Sequence(5' - 3'): AAA CGC GTG CGC TAC ACA GC			
5	Length: 21bp	Direction: Forward	Sequences: pRP backbone
Sequence(5' - 3'): AAT GAT AGA GTG GTA CCC TGC			
6	Length: 21bp	Direction: Forward	Sequences: pRP backbone
Sequence(5' - 3'): TAA ACG GCC ACA AGT TCA GCG			
7	Length: 19bp	Direction: Forward	Sequences: pRP backbone
Sequence(5' - 3'): CCA CTA CCA GCA GAA CAC C			
8	Length: 26bp	Direction: Forward	Sequences: pRP backbone
Sequence(5' - 3'): TTT CTC ATA GGC GTG TGC ACC TTC CG			
9	Length: 22bp	Direction: Forward	Sequences: pRP backbone
Sequence(5' - 3'): CCT GAT TCA CTC ATC GTT GAG C			
10	Length: 21bp	Direction: Reverse	Sequences: pRP backbone
Sequence(5' - 3'): TCT AAT TCC AGA CTA CTG GCG			
11	Length: 21bp	Direction: Reverse	Sequences: pRP backbone
Sequence(5' - 3'): CAA TTA TTA GGG AAG CGA CGC			
12	Length: 21bp	Direction: Reverse	Sequences: pRP backbone

Sequence(5' - 3'): AGG AAA CCA ACT TGT AGT GGC			
13	Length: 23bp	Direction: Reverse	Sequences: pRP backbone
Sequence(5' - 3'): ATG CAC CAA CGA GAC AAC CAA GC			
14	Length: 21bp	Direction: Reverse	Sequences: pRP backbone
Sequence(5' - 3'): TTG AAG TTC ACC TTG ATG CCG			
15	Length: 23bp	Direction: Reverse	Sequences: pRP backbone
Sequence(5' - 3'): AAG CAG TTT GAG TAT TTG GAG CC			
16	Length: 21bp	Direction: Reverse	Sequences: pRP backbone
Sequence(5' - 3'): TCA TTG ATA GGG TGT ACG CCG			
17	Length: 21bp	Direction: Forward	Sequences: pRP backbone
Sequence(5' - 3'): AGT AAA AGT AGC GCT TAC GGC			
18	Length: 21bp	Direction: Reverse	Sequences: pRP backbone
Sequence(5' - 3'): TCC AGA GCA ATG GTT TTA CCG			
19	Length: 19bp	Direction: Forward	Sequences: pRP backbone & <i>tetR</i> insert
Sequence(5' - 3'): ACC ATA CTC ACT TTT GCC C			
20	Length: 21bp	Direction: Forward	Sequences: pRP backbone & <i>tetR</i> insert
Sequence(5' - 3'): TTC CCG GAG TAA TCT GAT GGC			
21	Length: 21bp	Direction: Forward	Sequences: pRP backbone & <i>tetR</i> insert
Sequence(5' - 3'): ACA TTC AGA AGG AGT CGA CGC			
22	Length: 20bp	Direction: Forward	Sequences: pRP backbone & <i>tetR</i> insert
Sequence(5' - 3'): ATG AAG TGG TGC AAA ACG CG			
23	Length: 21bp	Direction: Reverse	Sequences: pRP backbone & <i>tetR</i> insert
Sequence(5' - 3'): ATC CCT TGA CTC AAA CAC GCG			
24	Length: 21bp	Direction: Forward	Sequences: pRP backbone & <i>tetR</i> insert
Sequence(5' - 3'): ACT GTG TTG ATA AGG GAC GGG			
25	Length: 21bp	Direction: Reverse	Sequences: pRP backbone & <i>tetR</i> insert
Sequence(5' - 3'): CAT TAA GCA GCT CTA ATG CGC			

Table C.1: List of primers used in sequencing reactions following standard and Gibson ligations in order to confirm success of reactions.

Appendix D

MATLAB code used to apply moving average smoothing technique

The following scripts were used in MATLAB (2012a) in order to impute data or apply the moving average filter to a data set of interest in order to smooth data and eliminate the effects of noise to better understand the underlying trend. The moving average filter used a window size of eight. Through 'CellA' is taken to be the data set of interest with size 1 row \times x columns where x is the number of data points. In the second script, x=60.

```
1 %Impute data to remove the data points quantified as zero
2
3 for(i=1:x)
4 if (CellA(i,x)<0.01)
5 CellA(i,x)=mean([CellA(i-3,x),CellA(i-2,x),CellA(i-1,x),CellA(i+1,x),
6 CellA(i+2,x),CellA(i+3,x)]);
7 end;
```

```
1 %Moving average filter, which takes into account beginning and end points
2
3 %Add extra column to each data set where moving average data will be
4 %printed
5 CellA=[CellA zeros(60,1)];
6
7 for(i=1:60)
8 if(i==1 )
9 CellA(i,3)=mean([CellA(i,2),CellA(i+1,2),CellA(i+2,2),CellA(i+3,2)
10 ,CellA(i+4,2),CellA(i+5,2),CellA(i+6,2),CellA(i+7,2),CellA(i+8,2)]);
11 elseif (i==2 )
12 CellA(i,3)=mean([CellA(i-1,2),CellA(i,2),CellA(i+1,2),CellA(i+2,2),
13 CellA(i+3,2),CellA(i+4,2),CellA(i+5,2),CellA(i+6,2),CellA(i+7,2),
14 CellA(i+8,2)]);
15 elseif(i==3 )
16 CellA(i,3)=mean([CellA(i-2,2),CellA(i-1,2),CellA(i,2),CellA(i+1,2),
17 CellA(i+2,2),CellA(i+3,2),CellA(i+4,2),CellA(i+5,2),CellA(i+6,2),
18 CellA(i+7,2),CellA(i+8,2)]);
19 elseif(i==4 )
20 CellA(i,3)=mean([CellA(i-3,2),CellA(i-2,2),CellA(i-1,2),CellA(i,2),
21 CellA(i+1,2),CellA(i+2,2),CellA(i+3,2),CellA(i+4,2),CellA(i+5,2),
22 CellA(i+6,2),CellA(i+7,2),CellA(i+8,2)]);
23 elseif(i==5 )
```

```

24     CellA(i,3)=mean([CellA(i-4,2),CellA(i-3,2),CellA(i-2,2),CellA(i-1,2),
25     CellA(i,2),CellA(i+1,2),CellA(i+2,2),CellA(i+3,2),CellA(i+4,2),
26     CellA(i+5,2),CellA(i+6,2),CellA(i+7,2),CellA(i+8,2)]);
27 elseif(i==6 )
28     CellA(i,3)=mean([CellA(i-5,2),CellA(i-4,2),CellA(i-3,2),CellA(i-2,2),
29     CellA(i-1,2),CellA(i,2),CellA(i+1,2),CellA(i+2,2),CellA(i+3,2),
30     CellA(i+4,2),CellA(i+5,2),CellA(i+6,2),CellA(i+7,2),CellA(i+8,2)]);
31 elseif(i==7 )
32     CellA(i,3)=mean([CellA(i-6,2),CellA(i-5,2),CellA(i-4,2),CellA(i-3,2),
33     CellA(i-2,2),CellA(i-1,2),CellA(i,2),CellA(i+1,2),CellA(i+2,2),
34     CellA(i+3,2),CellA(i+4,2),CellA(i+5,2),CellA(i+6,2),CellA(i+7,2),
35     CellA(i+8,2)]);
36 elseif(i==8 )
37     CellA(i,3)=mean([CellA(i-7,2),CellA(i-6,2),CellA(i-5,2),CellA(i-4,2),
38     CellA(i-3,2),CellA(i-2,2),CellA(i-1,2),CellA(i,2),CellA(i+1,2),
39     CellA(i+2,2),CellA(i+3,2),CellA(i+4,2),CellA(i+5,2),CellA(i+6,2),
40     CellA(i+7,2),CellA(i+8,2)]);
41
42 elseif(i==53 )
43     CellA(i,3)=mean([CellA(i-8,2),CellA(i-7,2),CellA(i-6,2),CellA(i-5,2),
44     CellA(i-4,2),CellA(i-3,2),CellA(i-2,2),CellA(i-1,2),CellA(i,2),
45     CellA(i+1,2),CellA(i+2,2),CellA(i+3,2),CellA(i+4,2),CellA(i+5,2),
46     CellA(i+6,2),CellA(i+7,2)]);
47 elseif(i==54 )
48     CellA(i,3)=mean([CellA(i-8,2),CellA(i-7,2),CellA(i-6,2),CellA(i-5,2),
49     CellA(i-4,2),CellA(i-3,2),CellA(i-2,2),CellA(i-1,2),CellA(i,2),
50     CellA(i+1,2),CellA(i+2,2),CellA(i+3,2),CellA(i+4,2),CellA(i+5,2),
51     CellA(i+6,2)]);
52 elseif(i==55 )
53     CellA(i,3)=mean([CellA(i-8,2),CellA(i-7,2),CellA(i-6,2),CellA(i-5,2),
54     CellA(i-4,2),CellA(i-3,2),CellA(i-2,2),CellA(i-1,2),CellA(i,2),
55     CellA(i+1,2),CellA(i+2,2),CellA(i+3,2),CellA(i+4,2),CellA(i+5,2)]);
56 elseif(i==56 )
57     CellA(i,3)=mean([CellA(i-8,2),CellA(i-7,2),CellA(i-6,2),CellA(i-5,2),
58     CellA(i-4,2),CellA(i-3,2),CellA(i-2,2),CellA(i-1,2),CellA(i,2),
59     CellA(i+1,2),CellA(i+2,2),CellA(i+3,2),CellA(i+4,2)]);
60 elseif(i==57 )
61     CellA(i,3)=mean([CellA(i-8,2),CellA(i-7,2),CellA(i-6,2),CellA(i-5,2),
62     CellA(i-4,2),CellA(i-3,2),CellA(i-2,2),CellA(i-1,2),CellA(i,2),
63     CellA(i+1,2),CellA(i+2,2),CellA(i+3,2)]);
64 elseif(i==58 )
65     CellA(i,3)=mean([CellA(i-8,2),CellA(i-7,2),CellA(i-6,2),CellA(i-5,2),
66     CellA(i-4,2),CellA(i-3,2),CellA(i-2,2),CellA(i-1,2),CellA(i,2),
67     CellA(i+1,2),CellA(i+2,2)]);
68 elseif(i==59 )
69     CellA(i,3)=mean([CellA(i-8,2),CellA(i-7,2),CellA(i-6,2),CellA(i-5,2),
70     CellA(i-4,2),CellA(i-3,2),CellA(i-2,2),CellA(i-1,2),CellA(i,2),
71     CellA(i+1,2)]);
72 elseif(i==60)
73     CellA(i,3)=mean([CellA(i-8,2),CellA(i-7,2),CellA(i-6,2),CellA(i-5,2),
74     CellA(i-4,2),CellA(i-3,2),CellA(i-2,2),CellA(i-1,2),CellA(i,2)]);
75 else
76     CellA(i,3)=mean([CellA(i-8,2),CellA(i-7,2),CellA(i-6,2),CellA(i-5,2),
77     CellA(i-4,2),CellA(i-3,2),CellA(i-2,2),CellA(i-1,2),CellA(i,2),
78     CellA(i+1,2),CellA(i+2,2),CellA(i+3,2),CellA(i+4,2),CellA(i+5,2),
79     CellA(i+6,2),CellA(i+7,2),CellA(i+8,2)]);
80 end;
81 end;

```

Appendix E

MATLAB code used for phase space reconstruction

This script was used in MATLAB (2012a) in order to reconstruct the phase space of the smoothed fluorescent trajectories as quantified from microscopy imaging data. The phase was reconstructed by removing the mean and linear trend from the smoothed data set and then plotting the detrended data using a delay of 8 against itself at delay 0.

```
1 %Key of Cell A matrix columns
2 %1=time
3 %2=original data
4 %3=smoothed data
5 %4=smoothed data with mean removed
6 %5=smoothed data with mean and trend removed
7 %6=phase space data on non-detrended data
8 %7=phase space data on detrended data
9
10 %Remove mean from smoothed data so as to make data revolve around 0
11 CellA=[CellA zeros(60,1)];
12 MeanA=mean(CellA(:,3));
13 for i=1:60
14 CellA(i,4)=CellA(i,3)-MeanA;
15 end;
16
17 %Fit linear line to the data
18 xA = CellA(:,1);
19 yA = CellA(:,4);
20 pA = polyfit(xA,yA,1);
21 x1A = CellA(:,1);
22 y1A = polyval(pA,x1A);
23
24 %Remove linear trend
25 CellA=[CellA zeros(60,1)];
26 for i=1:60
27 CellA(i,5)= CellA(i,4)-y1A(i,1);
28 end;
29
30 %Reconstruct phase space for detrended data
31 CellA=[CellA zeros(60,1)];
32 delay=8;
33 for i=1:(60-delay)
```

```
34     CellA(i,7)=CellA(i+delay,5);
35 end;
36 h=figure;
37 plot(CellA(1:(60-delay),5),CellA(1:(60-delay),7), 'LineWidth',1.5)
```

References

Agapakis, C. M. & Silver, P. A. (2009). Synthetic biology: exploring and exploiting genetic modularity through the design of novel biological networks. *Molecular bioSystems*, 5, 704–713. DOI: 10.1039/b901484e.

Agilent Technologies UK Ltd. (n.d). *X110-gold ultracompetent cells*. Retrieved from <http://www.genomics.agilent.com/en/product.jsp?cid=AG-PT-118&tabId=AG-PR-1098&requestid=484659> [Accessed 17 January 2014].

Aitcheson, N., Talbot, S., Shapiro, J., Hughes, K., Adkin, C., Butt, T., ... Rudenko, G. (2005). VSG switching in *Trypanosoma brucei*: antigenic variation analysed using RNAi in the absence of immune selection. *Mol. Microbiol.* 57(6), 1608–1622.

Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., ... Walter, P. (2010). *Essential cell biology* (3rd ed.). Garland Science.

Alibu, V., Storm, L., Haile, S., Clayton, C., & Horn, D. (2005). A doubly inducible system for RNA interference and rapid RNAi plasmid construction in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 139(1), 75–82.

Allgaier, M., Uphoff, H., Felske, A., & Wagner-Dobler, I. (2003). Aerobic anoxygenic photosynthesis in Roseobacter clade bacteria from diverse marine habitats. *Appl. Environ. Microbiol.* 69(9), 5051–5059.

Alper, J. (2009). Biotech in the basement. *Nat. Biotechnol.* 27(12), 1077–1078.

Alsfeld, S. & Horn, D. (2008). Single-locus targeting constructs for reliable regulated RNAi and transgene expression in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 161(1), 76–79.

Alsfeld, S. & Horn, D. (2012). Cell cycle regulated control of VSG expression site silencing by histones and histone chaperones ASF1A and CAF 1b in *Trypanosoma brucei*. *Nucleic Acids Res.* 40(20), 10150–10160.

Alsfeld, S., Kawahara, T., Glover, L., & Horn, D. (2005). Tagging a *T. brucei* rRNA locus improves stable transfection efficiency and circumvents inducible expression position effects. *Mol. Biochem. Parasitol.* 144(2), 142–148.

Andrianantoandro, E., Basu, S., Karig, D. K., & Weiss, R. (2006). Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol.* 2.

Anesiadis, N., Cluett, W. R., & Mahadevan, R. (2008). Dynamic metabolic engineering for increasing bioprocess productivity. *Metab. Eng.* 10(5), 255–266.

Anthony, R. (2001). Transformation of competent bacterial cells: electroporation. In *eLS*. Chichester: John Wiley and Sons, Ltd.

ApE. (2012). *Ape - a plasmid editor*. Retrieved from <http://biologylabs.utah.edu/jorgensen/wayned/ape/>.

Archer, S. K. (2009a). Standard culture medium allows clonal dilution of *Trypanosoma brucei* procyclic cells after auto conditioning. *Mol. Biochem. Parasitol.* 164(1), 100–103.

Archer, S. K. (2009b). Standard culture medium allows clonal dilution of *Trypanosoma brucei* procyclic cells after auto-conditioning. *Mol. Biochem. Parasitol.* 164(1), 100–103.

Arkin, A. P. (2013). A wise consistency: engineering biology for conformity, reliability, predictability. *Curr Opin Chem Biol*, 17(6), 893–901.

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.

Aubel, D. & Fussenegger, M. (2010). Mammalian synthetic biology—from tools to therapies. *Biosays*, 32(4), 332–345.

Aufderheide, K. J. (2008). An overview of techniques for immobilizing and viewing living cells. *Micron*, 39(2), 71–76.

Auslander, S., Auslander, D., Muller, M., Wieland, M., & Fussenegger, M. (2012). Programmable single cell mammalian biocomputers. *Nature*, 487(7405), 123–127.

Auslander, S. & Fussenegger, M. (2013). From gene switches to mammalian designer cells: present and future prospects. *Trends Biotechnol*. 31(3), 155–168.

Azami, H., Bozorgtabar, B., & Shiroie, M. (2011). Automatic signal segmentation using the fractal dimension and weighted moving average filter. *IJECS*, 11(6), 8–15.

Aziz, N. A., Anguelova, G. V., Marinus, J., Lammers, G. J., & Roos, R. A. C. (2010). Sleep and circadian rhythm alterations correlate with depression and cognitive impairment in huntington's disease. *Parkinsonism & related disorders*, 16(5), 345–350.

Baby, V., Matteau, D., Knight, T. F., & Rodrigue, S. (2013). Complete Genome Sequence of the *Mesoplasma florum* W37 Strain. *Genome Announc*, 1(6).

Bacchus, W., Aubel, D., & Fussenegger, M. (2013). Biomedically relevant circuit-design strategies in mammalian synthetic biology. *Mol. Syst. Biol.* 9, 691.

Bachmair, A., Finley, D., & Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino-terminal residue. *Science*, 234(4773), 179–186.

Bailey, J. E. (1998). Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities. *Biotechnology Progress*, 14(1), 8–20.

Bakker, B. M., Krauth Siegel, R. L., Clayton, C., Matthews, K., Girolami, M., Westerhoff, H. V., ... Barrett, M. P. (2010). The silicon trypanosome. *Parasitology*, 137(9), 1333–1341.

Bakker, B. M., Michels, P. A., Opperdoes, F. R., & Westerhoff, H. V. (1997). Glycolysis in blood-stream form *Trypanosoma brucei* can be understood in terms of the kinetics of the glycolytic enzymes. *J. Biol. Chem.* 272(6), 3207–3215.

Barkai, N. & Leibler, S. (2000). Biological rhythms: circadian clocks limited by noise. *Nature*, 403(6767), 267–268. DOI: 10.1038/35002258.

Barnes, C. P., Silk, D., Sheng, X., & Stumpf, M. P. (2011a). Bayesian design of synthetic biological systems. *PNAS*, 108(37), 15190–15195.

Barnes, C. P., Silk, D., & Stumpf, M. P. (2011b). Bayesian design strategies for synthetic biology. *Interface Focus*, 1(6), 895–908.

Barnes, D. K., Galgani, F., Thompson, R. C., & Barlaz, M. (2009). Accumulation and fragmentation of plastic debris in global environments. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 364(1526), 1985–1998.

Barrio, M., Burrage, K., Leier, A., & Tian, T. (2006). Oscillatory regulation of *hes1*: discrete stochastic delay modelling and simulation. *PLoS Comput Biol*, 2(9), e117.

Basile, G. & Peticca, M. (2009). Recombinant protein expression in *leishmania tarentolae*. *Mol Biotechnol*. 43(3), 273–278.

Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H., & Weiss, R. (2005). A synthetic multicellular system for programmed pattern formation. *Nature*, 434(7037), 1130–1134.

Basu, S., Mehreja, R., Thibierge, S., Chen, M. T., & Weiss, R. (2004). Spatiotemporal control of gene expression with pulse generating networks. *PNAS*, 101(17), 6355–6360.

Bates, R., Blyuss, O., & Zaikin, A. (2014). Stochastic resonance in an intracellular genetic perceptron. *Phys. Rev. E*, 89, 032716.

Batista, M., Marchini, F. K., Celedon, P. A., Fragoso, S. P., Probst, C. M., Preti, H., ... Krieger, M. A. (2010). A high throughput cloning system for reverse genetics in *Trypanosoma cruzi*. *BMC Microbiol.* 10, 259.

Batt, G., Yordanov, B., Weiss, R., & Belta, C. (2007). Robustness analysis and tuning of synthetic gene networks. *Bioinformatics*, 23(18), 2415–2422.

Bayer, T. S. & Smolke, C. D. (2005). Programmable ligand controlled riboregulators of eukaryotic gene expression. *Nat. Biotechnol.* 23(3), 337–343.

BBSRC. (2012). *20m investment for U.K. synthetic biology research*. Retrieved from <http://www.bbsrc.ac.uk/news/research-technologies/2012/121109-pr-investment-uk-synthetic-biology-research.aspx> [Accessed 29 April 2014].

BBSRC. (2014). *Uk establishes three new research centre*. Retrieved from <http://www.bbsrc.ac.uk/news/research-technologies/2014/140130-pr-new-synthetic-biology-research-centres.aspx> [Accessed 29 April 2014].

Becskei, A., Seraphin, B., & Serrano, L. (2001). Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J.* 20(10), 2528–2535.

Becskei, A. & Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, 405(6786), 590–593.

Beerli, R. R., Schopfer, U., Dreier, B., & Barbas, C. F. (2000). Chemically regulated zinc finger transcription factors. *J. Biol. Chem.* 275(42), 32617–32627.

Berens, C., Altschmied, L., & Hillen, W. (1992). The role of the N terminus in Tet repressor for tet operator binding determined by a mutational analysis. *J. Biol. Chem.* 267(3), 1945–1952.

Berla, B. M., Saha, R., Immethun, C. M., Maranas, C. D., Moon, T. S., & Pakrasi, H. B. (2013). Synthetic biology of cyanobacteria: unique challenges and opportunities. *Front Microbiol*, 4, 246.

Bernot, G., Comet, J., Richard, A., Chaves, M., Gouze, J., & Dayan, F. (2013). Modeling and Analysis of Gene Regulatory Networks. In F. Cazals and P. Kornprobst (Eds.) *Modeling in Computational Biology and Biomedicine*. Berlin Heidelberg:Springer-Verlag.

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., ... El-Sayed, N. M. (2005). The genome of the African trypanosome *Trypanosoma brucei*. *Science*, 309(5733), 416–422.

Beucher, S. & Lantuejoul, C. (1979, September). *Use of Watersheds in Contour Detection*. Paper presented at the International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation, Rennes, France.

Biliouris, K., Daoutidis, P., & Kaznessis, Y. N. (2011). Stochastic simulations of the tetracycline operon. *BMC Syst Biol*, 5, 9.

Birch, J. (2005). *Upstream mammalian cell processing: challenges and prospects*. Presentation at BioProcess International, Berlin, Germany.

Bloomfield, P. (2004). *Fourier analysis of time series: an introduction*. John Wiley and Sons.

Blumenthal, T. (2004). Operons in eukaryotes. *Brief Funct Genomic Proteomic*, 3(3), 199–211.

Board, T. S. (2012). *A synthetic biology roadmap for the U.K.* London: Technology Strategy Board. Retrieved from <http://www.rcuk.ac.uk/RCUK-prod/assets/documents/publications/SyntheticBiologyRoadmap.pdf> [Accessed 29 April 2014].

Borg, Y., Ullner, E., Alaga, A., Alsaedi, A., Nesbeth, D., & Zaikin, A. (2014). Complex and unexpected dynamics in simple genetic regulatory networks. *IJMP B*, 28(14).

Bouvier, L. A., Camara, M., Canepa, G. E., Miranda, M. R., & Pereira, C. A. (2013). Plasmid vectors and molecular building blocks for the development of genetic manipulation tools for *Trypanosoma cruzi*. *PLoS ONE*, 8(10), e80217.

Boyer, M. E., Stapleton, J. A., Kuchenreuther, J. M., Wang, C. W., & Swartz, J. R. (2008). Cell free synthesis and maturation of [FeFe] hydrogenases. *Biotechnol. Bioeng.* 99(1), 59–67.

Boyle, P. M., Burrill, D. R., Inniss, M. C., Agapakis, C. M., Deardon, A., Dewerd, J. G., ... Silver, P. A. (2012). A BioBrick compatible strategy for genetic modification of plants. *J. Biol. Eng.* 6(1), 8.

Bratsun, D., Volfson, D., Tsimring, L. S., & Hasty, J. (2005). Delay induced stochastic oscillations in gene regulation. *PNAS*, 102(41), 14593–14598.

Breitling, R. (2013). The LEXSY platform for recombinant protein expression. in a. de almeida et al.(ed.) *Farm animal proteomics 2013*. (pp. 45-48). Wageningen Academic Publishers.

Breitling, R., Klingner, S., Callewaert, N., Pietrucha, R., Geyer, A., Ehrlich, G., ... Alexandrov, K. (2002). Non pathogenic trypanosomatid protozoa as a platform for protein research and production. *Protein Expr. Purif.* 25(2), 209–218.

Brenner, K., Karig, D. K., Weiss, R., & Arnold, F. H. (2007). Engineered bidirectional communication mediates a consensus in a microbial biofilm consortium. *PNAS*, 104(44), 17300–17304.

Brinkhoff, T., Giebel, H. A., & Simon, M. (2008). Diversity, ecology, and genomics of the Roseobacter clade: a short overview. *Arch. Microbiol.* 189(6), 531–539.

Broadhead, R., Dawe, H. R., Farr, H., Griffiths, S., Hart, S. R., Portman, N., ... Gull, K. (2006). Flagellar motility is required for the viability of the bloodstream trypanosome. *Nature*, 440(7081), 224–227.

Bruhn, J. B., Gram, L., & Belas, R. (2007). Production of antibacterial compounds and biofilm formation by Roseobacter species are influenced by culture conditions. *Appl. Environ. Microbiol.* 73(2), 442–450.

Bryksina, A., Browna, A., Bakshb, M., Finn, M., & Barker, T. (2014). Learning from nature ? Novel synthetic biology approaches for biomaterial design. *Acta Biomater.* 10(4), 1761–1769.

Buchan, A., Collier, L. S., Neidle, E. L., & Moran, M. A. (2000). Key aromatic-ring-cleaving enzyme, protocatechuate 3,4-dioxygenase, in the ecologically important marine Roseobacter lineage. *Appl. Environ. Microbiol.* 66(11), 4662–4672.

Buchan, A., Gonzalez, J. M., & Moran, M. A. (2005). Overview of the marine Roseobacter lineage. *Appl. Environ. Microbiol.* 71(10), 5665–5677.

Buddruhs, N., Chertkov, O., Petersen, J., Fiebig, A., Chen, A., Pati, A., ... Klenk, H. P. (2013). Complete genome sequence of the marine methylhalide oxidizing *Leisingera* methylohalidivorans type strain (DSM 14336(T)), a representative of the Roseobacter clade. *Stand Genomic Sci*, 9(1), 128–141.

Byrne, J., Heidelberger, R., & Waxham, M. (2014). *From molecules to networks: an introduction to cellular and molecular neuroscience*. Academic Press.

Cai, X. & Wang, X. (2007). Stochastic modeling and simulation of gene networks, a review of the state of the art research on stochastic simulations. *IEEE Signal Process. Mag*, 24(1), 27–36.

Calvert, J., Schyfter, P., Elfick, A., & Endy, D. (2014). *Synthetic aesthetics, investigating synthetic biology's designs on nature*. MIT Press.

Calvo-Alvarez, E., Alvarez-Velilla, R., Fernandez-Prada, C., Balana-Fouce, R., & Reguera, R. M. (2015). Trypanosomatids see the light: recent advances in bioimaging research. *Drug Discov. Today*, 20(1), 114–121.

Cameron, D. E., Bashor, C. J., & Collins, J. J. (2014). A brief history of synthetic biology. *Nat. Rev. Microbiol.* 12(5), 381–390.

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

Carlson, R. (2014). *Time for new dna synthesis and sequencing cost curves*. Retrieved from <http://www.synthesis.cc/> [Accessed 01 March 2015].

Carothers, J. M., Goler, J. A., Juminaga, D., & Keasling, J. D. (2011). Model driven engineering of RNA devices to quantitatively program gene expression. *Science*, 334(6063), 1716–1719.

Carrington, M., Miller, N., Blum, M., Roditi, I., Wiley, D., & Turner, M. (1991). Variant specific glycoprotein of *Trypanosoma brucei* consists of two domains each having an independently conserved pattern of cysteine residues. *J. Mol. Biol.* 221(3), 823–835.

Carruthers, V. B. & Cross, G. A. (1992). High efficiency clonal growth of bloodstream and insect form *Trypanosoma brucei* on agarose plates. *PNAS*, 89(18), 8818–8821.

Chamakh-Ayari, R., Bras-Goncalves, R., Bahi-Jaber, N., Petididier, E., Markikou-Ouni, W., Aoun, K., ... Meddeb-Garnaoui, A. (2014). In vitro evaluation of a soluble *Leishmania* promastigote surface antigen as a potential vaccine candidate against human leishmaniasis. *PLoS ONE*, 9(5), e92708.

Chandran, D., Copeland, W., Sleight, S., & Sauro, H. (2008). Mathematical modeling and synthetic biology. *Drug Discov Today Dis Models*, 5(4), 299–309.

Chang, Y. C., Lin, C. L., & Jennawasin, T. (2013). Design of synthetic genetic oscillators using evolutionary optimization. *Evol. Bioinform. Online*, 9, 137–150.

Chen, B. S., Chang, C. H., & Lee, H. C. (2009). Robust synthetic biology design: stochastic game theory approach. *Bioinformatics*, 25(14), 1822–1830.

Chen, B. S. & Chen, P. W. (2010). GA based Design Algorithms for the Robust Synthetic Genetic Oscillators with Prescribed Amplitude, Period and Phase. *Gene Regul Syst Bio*, 4, 35–52.

Chen, B. S. & Hsu, C. Y. (2012). Robust synchronization control scheme of a population of nonlinear stochastic synthetic genetic oscillators under intrinsic and extrinsic molecular noise via quorum sensing. *BMC Syst Biol*, 6, 136.

Chen, L. & Aihara, K. (2002). A model of periodic oscillation for genetic regulatory systems. *IEEE T Circuits-I*, 49(10), 1429–1436.

Chen, M. T. & Weiss, R. (2005). Artificial cell cell communication in yeast *Saccharomyces cerevisiae* using signaling elements from *Arabidopsis thaliana*. *Nat. Biotechnol.* 23(12), 1551–1555.

Christie-Oleza, J. A., Pina-Villalonga, J. M., Bosch, R., Nogales, B., & Armengaud, J. (2012). Comparative proteogenomics of twelve *Roseobacter* exoproteomes reveals different adaptive strategies among these marine bacteria. *Mol. Cell Proteomics*, 11(2), M111.013110.

Church, G. M., Elowitz, M. B., Smolke, C. D., Voigt, C. A., & Weiss, R. (2014). Realizing the potential of synthetic biology. *Nat. Rev. Mol. Cell Biol.* 15(4), 289–294.

Cinquemani, E., Milius-Argeitis, A., Summers, S., & Lygeros, J. (2008). Stochastic dynamics of genetic networks: modelling and parameter identification. *Bioinformatics*, 24(23), 2748–2754.

Clayton, C. & Shapira, M. (2007). Posttranscriptional regulation of gene expression in trypanosomes and leishmanias. *Mol Biochem Parasit*, 156(2), 93–101.

Clayton, C. (1999). Genetic manipulation of kinetoplastida. *Parasitol Today*, 15(9), 372–378.

Corish, P. & Tyler-Smith, C. (1999). Attenuation of green fluorescent protein half-life in mammalian cells. *Protein Eng.* 12(12), 1035–1040.

Cross, G. A. (1975). Identification, purification and properties of clone specific glycoprotein antigens constituting the surface coat of *Trypanosoma brucei*. *Parasitology*, 71(3), 393–417.

Dacher, M., Morales, M. A., Pescher, P., Leclercq, O., Rachidi, N., Prina, E., ... Spath, G. F. (2014). Probing druggability and biological function of essential proteins in *Leishmania* combining facilitated null mutant and plasmid shuffle analyses. *Mol. Microbiol.* 93(1), 146–166.

D'Alvise, P. (2013). *Aquaculture application and ecophysiology of marine bacteria from the Roseobacter clade* (Doctoral dissertation, Technical University of Denmark). Retrieved from http://orbit.dtu.dk/files/54035617/phd.thesis_pauldalvise.pdf.

Danchin, A. (2012). Scaling up synthetic biology: Do not forget the chassis. *FEBS Lett.* 586(15), 2129–2137.

Daniels, J. P., Gull, K., & Wickstead, B. (2010). Cell biology of the trypanosome genome. *Microbiol. Mol. Biol. Rev.* 74(4), 552–569.

Danino, T., Mondragon-Palomino, O., Tsimring, L., & Hasty, J. (2010). A synchronized quorum of genetic clocks. *Nature*, 463(7279), 326–330.

Dantuma, N. P., Lindsten, K., Glas, R., Jellne, M., & Masucci, M. G. (2000). Short lived green fluorescent proteins for quantifying ubiquitin/proteasome dependent proteolysis in living cells. *Nat. Biotechnol.* 18(5), 538–543.

Das, S. (2009). *Handbook of research on computational methodologies in gene regulatory networks*. IGI Global.

Dash, H. R., Mangwani, N., Chakraborty, J., Kumari, S., & Das, S. (2013). Marine bacteria, potential candidates for enhanced bioremediation. *Appl. Microbiol. Biotechnol.* 97(2), 561–571.

Dasika, M. S. & Maranas, C. D. (2008). OptCircuit: an optimization based method for computational design of genetic circuits. *BMC Syst Biol*, 2, 24.

David, F. & Siewers, V. (2015). Advances in yeast genome engineering. *FEMS Yeast Research*, 15(1), 1–14.

De Boor, C. (2001). *A practical guide to splines*. Springer.

de Jong, H. (2003). Mathematical modeling of genetic regulatory networks. *Projet Comore, Inria*. Retrieved from http://www-sop.inria.fr/comore/main_fr.html.

de Jong, H. (2002). Modeling and simulation of genetic regulatory systems: a literature review. *J Comput Biol*, 9(1), 67–103.

DeAngelis, K. M., Firestone, M. K., & Lindow, S. E. (2007). Sensitive whole cell biosensor suitable for detecting a variety of N acyl homoserine lactones in intact rhizosphere microbial communities. *Appl. Environ. Microbiol.* 73(11), 3724–3727.

Department for Business, Innovation and Skills. (2013). *Over 60 million for synthetic biology*. Retrieved from <https://www.gov.uk/government/news/over-60-million-for-synthetic-biology> [Accessed 29 April 2014].

Dortay, H. & Mueller Roeber, B. (2010). A highly efficient pipeline for protein expression in Leishmania tarentolae using infrared fluorescence protein as marker. *Microb. Cell Fact.* 9, 29.

Dragosits, M., Nicklas, D., & Tagkopoulos, I. (2012). A synthetic biology approach to self-regulatory recombinant protein production in Escherichia coli. *J. Biol. Eng.* 6(1), 2.

Drozdz, M. & Clayton, C. (1999). Structure of a regulatory 3' untranslated region from trypanosoma brucei. *RNA*, 5(12), 1632–1644.

Drubin, D. A., Way, J. C., & Silver, P. A. (2007). Designing biological systems. *Genes Dev.* 21(3), 242–254.

Dueber, J. E., Yeh, B. J., Chak, K., & Lim, W. A. (2003). Reprogramming control of an allosteric signaling switch through modular recombination. *Science*, 301(5641), 1904–1908.

El Samad, H., Khammash, M., Petzold, L., & Gillespie, D. (2005). Stochastic modelling of gene regulatory networks. *Int J Robust Nonlin.* 15(15), 691–711.

Elifantz, H., Horn, G., Ayon, M., Cohen, Y., & Minz, D. (2013). Rhodobacteraceae are the key members of the microbial community of the initial biofilm formed in Eastern Mediterranean coastal seawater. *FEMS Microbiol. Ecol.* 85(2), 348–357.

Elowitz, M. B. & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767), 335–338.

Elowitz, M. B., Levine, A. J., Siggia, E. D., & Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science*, 297(5584), 1183–1186.

Endler, L., Rodriguez, N., Juty, N., Chelliah, V., Laibe, C., Li, C., & Le Novere, N. (2009). Designing and encoding models for synthetic biology. *J R Soc Interface, 6 Suppl 4*, S405–417.

Endy, D. (2007). *2003 synthetic biology study*. Retrieved from <http://dspace.mit.edu/handle/1721.1/38455> [Accessed 29 April 2014].

ERASynBio. (2014). *Next steps for european synthetic biology, a strategic vision from erasynbio*. :ERASynBio. Retrieved from <<http://www.bbsrc.ac.uk/web/FILES/Publications/1404-erasynbio-strategic-vision.pdf>> [Accessed 30 April 2014].

Erickson, B., Singh, R., & Winters, P. (2011). Synthetic biology: regulating industry uses of new biotechnologies. *Science*, 333(6047), 1254–1256.

Erneux, T. (2009). *Applied delay differential equations*. Springer Science and Business Media.

ESPRC. (2014). *EPSRC support by research area in synthetic biology*. Retrieved from <<http://gow.epsrc.ac.uk/NGBOChooseTTS.aspx?Mode=ResearchArea&ItemDesc=Synthetic+Biology>> [Accessed 29 April 2014].

ESPRC. (n.d.). *Synthetic biology*. Retrieved from <<http://www.epsrc.ac.uk/research/ourportfolio/researchareas/Pages/synthbio.aspx>> [Accessed 29 April 2014].

Evan, G. I., Lewis, G. K., Ramsay, G., & Bishop, J. M. (1985). Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol. Cell. Biol.* 5(12), 3610–3616.

Feng, X. J., Hooshangi, S., Chen, D., Li, G., Weiss, R., & Rabitz, H. (2004). Optimizing genetic circuits by global sensitivity analysis. *Biophys. J.* 87(4), 2195–2202.

Filippi, S., Barnes, C. P., Cornebise, J., & Stumpf, M. P. (2013). On optimality of kernels for approximate Bayesian computation using sequential Monte Carlo. *Stat Appl Genet Mol Biol*, 12(1), 87–107.

Floury, J., Jeanson, S., Madec, M. N., & Lortal, S. (2013). Porosity of *Lactococcus lactis* subsp. *lactis* LD61 colonies immobilised in model cheese. *Int. J. Food Microbiol.* 163(2-3), 64–70.

Forster, A. C. & Church, G. M. (2006). Towards synthesis of a minimal cell. *Mol. Syst. Biol.* 2, 45.

Francois, P. & Hakim, V. (2004). Design of genetic networks with specified functions by evolution in silico. *PNAS*, 101(2), 580–585.

Fraser, A. M. & Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Phys. Rev. A*, 33(2), 1134–1140.

Freire, E. R., Malvezzi, A. M., Vashisht, A. A., Zuberek, J., Saada, E. A., Langousis, G., ... Campbell, D. A. (2014). Trypanosoma brucei translation initiation factor homolog EIF4E6 forms a tripartite cytosolic complex with EIF4G5 and a capping enzyme homolog. *Eukaryotic Cell*, 13(7), 896–908.

Freitag, A. & Pfeffer, M. J. (2013). Process, not product: investigating recommendations for improving citizen science "success". *PLoS ONE*, 8(5), e64079.

Friedland, A. E., Lu, T. K., Wang, X., Shi, D., Church, G., & Collins, J. J. (2009). Synthetic gene networks that count. *Science*, 324(5931), 1199–1202.

Fritzsche, C., Sitz, M., Weiland, N., Breitling, R., & Pohl, H. (2007). Characterization of the growth behavior of leishmania tarentolae: a new expression system for recombinant proteins. *J Basic Microb*, 47(5), 384–393. DOI: 10.1002/jobm.200710111.

Fung, E., Wong, W. W., Suen, J. K., Bulter, T., Lee, S., & Liao, J. C. (2005). A synthetic gene metabolic oscillator. *Nature*, 435(7038), 118–122. DOI: 10.1038/nature03508.

Gao, H., Zhuo, Y., Ashforth, E., & Zhang, L. (2010a). Engineering of a genome reduced host: practical application of synthetic biology in the overproduction of desired secondary metabolites. *Protein Cell*, 1(7), 621–626.

Gao, S., Hartman, J. L., Carter, J. L., Hessner, M. J., & Wang, X. (2010b). Global analysis of phase locking in gene expression during cell cycle: the potential in network modeling. *BMC Syst. Biol.* 4, 167.

Garcia-Ojalvo, J., Elowitz, M. B., & Strogatz, S. H. (2004). Modeling a synthetic multicellular clock: repressilators coupled by quorum sensing. *PNAS*, 101(30), 10955–10960.

Gardner, T. S., Cantor, C. R., & Collins, J. J. (2000). Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, 403(6767), 339–342. DOI: 10.1038/35002131.

Gibson, D. G., Young, L., Chuang, R. Y., Craig Venter, J., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*, 6(5), 343–345.

Gibson, M. & Bruck, J. (2000). Efficient exact stochastic simulation of chemical systems with many species and many channels. *J Phys Chem A*, 104(9), 1876–1889.

Giebel, H. A., Kalhoefer, D., Lemke, A., Thole, S., Gahl-Janssen, R., Simon, M., & Brinkhoff, T. (2011). Distribution of Roseobacter RCA and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. *ISME J*, 5(1), 8–19.

Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J Comput Phys*, 22(4), 403–434.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* 81(25), 2340–2361.

Gillespie, D. T. (2000). The chemical langevin equation. *J. Phys. Chem.* 113(1), 297–306.

Gillespie, D. T. (2001). Approximate accelerated stochastic simulation of chemically reacting systems. *J. Phys. Chem.* 115(4), 1716–1733.

Gimpel, J. A., Specht, E. A., Georgianna, D. R., & Mayfield, S. P. (2013). Advances in microalgae engineering and synthetic biology applications for biofuel production. *Curr Opin Chem Biol*, 17(3), 489–495.

Glass, J. I., Assad Garcia, N., Alperovich, N., Yooseph, S., Lewis, M. R., Maruf, M., ... Craig Venter, J. (2006). Essential genes of a minimal bacterium. *PNAS*, 103(2), 425–430.

Goldbeter, A. (2002). Computational approaches to cellular rhythms. *Nature*, 420(6912), 238–245.

Gonda, D. K., Bachmair, A., Wunning, I., Tobias, J. W., Lane, W. S., & Varshavsky, A. (1989). Universality and structure of the N-end rule. *J. Biol. Chem.* 264(28), 16700–16712.

Gonzalez, J. M., Kiene, R. P., & Moran, M. A. (1999). Transformation of sulfur compounds by an abundant lineage of marine bacteria in the alpha subclass of the class Proteobacteria. *Appl. Environ. Microbiol.* 65(9), 3810–3819.

Gonzalez, R. & Woods, R. (2002). *Digital image processing* (2nd edition). Beijing: Publishing House of Electronics Industry.

Gonze, D. (2013). Modeling the effect of cell division on genetic oscillators. *J. Theor. Biol.* 325, 22–33.

Gonze, D., Bernard, S., Waltermann, C., Kramer, A., & Herz, H. (2005). Spontaneous synchronization of coupled circadian oscillators. *Biophys. J.* 89(1), 120–129.

Goodwin, B. C. (1963). *Temporal organization in cells; a dynamic theory of cellular control processes*. Retrieved from <http://www.archive.org/details/temporalorganiza00good>. London: Academic Press.

Gossen, M. & Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline responsive promoters. *PNAS*, 89(12), 5547–5551.

Gramelsberger, G. (2013). The simulation approach in synthetic biology. *Stud Hist Philos Biol Biomed Sci*, 44(2), 150–157.

Grant, C. (1999). Definitions of stability. *Theory of Ordinary Differential Equations*, [pdf] Brigham Young University. Retrieved from <http://www.math.byu.edu/~grant/courses/m634/f99/lec20.pdf> [Accessed 19 February 2012].

Gray, K. M., Passador, L., Iglesias, B. H., & Greenberg, E. P. (1994). Interchangeability and specificity of components from the quorum sensing regulatory systems of *Vibrio fischeri* and *Pseudomonas aeruginosa*. *J. Bacteriol.* 176(10), 3076–3080.

Griffith, J. S. (1968). Mathematics of cellular control processes. I. Negative feedback to one gene. *J. Theor. Biol.* 20(2), 202–208.

Gruszynski, A. E., van Deursen, F. J., Albareda, M. C., Best, A., Chaudhary, K., Cliffe, L. J., ... Bangs, J. D. (2006). Regulation of surface coat exchange by differentiating African trypanosomes. *Mol. Biochem. Parasitol.* 147(2), 211–223.

Gu, P., Yang, F., Su, T., Wang, Q., Liang, Q., & Qi, Q. (2015). A rapid and reliable strategy for chromosomal integration of gene(s) with multiple copies. *Sci Rep*, 5, 9684.

Gu, X., Reid, D., Higham, D. J., & Gilbert, D. (2013). Mathematical modelling of polyamine metabolism in bloodstream-form *Trypanosoma brucei*: an application to drug target identification. *PLoS ONE*, 8(1), e53734.

Guantes, R. & Poyatos, J. F. (2006). Dynamical principles of two component genetic oscillators. *PLoS Comput Biol*, 2(3), e30. DOI: 10.1371/journal.pcbi.0020030.

Guido, N. J., Wang, X., Adalsteinsson, D., McMillen, D., Hasty, J., Cantor, C. R., ... Collins, J. J. (2006). A bottom up approach to gene regulation. *Nature*, 439(7078), 856–860.

Gulati, S., Rouilly, V., Niu, X., Chappell, J., Kitney, R. I., Edel, J. B., ... deMello, A. J. (2009). Opportunities for microfluidic technologies in synthetic biology. *J R Soc Interface*, 6 Suppl 4, 493–506.

Gupta, S., Bram, E. E., & Weiss, R. (2013). Genetically programmable pathogen sense and destroy. *ACS Synth Biol*, 2(12), 715–723.

Hahnke, S., Brock, N. L., Zell, C., Simon, M., Dickschat, J. S., & Brinkhoff, T. (2013). Physiological diversity of Roseobacter clade bacteria cooccurring during a phytoplankton bloom in the North Sea. *Syst. Appl. Microbiol.* 36(1), 39–48.

Haldimann, A. & Wanner, B. L. (2001). Conditional replication, integration, excision, and retrieval plasmid host systems for gene structure function studies of bacteria. *J. Bacteriol.* 183(21), 6384–6393.

Hamamoto, S., Okuma, T., Yamamoto, A., Kageyama, K., Ueki, A., Matsuoka, T., & Miki, Y. (2015). Combination radiofrequency ablation and local injection of the immunostimulant bacillus Calmette-Guerin induces antitumor immunity in the lung and at a distant VX2 tumor in a rabbit model. *J Vasc Interv Radiol*, 26(2), 271–278.

Hand, A. J., Sun, T., Barber, D. C., Hose, D. R., & MacNeil, S. (2009). Automated tracking of migrating cells in phase-contrast video microscopy sequences using image registration. *J Microsc*, 234(1), 62–79.

Hardman-Mountford, N., Polimene, L., Hirata, T., Brewin, R., & Aiken, J. (2013). Impacts of light shading and nutrient enrichment geoengineering approaches on the productivity of a stratified, oligotrophic ocean ecosystem. *J. R. Soc. Interface*, 10(89).

Hasty, J., Dolnik, M., Rottschufer, V., & Collins, J. J. (2002a). Synthetic gene network for entraining and amplifying cellular oscillations. *Phys Rev Lett*, 88(14), pages. DOI: 10.1103/PhysRevLett.88.148101.

Hasty, J., McMillen, D., & Collins, J. J. (2002b). Engineered gene circuits. *Nature*, 420(6912), 224–230. DOI: 10.1038/nature01257.

Haupts, U., Maiti, S., Schwille, P., & Webb, W. W. (1998). Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy. *PNAS*, 95(23), 13573–13578.

Hayashi, K., Morooka, N., Yamamoto, Y., Fujita, K., Isono, K., Choi, S., ... Horiuchi, T. (2006). Highly accurate genome sequences of *Escherichia coli* K12 strains MG1655 and W3110. *Mol. Syst. Biol.* 2, 2006.0007.

Hayden, E. C. (2014). Synthetic biology firms shift focus. *Nature*, 505(7485), 598.

Heddergott, N., Kruger, T., Babu, S. B., Wei, A., Stellamanns, E., Uppaluri, S., ... Engstler, M. (2012). Trypanosome motion represents an adaptation to the crowded environment of the vertebrate bloodstream. *PLoS Pathog*, 8(11), e1003023.

Heidorn, T., Camsund, D., Huang, H., Lindberg, P., Oliveira, P., Stensjo, K., & Lindblad, P. (2011). Chapter twenty-four - synthetic biology in cyanobacteria: engineering and analyzing novel functions. In C. Voigt (Ed.), *Synthetic biology, Part A* (Vol. 497, pp. 539–579). Methods in Enzymology. Academic Press.

Heinemann, M. & Panke, S. (2006). Synthetic biology, putting engineering into biology. *Bioinformatics*, 22(22), 2790–2799.

Hidalgo-Ruz, V. & Thiel, M. (2013). Distribution and abundance of small plastic debris on beaches in the SE Pacific (Chile): a study supported by a citizen science project. *Mar. Environ. Res.* 87-88, 12–18.

Hillen, W. & Berens, C. (1994). Mechanisms underlying expression of Tn10 encoded tetracycline resistance. *Annu. Rev. Microbiol.* 48, 345–369.

Hillen, W., Gatz, C., Altschmied, L., Schollmeier, K., & Meier, I. (1983). Control of expression of the Tn10 encoded tetracycline resistance genes. Equilibrium and kinetic investigation of the regulatory reactions. *J. Mol. Biol.* 169(3), 707–721.

Hillen, W., Unger, B., & Klock, G. (1982). Analysis of tet operator-TET repressor complexes by thermal denaturation studies. *Nucleic Acids Res.* 10(19), 6085–6097.

Hirokawa, Y., Kawano, H., Tanaka-Masuda, K., Nakamura, N., Nakagawa, A., Ito, M., ... Ogasawara, N. (2013). Genetic manipulations restored the growth fitness of reduced-genome *Escherichia coli*. *J. Biosci. Bioeng.* 116(1), 52–58.

Hirumi, H. & Hirumi, K. (1984). Continuous cultivation of animal infective bloodstream forms of an East African *Trypanosoma congolense* stock. *Ann Trop Med Parasitol*, 78(3), 327–330.

Hochachka, W. M., Fink, D., Hutchinson, R. A., Sheldon, D., Wong, W. K., & Kelling, S. (2012). Data intensive science applied to broad scale citizen science. *Trends Ecol. Evol. (Amst.)* 27(2), 130–137.

Hodgman, C. E. & Jewett, M. C. (2012). Cell free synthetic biology: thinking outside the cell. *Metab. Eng.* 14(3), 261–269.

Holtz, W. J. & Keasling, J. D. (2010). Engineering static and dynamic control of synthetic pathways. *Cell*, 140(1), 19–23.

Hong, S. H., Hegde, M., Kim, J., Wang, X., Jayaraman, A., & Wood, T. K. (2012). Synthetic quorum sensing circuit to control consortial biofilm formation and dispersal in a microfluidic device. *Nat Commun*, 3, 613.

Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., ... Kummer, U. (2006). COPASI—a COmplex PAthway SImulator. *Bioinformatics*, 22(24), 3067–3074.

Horn, D. & McCulloch, R. (2010). Molecular mechanisms underlying the control of antigenic variation in African trypanosomes. *Curr. Opin. Microbiol.* 13(6), 700–705.

Hotz, H., Lorenz, P., Fischer, R., Krieger, S., & Clayton, C. (1995). Role of 3' untranslated regions in the regulation of hexose transporter mRNAs in *trypanosoma brucei*. *Mol Biochem Parasit.* 75(1), 1–14.

Hu, M. C. & Davidson, N. (1987). The inducible lac operator repressor system is functional in mammalian cells. *Cell*, 48(4), 555–566.

Huang, Z., Kaltenbrunner, S., Simkova, E., Stanek, D., Lukes, J., & Hashimi, H. (2014). The Dynamics of Mitochondrial RNA binding Protein Complex in *Trypanosoma brucei* and its Petite Mutant Under Optimized Immobilization Conditions. *Eukaryotic Cell*, 13(9), 1232–1240.

Hug, M., Carruthers, V. B., Hartmann, C., Sherman, D. S., Cross, G. A., & Clayton, C. (1993). A possible role for the 3'-untranslated region in developmental regulation in *trypanosoma brucei*. *Mol Biochem Parasit.* 61(1), 87–95.

Hutchinson, O. C., Picozzi, K., Jones, N. G., Mott, H., Sharma, R., Welburn, S. C., & Carrington, M. (2007). Variant surface glycoprotein gene repertoires in *trypanosoma brucei* have diverged to become strain specific. *BMC Genomics*, 8(1), 234.

Hutchinson, S., Alsford, S., Glover, L., & Horn, D. (2014). Interference among homologous sequences maintains vsg allelic exclusion in african trypanosomes. *In prep.*

Inverso, J. A., Uphoff, T. S., Johnson, S. C., Paulnock, D. M., & Mansfield, J. M. (2010). Biological variation among african trypanosomes: I. Clonal expression of virulence is not linked to the variant surface glycoprotein or the variant surface glycoprotein gene telomeric expression site. *DNA Cell Biol.* 29(5), 215–227.

IPCC. (2012). *Meeting report of the intergovernmental panel on climate change expert meeting on geoengineering* (O. Edenhofer, R. Pichs-Madruga, Y. Sokona, C. Field, V. Barros, T. Stocker, ... M. Mastrandrea, Eds.). pp. 99. Potsdam, Germany: IPCC Working Group III Technical Support Unit, Potsdam Institute for Climate Impact Research.

Irie, K., Kitagawa, K., Nagura, H., Imai, T., Shimomura, T., & Fujiyoshi, Y. (2010). Comparative study of the gating motif and C type inactivation in prokaryotic voltage gated sodium channels. *J. Biol. Chem.* 285(6), 3685–3694.

Isaacs, F. J., Dwyer, D. J., Ding, C., Pervouchine, D. D., Cantor, C. R., & Collins, J. J. (2004). Engineered riboregulators enable post transcriptional control of gene expression. *Nat. Biotechnol.* 22(7), 841–847.

Izhikevich, E. (2007). *Dynamical systems in neuroscience*. Computational neuroscience. MIT Press.

Jackson, A. P. (2014). Genome evolution in trypanosomatid parasites. *Parasitology*, 1–17.

Jain, R. & Babbar, S. B. (2006). Xanthan gum: an economical substitute for agar in plant tissue culture media. *Plant Cell Rep.* 25(2), 81–84.

Jaschke, P. R., Saer, R. G., Noll, S., & Beatty, J. T. (2011). Modification of the genome of Rhodobacter sphaeroides and construction of synthetic operons. *Meth. Enzymol.* 497, 519–538.

Jena Bioscience GmbH. (n.d.). *LEXSY Eukaryotic Protein Expression*. Retrieved from http://www.jenabioscience.com/images/b3e879b381/Lexsy_brochure_web.pdf [Accessed 29 May 2014].

Jena Bioscience GmbH. (n.d.). Lexsy eukaryotic protein expression in Leishmania tarentolae. Retrieved from http://www.jenabioscience.com/cms/en/1/browse/1838_eukaryotic_expression_system_lexsy.html [Accessed 23 May 2014].

Jensen, B. C., Ramasamy, G., Vasconcelos, E. J., Ingolia, N. T., Myler, P. J., & Parsons, M. (2014). Extensive stage-regulation of translation revealed by ribosome profiling of Trypanosoma brucei. *BMC Genomics*, 15, 911.

Jiang, H., Salehi, S. A., Riedel, M. D., & Parhi, K. K. (2013). Discrete-time signal processing with DNA. *ACS Synth Biol.* 2(5), 245–254.

Johnson, M. & Principe, J. (2005). Modeling and detection of limit-cycle oscillations using adaptable linear models. *Journal of Aircraft*, 42(6), 1575–1587.

Jordan, R. C., Gray, S. A., Howe, D. V., Brooks, W. R., & Ehrenfeld, J. G. (2011). Knowledge gain and behavioral change in citizen science programs. *Conserv. Biol.* 25(6), 1148–1154.

Jorgensen, E. D. & Grushkin, D. (2011). Engage with, don't fear, community labs. *Nat. Med.* 17(4), 411.

Jorgensen, M. L., Friis, N. A., Just, J., Madsen, P., Petersen, S. V., & Kristensen, P. (2014). Expression of single chain variable fragments fused with the Fc region of rabbit IgG in Leishmania tarentolae. *Microb. Cell Fact.* 13.

Kaern, M., Blake, W. J., & Collins, J. (2003). The engineering of gene regulatory networks. *Annu Rev Biochem*, 5(1), 179–206.

Kaern, M., Elston, T. C., Blake, W. J., & Collins, J. J. (2005). Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.* 6(6), 451–464. PMID: 15883588.

Kahl, L. J. & Endy, D. (2013). A survey of enabling technologies in synthetic biology. *J. Biol. Eng.* 7(1), 13.

Kain, S. R., Adams, M., Kondepudi, A., Yang, T. T., Ward, W. W., & Kitts, P. (1995). Green fluorescent protein as a reporter of gene expression and protein localization. *BioTechniques*, 19(4), 650–655.

Kalhoefer, D., Thole, S., Voget, S., Lehmann, R., Liesegang, H., Wollher, A., ... Brinkhoff, T. (2011). Comparative genome analysis and genome-guided physiological analysis of *Roseobacter litoralis*. *BMC Genomics*, 12, 324.

Kamionka, A., Sehnal, M., Scholz, O., & Hillen, W. (2004). Independent regulation of two genes in *Escherichia coli* by tetracyclines and Tet repressor variants. *J. Bacteriol.* 186(13), 4399–4401.

Kanter, G., Yang, J., Voloshin, A., Levy, S., Swartz, J. R., & Levy, R. (2007). Cell free production of scFv fusion proteins: an efficient approach for personalized lymphoma vaccines. *Blood*, 109(8), 3393–3399.

Kaufmann, A., Mickoleit, M., Weber, M., & Huisken, J. (2012). Multilayer mounting enables long term imaging of zebrafish development in a light sheet microscope. *Development*, 139(17), 3242–3247.

Kawrykow, A., Roumanis, G., Kam, A., Kwak, D., Leung, C., Wu, C., ... Waldspuhl, J. (2012). Phylo: a citizen science approach for improving multiple sequence alignment. *PLoS ONE*, 7(3), e31362.

Kaznessis, Y. N. (2007). Models for synthetic biology. *BMC Syst Biol*, 1, 47.

Kean, S. (2011). A lab of their own. *Science*, 333(6047), 1240–1241.

Keasling, J. D. (2008). Synthetic biology for synthetic chemistry. *ACS Chem. Biol.* 3(1), 64–76.

Keating, S. M. & Le Novere, N. (2013). Supporting SBML as a model exchange format in software applications. *Methods Mol. Biol.* 1021, 201–225.

Kelwick, R., MacDonald, J. T., Webb, A. J., & Freemont, P. (2014). Developments in the tools and methodologies of synthetic biology. *Front Bioeng Biotechnol*, 2, 60.

Kemmer, C., Fluri, D. A., Witschi, U., Passeraub, A., Gutzwiller, A., & Fussenegger, M. (2011). A designer network coordinating bovine artificial insemination by ovulation triggered release of implanted sperms. *J Control Release*, 150(1), 23–29.

Kennel, M. B., Brown, R., & Abarbanel, H. D. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. A*, 45(6), 3403–3411.

Kepler, T. B. & Elston, T. C. (2001). Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys J*, 81(6), 3116–3136.

Kerkhoven, E. J., Achcar, F., Alibu, V. P., Burchmore, R. J., Gilbert, I. H., Trybio, M., ... Barrett, M. P. (2013). Handling uncertainty in dynamic models: the pentose phosphate pathway in *Trypanosoma brucei*. *PLoS Comput. Biol.* 9(12), e1003371.

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.

Khalil, A. S. & Collins, J. J. (2010). Synthetic biology: applications come of age. *Nat Rev Genet*, 11(5), 367–379.

Kim, E., Sun, L., Gabel, C. V., & Fang Yen, C. (2013). Long term imaging of *Caenorhabditis elegans* using nanoparticle mediated immobilization. *PLoS ONE*, 8(1), e53419.

Kim, I. K., Roldao, A., Siewers, V., & Nielsen, J. (2012). A systems-level approach for metabolic engineering of yeast cell factories. *FEMS Yeast Res.* 12(2), 228–248.

Kim, J. & Winfree, E. (2011). Synthetic in vitro transcriptional oscillators. *Mol. Syst. Biol.* 7, 465.

Kitney, R. & Freemont, P. (2012). Synthetic biology the state of play. *FEBS Lett.* 586(15), 2029–2036.

Kobayashi, H., Kaern, M., Araki, M., Chung, K., Gardner, T. S., Cantor, C. R., & Collins, J. J. (2004). Programmable cells: interfacing natural and engineered gene networks. *PNAS*, 101(22), 8414–8419.

Koseska, A., Volkov, E., & Kurths, J. (2011). Synthetic multicellular oscillatory systems: controlling protein dynamics with genetic circuits. *Phys Scripta*, 84(4), 045007.

Kovach, M. E., Phillips, R. W., Elzer, P. H., Roop, R. M., & Peterson, K. M. (1994). pBBR1MCS: a broad host range cloning vector. *BioTechniques*, 16(5), 800–802.

Kraeva, N., Ishemgulova, A., Luke, J., & Yurchenko, V. (2014). Tetracycline-inducible gene expression system in *Leishmania mexicana*. *Mol. Biochem. Parasitol.* 198(1), 11–13.

Kramer, B. P., Viretta, A. U., Daoud El Baba, M., Aubel, D., Weber, W., & Fussenegger, M. (2004). An engineered epigenetic transgene switch in mammalian cells. *Nat. Biotechnol.* 22(7), 867–870.

Kulkarni, M. M., Karafova, A., Kamysz, W., Schenkman, S., Pelle, R., & McGwire, B. S. (2013). Secreted trypanosome cyclophilin inactivates lytic insect defense peptides and induces parasite calcineurin activation and infectivity. *J. Biol. Chem.* 288(12), 8772–8784.

Kuznetsov, Y. (2004). *Elements of applied bifurcation theory*. Springer.

Kwok, R. (2010). Five hard truths for synthetic biology. *Nature*, 463(7279), 288–290.

La Flamme, A. C., Buckner, F. S., Swindle, J., Ajioka, J., & Van Voorhis, W. C. (1995). Expression of mammalian cytokines by trypanosoma cruzi indicates unique signal sequence requirements and processing. *Mol Biochem Parasit.* 75(1), 25–31.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227(5259), 680–685.

Lafay, B., Ruimy, R., de Traubenberg, C. R., Breitmayer, V., Gauthier, M. J., & Christen, R. (1995). Roseobacter algicola sp. nov., a new marine bacterium isolated from the phycosphere of the toxin producing dinoflagellate *Prorocentrum lima*. *Int. J. Syst. Bacteriol.* 45(2), 290–296.

Lakshmanan, M. & Rajaseekar, S. (2012). *Nonlinear dynamics: integrability, chaos and patterns*. Advanced Texts in Physics. Springer Berlin Heidelberg.

Lamendella, R., Strutt, S., Borglin, S., Chakraborty, R., Tas, N., Mason, O., ... Jansson, J. (2014). Assessment of the deepwater horizon oil spill impact on gulf coast microbial communities. *Front Microbiol.* 5(130).

Landrain, T., Meyer, M., Perez, A. M., & Sussan, R. (2013). Do it yourself biology: challenges and promises for an open science and technology movement. *Syst Synth Biol.* 7(3), 115–126.

Lang, M., Marquez-Lago, T. T., Stelling, J., & Waldherr, S. (2011). Autonomous synchronization of chemically coupled synthetic oscillators. *Bull. Math. Biol.* 73(11), 2678–2706.

Langousis, G. & Hill, K. L. (2014). Motility and more: the flagellum of *Trypanosoma brucei*. *Nat. Rev. Microbiol.* 12(7), 505–518.

Larson, M. H., Gilbert, L. A., Wang, X., Lim, W. A., Weissman, J. S., & Qi, L. S. (2013). CRISPR interference (CRISPRi) for sequence specific control of gene expression. *Nat Protoc.* 8(11), 2180–2196.

Ledford, H. (2010). Garage biotech: Life hackers. *Nature*, 467(7316), 650–652.

Lee, S. H., Stephens, J. L., & Englund, P. T. (2007). A fatty-acid synthesis mechanism specialized for parasitism. *Nat. Rev. Microbiol.* 5(4), 287–297.

Leech, V., Quail, M. A., & Melville, S. E. (2004). Separation, digestion, and cloning of intact parasite chromosomes embedded in agarose. *Methods Mol. Biol.* 270, 335–352.

Leier, A., Marquez-Lago, T. T., & Burrage, K. (2008). Generalized binomial tleap method for biochemical kinetics incorporating both delay and intrinsic noise. *J. Phys. Chem.* 128(20), 205107–205114.

Leite, M. C. A. & Wang, Y. (2010). Multistability, oscillations and bifurcations in feedback loops. *Math Biosci Eng.* 7(1), 83–97.

Lenk, S., Moraru, C., Hahnke, S., Arnds, J., Richter, M., Kube, M., ... Mussmann, M. (2012). Roseobacter clade bacteria are abundant in coastal sediments and encode a novel combination of sulfur oxidation genes. *ISME J.*, 6(12), 2178–2187.

Leonard, J. N. (2014). The rise of mammals. *ACS Synth Biol.*, 3(12), 878–879.

Leung, K. F., Riley, F. S., Carrington, M., & Field, M. C. (2011). Ubiquitylation and developmental regulation of invariant surface protein expression in trypanosomes. *Eukaryotic Cell*, 10(7), 916–931.

Li, W., Krishna, S., Pigolotti, S., Mitarai, N., & Jensen, M. H. (2012). Switching between oscillations and homeostasis in competing negative and positive feedback motifs. *J. Theor. Biol.* 307, 205–210.

Li, X., Zhao, X., Fang, Y., Jiang, X., Duong, T., Fan, C., ... Kain, S. R. (1998). Generation of destabilized green fluorescent protein as a transcription reporter. *J. Biol. Chem.* 273(52), 34970–34975.

Lienert, F., Lohmueller, J. J., Garg, A., & Silver, P. A. (2014). Synthetic biology in mammalian cells: next generation research tools and therapeutics. *Nat. Rev. Mol. Cell Biol.* 15(2), 95–107.

Liepe, J., Barnes, C., Cule, E., Erguler, K., Kirk, P., Toni, T., & Stumpf, M. P. (2010). ABC SysBio approximate Bayesian computation in Python with GPU support. *Bioinformatics*, 26(14), 1797–1799.

Liepe, J., Kirk, P., Filippi, S., Toni, T., Barnes, C. P., & Stumpf, M. P. (2014). A framework for parameter estimation and model selection from experimental data in systems biology using approximate Bayesian computation. *Nat Protoc.*, 9(2), 439–456.

Life Technologies Ltd. (n.d.). *One shot®TOP10 chemically competent E. coli*. Retrieved from <http://http://www.lifetechnologies.com/order/catalog/product/C404010?CID=cvc-top10-competent-cells-c1b1> [Accessed 17 January 2014].

Lin, L. & Xu, J. (2013). Dissecting and engineering metabolic and regulatory networks of thermophilic bacteria for biofuel production. *Biotechnol. Adv.* 31(6), 827–837.

Littmann, M. & Suomela, T. (2014). Crowdsourcing, the great meteor storm of 1833, and the founding of meteor science. *Endeavour*, 38(2), 130–138.

Liu, C., Fu, X., Liu, L., Ren, X., Chau, C. K., Li, S., ... Huang, J.-D. (2011). Sequential establishment of stripe patterns in an expanding cell population. *Science*, 334(6053), 238–241.

Liu, L., Redden, H., & Alper, H. S. (2013). Frontiers of yeast metabolic engineering: diversifying beyond ethanol and *Saccharomyces*. *Curr. Opin. Biotechnol.* 24(6), 1023–1030.

Lu, Z., Deng, Y., Van Nostrand, J., He, Z., Voordeckers, J., Zhou, A., ... Zhou, J. (2012). Microbial gene functions enriched in the deepwater horizon deep sea oil plume. *ISME J.* 6(2), 451–60.

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 Res.* 25(6), 1203–1210.

Ma, Y. F., Weiss, L. M., & Huang, H. (2012). A method for rapid regulation of protein expression in *Trypanosoma cruzi*. *Int. J. Parasitol.* 42(1), 33–37.

MacGregor, P., Rojas, F., Dean, S., & Matthews, K. R. (2013). Stable transformation of pleomorphic bloodstream form *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 190(2), 60–62.

MacLean, L., Myburgh, E., Rodgers, J., & Price, H. P. (2013). Imaging African trypanosomes. *Parasite Immunol.* 35(9-10), 283–294.

Maldonado, E., Hampsey, M., & Reinberg, D. (1999). Repression: targeting the heart of the matter. *Cell*, 99(5), 455–458.

Mann, H. & Whitney, D. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *Ann. Math. Stat.* 18(1), 50–60.

Mansfield, J. M. & Paulnock, D. M. (2008). Genetic manipulation of African trypanosomes as a tool to dissect the immunobiology of infection. *Parasite Immunol.* 30(4), 245–253.

Marcello, L. & Barry, J. D. (2007). Analysis of the VSG gene silent archive in *Trypanosoma brucei* reveals that mosaic gene expression is prominent in antigenic variation and is favored by archive substructure. *Genome Res.* 17(9), 1344–1352.

Marinho, F. A., Goncalves, K. C., Oliveira, S. S., Goncalves, D. S., Matteoli, F. P., Seabra, S. H., ... Branquinha, M. H. (2014). The calpain inhibitor MDL28170 induces the expression of apoptotic markers in *Leishmania amazonensis* promastigotes. *PLoS ONE*, 9(1), e87659.

Markson, J. S. & Elowitz, M. B. (2014). Synthetic biology of multicellular systems: new platforms and applications for animal cells and organisms. *ACS Synth Biol.* 3(12), 875–876.

MATLAB. (2012). *Version 7.14.0.737 (r2012a)*. Natick, Massachusetts: The MathWorks Inc.

Matthews, K. R. (2005). The developmental cell biology of *Trypanosoma brucei*. *J. Cell. Sci.* 118(Pt 2), 283–290.

McClung, C. R. (2006). Plant circadian rhythms. *Plant Cell*, 18(4), 792–803.

McIsaac, R., Gibney, P., Chandran, S., Benjamin, K., & Botstein, D. (2014). Synthetic biology tools for programming gene expression without nutritional perturbations in *Saccharomyces cerevisiae*. *Nucl. Acids Res.* 42(6), e48.

McKane, A., Nagy, J., Newman, T., & Stefanini, M. (2007). Amplified biochemical oscillations in cellular systems. *Journal of Statistical Physics*, 128(1-2), 165–191.

McLatchie, A. P., Burrell-Saward, H., Myburgh, E., Lewis, M. D., Ward, T. H., Mottram, J. C., ... Taylor, M. C. (2013). Highly sensitive *in vivo* imaging of *Trypanosoma brucei* expressing "red-shifted" luciferase. *PLoS Negl Trop Dis*, 7(11), e2571.

Medford, J. I. & Prasad, A. (2014). Plant Science. Plant synthetic biology takes root. *Science*, 346(6206), 162–163.

Medlock, J., Atkins, K. E., Thomas, D. N., Aksoy, S., & Galvani, A. P. (2013). Evaluating paratransgenesis as a potential control strategy for African trypanosomiasis. *PLoS Negl Trop Dis*, 7(8), e2374.

Melville, S. E. (Ed.). (2010). *Parasite genomics protocols* (1st ed.). Humana Press.

Menaker, M. (1996). Colin S Pittendrigh 1918-1996. *Nature*, 381, 24.

Menezes, A. A., Cumbers, J., Hogan, J. A., & Arkin, A. P. (2015). Towards synthetic biological approaches to resource utilization on space missions. *J R Soc Interface*, 12(102), 20140715.

Menolascina, F., Siciliano, V., & di Bernardo, D. (2012). Engineering and control of biological systems: A new way to tackle complex diseases. *FEBS Lett.* 586(15), 2122–2128.

Menzella, H. G. (2011). Comparison of two codon optimization strategies to enhance recombinant protein production in *Escherichia coli*. *Microb. Cell Fact.* 10, 15.

Miller, T. R. & Belas, R. (2006). Motility is involved in *Silicibacter* sp. TM1040 interaction with dinoflagellates. *Environ. Microbiol.* 8(9), 1648–1659.

Mirza, S. S., Qazi, J. I., Zhao, Q., & Chen, S. (2013). Photo biohydrogen production potential of *Rhodobacter capsulatus* PK from wheat straw. *Biotechnol Biofuels*, 6(1), 144.

Mirzaahmadi, S., Asaadi Tehrani, G., Bandehpour, M., Davoudi, N., Tahmasbi, L., Hosseinzadeh, N., ... Kazemi, B. (2011). Expression of recombinant human coagulation factor VII by the Lizard *Leishmania* expression system. *J. Biomed. Biotechnol.* 2011, 873874.

Mistry, H. (2007). *Stochastic modeling of gene regulatory networks* (Doctoral dissertation, The University of Manchester). Retrieved from <http://eprints.ma.man.ac.uk/794/>.

Mitra, S., Sana, B., & Mukherjee, J. (2014). Ecological Roles and Biotechnological Applications of Marine and Intertidal Microbial Biofilms. *Adv. Biochem. Eng. Biotechnol.*

Mondragon-Palomino, O., Danino, T., Selimkhanov, J., Tsimring, L., & Hasty, J. (2011). Entrainment of a population of synthetic genetic oscillators. *Science*, 333(6047), 1315–1319.

Monod, J. (1971). *Chance and necessity: an essay on the natural philosophy of modern biology*. New York: Knopf.

Moon, T. S., Lou, C., Tamsir, A., Stanton, B. C., & Voigt, C. A. (2012). Genetic programs constructed from layered logic gates in single cells. *Nature*, 491(7423), 249–253.

Moore, R., Chandrasas, A., & Bleris, L. (2014). Transcription activator-like effectors: a toolkit for synthetic biology. *ACS Synth Biol*, 3(10), 708–716.

Moran, M. A., Belas, R., Schell, M. A., Gonzalez, J. M., Sun, F., Sun, S., ... Buchan, A. (2007). Ecological genomics of marine Roseobacters. *Appl. Environ. Microbiol.* 73(14), 4559–4569.

Morrison, L. J., Marcello, L., & McCulloch, R. (2009). Antigenic variation in the African trypanosome: molecular mechanisms and phenotypic complexity. *Cell. Microbiol.* 11(12), 1724–1734.

Morrissey, E. R., Juarez, M. A., Denby, K. J., & Burroughs, N. J. (2011). Inferring the time-invariant topology of a nonlinear sparse gene regulatory network using fully Bayesian spline autoregression. *Biostatistics*, 12(4), 682–694.

Muller, K. M. & Arndt, K. M. (2012). Standardization in synthetic biology. *Methods Mol. Biol.* 813, 23–43.

Muller, S., Hofbauer, J., Endler, L., Flamm, C., Widder, S., & Schuster, P. (2006). A generalized model of the repressilator. *J Math Biol.* 53(6), 905–937.

Na, D., Yoo, S. M., Chung, H., Park, H., Park, J. H., & Lee, S. Y. (2013). Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. *Nat. Biotechnol.* 31(2), 170–174.

Neddermann, P., Gargioli, C., Muraglia, E., Sambucini, S., Bonelli, F., De Francesco, R., & Cortese, R. (2003). A novel, inducible, eukaryotic gene expression system based on the quorum sensing transcription factor TraR. *EMBO Rep.* 4(2), 159–165.

Nelson, B. (2014). Synthetic Biology: Cultural Divide. *Nature*, 509(7499), 152–154.

Newton, R. J., Griffin, L. E., Bowles, K. M., Meile, C., Gifford, S., Givens, C. E., ... Moran, M. A. (2010). Genome characteristics of a generalist marine bacterial lineage. *ISME J*, 4(6), 784–798.

Ni, M., Tong, W. H., Choudhury, D., Rahim, N. A., Iliescu, C., & Yu, H. (2009). Cell culture on MEMS platforms: a review. *Int J Mol Sci.* 10(12), 5411–5441.

Nielsen, A. A., Segall-Shapiro, T. H., & Voigt, C. A. (2013a). Advances in genetic circuit design: novel biochemistries, deep part mining, and precision gene expression. *Curr Opin Chem Biol.* 17(6), 878–892.

Nielsen, J., Larsson, C., van Maris, A., & Pronk, J. (2013b). Metabolic engineering of yeast for production of fuels and chemicals. *Curr. Opin. Biotechnol.* 24(3), 398–404.

Nikel, P. I., Martinez-Garcia, E., & de Lorenzo, V. (2014). Biotechnological domestication of pseudomonads using synthetic biology. *Nat. Rev. Microbiol.* 12(5), 368–379.

Nistala, G. J., Wu, K., Rao, C. V., & Bhalerao, K. D. (2010). A modular positive feedback based gene amplifier. *J. Biol. Eng.* 4, 4.

Novak, B. & Tyson, J. J. (2008). Design principles of biochemical oscillators. *Nat Rev Mol Cell Biol*, 9(12), 981–991.

Oberholzer, M., Lopez, M. A., McLelland, B. T., & Hill, K. L. (2010). Social motility in African trypanosomes. *PLoS Pathog.* 6(1), e1000739.

O'Brien, E. L., Itallie, E. V., & Bennett, M. R. (2012). Modeling synthetic gene oscillators. *Math Biosci.* 236(1), 1–15.

O'Connor, S. E. & Brutnell, T. P. (2014). Editorial overview: Growing the future: synthetic biology in plants. *Curr. Opin. Plant Biol.* 19, iv–v.

Oppenheim, A., Schafer, R., & Buck, J. (1999). *Discrete-time signal processing*. Prentice-Hall.

Orosz, G., Moehlis, J., & Murray, R. (2010). Controlling biological networks by time delayed signals. *Philosophical Transactions A*, 368, 439–454.

Osella, M. & Lagomarsino, M. C. (2013). Growth rate dependent dynamics of a bacterial genetic oscillator. *Phys Rev E Stat Nonlin Soft Matter Phys*, 87(1), 012726.

Pace Pereira Lima, G., Arnoux da Silva Campos, R., Gomes Willadino, L., Camara, T., & Vianello, F. (2012). Recent advances in plant in vitro culture. In A. Leva (Ed.), (Chap. 9). Retrieved from <<http://www.intechopen.com/books/recent-advances-in-plant-in-vitro-culture/polyamines-gelling-agents-in-tissue-culture-micropropagation-of-medicinal-plants-and-bioreactors>>.

Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., . . . Newman, J. D. (2013). High level semi synthetic production of the potent antimalarial artemisinin. *Nature*, 496(7446), 528–532.

Pal, N. & Pal, S. (1993). A review on image segmentation techniques. *Pattern Recognit.* 26(9), 1277–1294.

Palenchar, J. B. & Bellofatto, V. (2006). Gene transcription in trypanosomes. *Mol. Biochem. Parasitol.* 146(2), 135–141.

Panke, S. (2005). Synthetic biology: the (long) road to bioengineering. *34th Meeting of the European Group on Ethics in Science and New Technologies, Brussels, Belgium, 22 April 2009.* N.A: N.A.

Pappas, K. M., Weingart, C. L., & Winans, S. C. (2004). Chemical communication in proteobacteria: biochemical and structural studies of signal synthases and receptors required for intercellular signalling. *Mol. Microbiol.* 53(3), 755–769.

Park, S. H., Nguyen, T. N., Kirkham, J. K., Lee, J. H., & Gunzl, A. (2011). Transcription by the multifunctional RNA polymerase I in *Trypanosoma brucei* functions independently of RPB7. *Mol. Biochem. Parasitol.* 180(1), 35–42.

Park, S. H., Zarrinpar, A., & Lim, W. A. (2003). Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms. *Science*, 299(5609), 1061–1064.

Pays, E. (Ed.). (1999). The trypanosome surface: Proceedings of the third francqui colloquium, Brussels: De Boeck Superieur.

Peacock, L., Bailey, M., & Gibson, W. (2005). Tetracycline induction of gene expression in *trypanosoma brucei* within the tsetse fly vector. *Mol Biochem Parasit.* 140(2), 247–249.

Pearce, J. M. (2012). Materials science. Building research equipment with free, open source hardware. *Science*, 337(6100), 1303–1304.

Pelkmans, L. (2012). Cell Biology. Using cell-to-cell variability—a new era in molecular biology. *Science*, 336(6080), 425–426.

Peralta-Yahya, P., Zhang, F., del Cardayre, S., & Keasling, J. (2012). Microbial engineering for the production of advanced biofuels. *Nature*, 488(7411), 320–328.

Perkel, J. M. (2014). Seamlessly rewriting the lab cloning manual. *BioTechniques*, 56(1), 12–14.

Perry, N. & Ninfa, A. J. (2012). Synthetic networks: oscillators and toggle switches for *Escherichia coli*. *Methods Mol. Biol.* 813, 287–300.

Petersen, J., Frank, O., Goker, M., & Pradella, S. (2013). Extrachromosomal, extraordinary and essential the plasmids of the Roseobacter clade. *Appl. Microbiol. Biotechnol.* 97(7), 2805–2815.

Piekarski, T., Buchholz, I., Drepper, T., Schobert, M., Wagner-Doebler, I., Tielen, P., & Jahn, D. (2009). Genetic tools for the investigation of Roseobacter clade bacteria. *BMC Microbiol.* 9, 265.

Pion, C., Courtois, V., Husson, S., Bernard, M. C., Nicolai, M. C., Talaga, P., . . . Legastelois, I. (2014). Characterization and immunogenicity in mice of recombinant influenza haemagglutinins produced in *Leishmania tarentolae*. *Vaccine*, 32(43), 5570–5576.

Pitt, T. L. & Bradley, D. E. (1975). The antibody response to the flagella of *Pseudomonas aeruginosa*. *J. Med. Microbiol.* 8(1), 97–106.

Podlipaev, S. (2001). The more insect trypanosomatids under study-the more diverse Trypanosomatidae appears. *Int. J. Parasitol.* 31(5-6), 648–652.

Pomerening, J. R., Kim, S. Y., & Ferrell Jr., J. E. (2005). Systems level dissection of the cell cycle oscillator: bypassing positive feedback produces damped oscillations. *Cell*, 122(4), 565–578.

Price, H. P., MacLean, L., Morrison, J., O'Toole, P. J., & Smith, D. F. (2010). Validation of a new method for immobilising kinetoplastid parasites for live cell imaging. *Mol. Biochem. Parasitol.* 169(1), 66–69.

Purcell, O., Savery, N. J., Grierson, C. S., & di Bernardo, M. (2010). A comparative analysis of synthetic genetic oscillators. *J. R. Soc. Interface*. DOI: 10.1098/rsif.2010.0183.

Radeck, J., Kraft, K., Bartels, J., Cikovic, T., Durr, F., Emenegger, J., ... Mascher, T. (2013). The *Bacillus* BioBrick Box: generation and evaluation of essential genetic building blocks for standardized work with *Bacillus subtilis*. *J. Biol. Eng.* 7(1), 29.

Ralston, K. S., Kabututu, Z. P., Melehani, J. H., Oberholzer, M., & Hill, K. L. (2009). The *Trypanosoma brucei* flagellum: moving parasites in new directions. *Annu. Rev. Microbiol.* 63, 335–362.

Raser, J. M. & O'Shea, E. K. (2004). Control of stochasticity in eukaryotic gene expression. *Science*, 304(5678), 1811–1814.

Raser, J. M. & O'Shea, E. K. (2005). Noise in gene expression: origins, consequences, and control. *Science*, 309(5743), 2010–2013.

Raymond, F., Boisvert, S., Roy, G., Ritt, J. F., Legare, D., Isnard, A., ... Corbeil, J. (2012). Genome sequencing of the lizard parasite *Leishmania tarentolae* reveals loss of genes associated to the intracellular stage of human pathogenic species. *Nucleic Acids Res.* 40(3), 1131–1147.

Rayner, S., Heyward, C., Kruger, T., Pidgeon, N., Redgwell, C., & Savulescu, J. (2013). The oxford principles. *Climatic Change*, 121(3), 499–512.

Registry of Standard Biological Parts. (2014). *Frequently used parts*. Retrieved from http://parts.igem.org/Frequently_Used_Parts [Accessed 7 May 2014].

Renaud, O., Herbomel, P., & Kissa, K. (2011). Studying cell behavior in whole zebrafish embryos by confocal live imaging: application to hematopoietic stem cells. *Nat Protoc*, 6(12), 1897–1904.

Riedel, T., Teshima, H., Petersen, J., Fiebig, A., Davenport, K., Daligault, H., ... Klenk, H. P. (2013). Genome sequence of the *Leisingera aquimarina* type strain (DSM 24565(T)), a member of the marine *Roseobacter* clade rich in extrachromosomal elements. *Stand Genomic Sci*, 8(3), 389–402.

Rivadeneyra, A., Gonzalez-Martinez, A., Gonzalez-Lopez, J., Martin-Ramos, D., Martinez-Toledo, M. V., & Rivadeneyra, M. A. (2014). Precipitation of phosphate minerals by microorganisms isolated from a fixed biofilm reactor used for the treatment of domestic wastewater. *Int J Environ Res Public Health*, 11(4), 3689–3704.

Rossi, F. M., Guicherit, O. M., Spicher, A., Kringstein, A. M., Fatyol, K., Blakely, B. T., & Blau, H. M. (1998). Tetracycline-regulatable factors with distinct dimerization domains allow reversible growth inhibition by p16. *Nat. Genet.* 20(4), 389–393.

Rudenko, G. (2011). African trypanosomes: the genome and adaptations for immune evasion. *Essays Biochem.* 51, 47–62. PMID: 22023441.

Ruder, W. C., Lu, T., & Collins, J. J. (2011). Synthetic biology moving into the clinic. *Science*, 333(6047), 1248–1252. DOI: 10.1126/science.1206843.

Ruoff, P., Vinsjevik, M., Monnerjahn, C., & Rensing, L. (2001). The Goodwin model: simulating the effect of light pulses on the circadian sporulation rhythm of *Neurospora crassa*. *J. Theor. Biol.* 209(1), 29–42.

Saadatpour, A. & Albert, R. (2013). Boolean modeling of biological regulatory networks: a methodology tutorial. *Methods*, 62(1), 3–12.

Salis, H. & Kaznessis, Y. (2005). Accurate hybrid stochastic simulation of a system of coupled chemical or biochemical reactions. *J. Phys. Chem.* 122(5), 54103.

Samanovic, M., Molina-Portela, M. P., Chessler, A. D., Burleigh, B. A., & Raper, J. (2009). Trypanosome lytic factor, an antimicrobial high-density lipoprotein, ameliorates Leishmania infection. *PLoS Pathog.* 5(1), e1000276.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (2001). *Molecular cloning: a laboratory manual* (Third). Cold Spring Harbor Laboratory.

San-Miguel, A. & Lu, H. (2013). Microfluidics as a tool for *C. elegans* research. *WormBook*, 1–19.

Santo, M., Weitsman, R., & Sivan, A. (2013). The role of the copper binding enzyme - laccase - in the biodegradation of polyethylene by the actinomycete *Rhodococcus ruber*. *International Biodeterioration & Biodegradation*, 84, 204–210.

Sato-Takabe, Y., Hamasaki, K., & Suzuki, K. (2014). Photosynthetic competence of the marine aerobic anoxygenic phototrophic bacterium *Roseobacter* sp. under organic substrate limitation. *Microbes Environ.* 29(1), 100–103.

Saukshmya, T. & Chugh, A. (2010). Intellectual property rights in synthetic biology: an anti thesis to open access to research? *Syst Synth Biol.* 4(4), 241–245.

Sayut, D. J., Kambam, P. K. R., & Sun, L. (2007). Engineering and applications of genetic circuits. *Mol. BioSyst.* 3(12), 835–840.

Schimanski, B., Laufer, G., Gontcharova, L., & Gunzl, A. (2004). The *Trypanosoma brucei* spliced leader RNA and rRNA gene promoters have interchangeable TbSNAP50 binding elements. *Nucleic Acids Res.* 32(2), 700–709.

Schmidt, C. W. (2010). Synthetic biology, environmental health implications of a new field. *Environ. Health Perspect.* 118(3), 118–123.

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*, 9(7), 671–675.

Schonenberger, B. (1979). Cultivation and in vitro cloning or procyclic culture forms of trypanosoma brucei in a semi defined medium. short communication. *Acta Trop.* 36(3), 289–92.

Schurch, N., Furger, A., Kurath, U., & Roditi, I. (1997). Contributions of the procyclin 3' untranslated region and coding region to the regulation of expression in bloodstream forms of trypanosoma brucei. *Mol Biochem Parasit.* 89(1), 109–121.

Schwartz, R. (2008). *Biological modeling and simulation: a survey of practical models, algorithms, and numerical methods*. Cambridge, Massachusetts, U.S.A: MIT Press.

Scudellari, M. (2013). *Biology hacklabs*. Retrieved from <http://www.the-scientist.com/?articles-view/articleNo/34469/title/Biology-Hacklabs/> [Accessed 30 April 2014].

Sepulveda, S., Valenzuela, L., Ponce, I., Sierra, S., Bahamondes, P., Ramirez, S., ... Cabrera, G. (2014). Expression, functionality, and localization of apurinic/apyrimidinic endonucleases in replicative and non replicative forms of *Trypanosoma cruzi*. *J. Cell. Biochem.* 115(2), 397–409.

Serricchio, M. & Butikofer, P. (2011). *Trypanosoma brucei*: a model micro-organism to study eukaryotic phospholipid biosynthesis. *FEBS Journal*, 278(7), 1035–1046.

Seydel, R. (1994). *Practical bifurcation and stability analysis: from equilibrium to chaos*. Springer.

Seyfried, G., Pei, L., & Schmidt, M. (2014). European do-it-yourself (DIY) biology: beyond the hope, hype and horror. *Bioessays*, 36(6), 548–551.

Shaner, N. C., Steinbach, P. A., & Tsien, R. Y. (2005). A guide to choosing fluorescent proteins. *Nat. Methods*, 2(12), 905–909.

Sheinman, M. & Kafri, Y. (2012). How does the DNA sequence affect the Hill curve of transcriptional response? *Phys Biol.* 9(5), 056006.

Sheskin, D. (2003). *Handbook of parametric and nonparametric statistical procedures* (Third). London: CRC Press.

Shetty, R., Lizarazo, M., Rettberg, R., & Knight, T. F. (2011). Assembly of BioBrick standard biological parts using three antibiotic assembly. *Meth. Enzymol.* 498, 311–326.

Shiba, T. (1991). Roseobacter litoralis gen. nov., sp. nov., and Roseobacter denitrificans sp. nov., aerobic pink-pigmented bacteria which contain bacteriochlorophyll a. *Systematic and Applied Microbiology*, 14(2), 140–145.

Siegel, T. N., Gunasekera, K., Cross, G. A., & Ochsenreiter, T. (2011). Gene expression in *Trypanosoma brucei*: lessons from high-throughput RNA sequencing. *Trends Parasitol.* 27(10), 434–441.

Simpson, Frech, G. C., & Maslov, D. A. (1991). RNA editing in trypanosomatid mitochondria. *Annual Review of Microbiology*, 45, 99–121.

Smith, M. T., Wilding, K. M., Hunt, J. M., Bennett, A. M., & Bundy, B. C. (2014). The emerging age of cell-free synthetic biology. *FEBS Lett.* 588(17), 2755–2761.

Smith, T. K., Vasileva, N., Gluenz, E., Terry, S., Portman, N., Kramer, S., ... Rudenko, G. (2009). Blocking variant surface glycoprotein synthesis in *Trypanosoma brucei* triggers a general arrest in translation initiation. *PLoS ONE*, 4(10), e7532.

Smolen, P., Baxter, D. A., & Byrne, J. H. (2000). Mathematical modeling of gene networks. *Neuron*, 26(3), 567–580.

Snijder, B. & Pelkmans, L. (2011). Origins of regulated cell-to-cell variability. *Nat. Rev. Mol. Cell Biol.* 12(2), 119–125.

Stephanopoulos, G. (2012). Synthetic biology and metabolic engineering. *ACS Synth Biol.* 1(11), 514–525.

Stockdale, C., Swiderski, M. R., Barry, J. D., & McCulloch, R. (2008). Antigenic variation in *Trypanosoma brucei*: joining the DOTs. *PLoS Biol.* 6(7), e185.

Strelkowa, N. & Barahona, M. (2010). Switchable genetic oscillator operating in quasi stable mode. *J. R. Soc. Interface*, 7(48), 1071–1082. DOI: 10.1098/rsif.2009.0487.

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. DOI: 10.1038/nature07389.

Sugino, M. & Niimi, T. (2012). Expression of multisubunit proteins in *Leishmania tarentolae*. *Methods Mol. Biol.* 824, 317–325.

Swingley, W. D., Sadekar, S., Mastrian, S. D., Matthies, H. J., Hao, J., Ramos, H., ... Touchman, J. W. (2007). The complete genome sequence of *Roseobacter denitrificans* reveals a mixotrophic rather than photosynthetic metabolism. *J. Bacteriol.* 189(3), 683–690.

Tabor, J. J. (2012). Modular gene circuit design takes two steps forward. *Nat. Methods*, 9(11), 1061–1063.

Tabor, J. J., Salis, H. M., Simpson, Z. B., Chevalier, A. A., Levskaya, A., Marcotte, E. M., ... Ellington, A. D. (2009). A synthetic genetic edge detection program. *Cell*, 137(7), 1272–1281.

Takens, F. (1981). Detecting strange attractors in turbulence. *Lect. Notes Math.* 898, 366–381.

Tamsir, A., Tabor, J. J., & Voigt, C. A. (2011). Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'. *Nature*, 469(7329), 212–215.

Teschl, G. (2012). *Ordinary differential equations and dynamical systems*. American Mathematical Society.

The Economist. (2010). *Synthetic biology: and man made life*. Retrieved from <http://www.economist.com/node/16163154> [Accessed 1 May 2014].

Tigges, M., Denervaud, N., Greber, D., Stelling, J., & Fussenegger, M. (2010). A synthetic low frequency mammalian oscillator. *Nucleic Acids Research*, 38(8), 2702–2711.

Tigges, M., Marquez-Lago, T. T., Stelling, J., & Fussenegger, M. (2009). A tunable synthetic mammalian oscillator. *Nature*, 457(7227), 309–312. DOI: 10.1038/nature07616.

Tsai, T. Y., 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*, 321(5885), 126–129. DOI: 10.1126/science.1156951.

Tucker, J. (2012). *Innovation, dual use, and security: managing the risks of emerging biological and chemical technologies*. MIT Press.

Tuttle, L. M., Salis, H., Tomshine, J., & Kaznessis, Y. N. (2005). Model Driven designs of an oscillating gene network. *Biophys J*, 89(6), 3873–3883.

Tyson, J. & Othmer, H. (1978). The dynamics of feedback control circuits in biochemical pathways. *Prog. Theor. Biol.* 5, 1–62.

Vacchina, P. & Morales, M. A. (2014). In vitro screening test using *Leishmania* promastigotes stably expressing mCherry protein. *Antimicrob. Agents Chemother.* 58(3), 1825–1828.

van Doren, D., Koenigstein, S., & Reiss, T. (2013). The development of synthetic biology: a patent analysis. *Syst. Synth. Biol.* 7(4), 209–220.

Vandermeulen, G., Marie, C., Scherman, D., & Preat, V. (2011). New generation of plasmid backbones devoid of antibiotic resistance marker for gene therapy trials. *Mol. Ther.* 19(11), 1942–1949.

Vardi, Y., Ying, Z., & Zhang, C. (2001). Two-sample tests for growth curves under dependent right censoring. *Biometrika*, 88(4), 949–960.

Vasquez, J. J., Hon, C. C., Vanselow, J. T., Schlosser, A., & Siegel, T. N. (2014). Comparative ribosome profiling reveals extensive translational complexity in different *Trypanosoma brucei* life cycle stages. *Nucleic Acids Res.* 42(6), 3623–3637.

Verlinde, C. L., Hannaert, V., Blonski, C., Willson, M., Perie, J. J., Fothergill-Gilmore, L. A., ... Michels, P. A. (2001). Glycolysis as a target for the design of new anti-trypanosome drugs. *Drug Resist. Updat.* 4(1), 50–65.

Verner, Z., Paris, Z., & Lukes, J. (2010). Mitochondrial membrane potential-based genome-wide RNAi screen of *Trypanosoma brucei*. *Parasitol. Res.* 106(5), 1241–1244.

Vetter, B. V., Pantidos, N., & Edmundson, M. (2014). 2nd congress on applied synthetic biology in Europe (Malaga, Spain, November 2013). *N Biotechnol.* 31(3), 221–229.

Vilar, J. M., Kueh, H. Y., Barkai, N., & Leibler, S. (2002). Mechanisms of noise-resistance in genetic oscillators. *PNAS*, 99(9), 5988–5992.

Vinson, V. & Pennisi, E. (2011). Synthetic biology. The allure of synthetic biology. Introduction. *Science*, 333(6047), 1235.

Voigt, C. (2011). *Synthetic biology: methods for part/device characterization and chassis engineering*. Academic Press.

Wagner, E. G. & Simons, R. W. (1994). Antisense RNA control in bacteria, phages, and plasmids. *Annu. Rev. Microbiol.* 48, 713–742.

Wagner-Dobler, I., Rheims, H., Felske, A., El-Ghezal, A., Flade-Schroder, D., Laatsch, H., ... Tindall, B. J. (2004). Oceanibulbus indolifex gen. nov., sp. nov., a North Sea alphaproteobacterium that produces bioactive metabolites. *Int. J. Syst. Evol. Microbiol.* 54(Pt 4), 1177–1184.

Wang, G. P. & Bushman, F. D. (2006). A statistical method for comparing viral growth curves. *J. Virol. Methods*, 135(1), 118–123.

Wang, R., Jingb, Z., & Chen, L. (2005). Modelling periodic oscillation in gene regulatory networks by cyclic feedback systems. *Bull Math Biol.* 67, 339–367. DOI: 10.1016/j.bulm.2004.07.005.

Wang, Y. N., Wang, M., & Field, M. C. (2010). *Trypanosoma brucei*: trypanosome-specific endoplasmic reticulum proteins involved in variant surface glycoprotein expression. *Exp. Parasitol.* 125(3), 208–221.

Weber, W., Luzi, S., Karlsson, M., Sanchez-Bustamante, C. D., Frey, U., Hierlemann, A., & Fussenegger, M. (2009). A synthetic mammalian electro genetic transcription circuit. *Nucleic Acids Res.* 37(4), e33.

Weber, W., Schoenmakers, R., Spielmann, M., El-Baba, M. D., Folcher, M., Keller, B., ... Fussenegger, M. (2003). Streptomyces derived quorum sensing systems engineered for adjustable transgene expression in mammalian cells and mice. *Nucleic Acids Res.* 31(14), e71.

Weisse, S., Heddergott, N., Heydt, M., Pflasterer, D., Maier, T., Haraszti, T., ... Rosenhahn, A. (2012). A quantitative 3D motility analysis of *Trypanosoma brucei* by use of digital in line holographic microscopy. *PLoS ONE*, 7(5), e37296.

Weldemichael, D. A. & Grossberg, G. T. (2010). Circadian rhythm disturbances in patients with alzheimer's disease: a review. *International Journal of Alzheimer's Disease*, 2010, 1–9. DOI: 10.4061/2010/716453.

Whitehead, N. A., Barnard, A. M., Slater, H., Simpson, N. J., & Salmond, G. P. (2001). Quorum sensing in Gram negative bacteria. *FEMS Microbiol. Rev.* 25(4), 365–404.

Wilkinson, D. (2011). *Stochastic modelling for systems biology* (2nd ed.). CRC Press.

Willets, D. (2013). *Eight great technologies*. London: Policy Exchange. Retrieved from <http://policyexchange.org.uk/images/publications/eight20great20technologies.pdf> [Accessed 29 April 2014].

Williams, P. G. (2009). Panning for chemical gold, marine bacteria as a source of new therapeutics. *Trends Biotechnol.* 27(1), 45–52.

Win, M. N., Liang, J. C., & Smolke, C. D. (2009). Frameworks for programming biological function through RNA parts and devices. *Chem. Biol.* 16(3), 298–310.

Wittner, M., Squillante, L., Nadler, J. P., & Tanowitz, H. B. (1982). *Trypanosoma cruzi*: colony formation and clonal growth in agar. *Exp. Parasitol.* 53(2), 255–261.

Wu, J. D., Hsueh, H. C., Huang, W. T., Liu, H. S., Leung, H. W., Ho, Y. R., ... Lai, M. D. (1997). The inducible lactose operator repressor system is functional in the whole animal. *DNA Cell Biol.* 16(1), 17–22.

Xiao, M. & Cao, J. (2008). Genetic oscillation deduced from hopf bifurcation in a genetic regulatory network with delays. *Math Biosci.* 215(1), 55–63.

Xiao, S., Shiloach, J., & Betenbaugh, M. J. (2014). Engineering cells to improve protein expression. *Curr. Opin. Struct. Biol.* 26C, 32–38.

Yang, F., Mitra, P., Zhang, L., Prak, L., Verhertbruggen, Y., Kim, J. S., ... Loque, D. (2013a). Engineering secondary cell wall deposition in plants. *Plant Biotechnol. J.* 11(3), 325–335.

Yang, X., Lau, K. Y., Sevim, V., & Tang, C. (2013b). Design principles of the yeast G1S switch. *PLoS Biol.* 11(10), e1001673.

Ye, H., Daoud-El Baba, M., Peng, R. W., & Fussenegger, M. (2011). A synthetic optogenetic transcription device enhances blood glucose homeostasis in mice. *Science*, 332(6037), 1565–1568.

Ye, H., Aubel, D., & Fussenegger, M. (2013). Synthetic mammalian gene circuits for biomedical applications. *Current Opinion in Chemical Biology*, 17(6), 910–917.

Yoda, M., Ushikubo, T., Inoue, W., & Sasai, M. (2007). Roles of noise in single and coupled multiple genetic oscillators. *J. Phys. Chem.* 126(11), 115101–115111. DOI: 10.1063/1.2539037.

Yokobayashi, Y., Weiss, R., & Arnold, F. H. (2002). Directed evolution of a genetic circuit. *PNAS*, 99(26), 16587–16591.

Yongling, S. & Su-Shing, C. (2006). Item sets based graph mining algorithm and application in genetic regulatory networks. In *Granular Computing, 2006 IEEE International Conference on* (pp. 337–340).

You, L., Cox, R. S., Weiss, R., & Arnold, F. H. (2004). Programmed population control by cell cell communication and regulated killing. *Nature*, 428(6985), 868–871.

Zan, J., Liu, Y., Fuqua, C., & Hill, R. T. (2014). Acylhomoserine lactone quorum sensing in the Roseobacter clade. *Int J Mol Sci*, 15(1), 654–669.

Zhang, C., Khoshmanesh, K., Mitchell, A., & Kalantar-Zadeh, K. (2010). Dielectrophoresis for manipulation of micro/nano particles in microfluidic systems. *Anal Bioanal Chem*, 396(1), 401–420.

Zhang, G., Gurtu, V., & Kain, S. R. (1996). An enhanced green fluorescent protein allows sensitive detection of gene transfer in mammalian cells. *Biochem. Biophys. Res. Commun.* 227(3), 707–711.

Zhang, W. & Nielsen, D. R. (2014). Synthetic biology applications in industrial microbiology. *Front Microbiol*, 5, 451.

Zheng, Y. & Sriram, G. (2010). Mathematical modeling: bridging the gap between concept and realization in synthetic biology. *J Biomed Biotechnol*, 2010.

Zhou, T., Zhang, J., Yuan, Z., & Chen, L. (2008). Synchronization of genetic oscillators. *Chaos*, 18(3).