UNIVERSITY OF MANCHESTER

Engineering Tuneable Gene Circuits in Yeast

by

Stephen Checkley

A thesis submitted to the University of Manchester for the degree of Doctor of Science

in the school of
Chemical Engineering and Analytical Science
Systems Biology Doctoral Training Centre

2011

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AP	Alkaline Phosphatase
bp	base pairs
cDNA	Complementary deoxyribonucleic acid
CFP	Cyan Fluorescent Protein
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonuceoside triphosphate
EDTA	Ethylene diaminetetra-acetic acid
E. coli	*Escherichia coli*
FITC	Fluorescein Isothiocyanate
g	grams
g	relative centrifugal force
G-Protein	guanine nucleotide-binding proteins
GDP	guanosine diphosphate
GFP	Green Fluorescence Protein
GTP	guanosine triphosphate
HIS3	Imidazoleglycerol-phosphate dehydratase
HRP	Horseradish peroxidase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRP	Iron Response Protein
IRE	Iron Responsive Element
Kb	kilobase pairs
K_d	Dissociation constant

K_i	Inhibition constant
L	Litre
LB	Luria-Bertani
LexAop	LexA Operator Sequence
LiAc	Lithium acetate
LEU2	Beta-isopropylmalate dehydrogenase
M	Molar
MAPK	Mitogen Activated Protein Kinase
MCA	Metabolic Control Analysis
MCR	Multi-cloning region
min	Minute
nt	Nucleotide
OD	Optical Density
ODE	Ordinary Differential Equation
PCR	Polymerase Chain Reaction
P_{DCD1}	Yeast *DCD1* gene promoter
P_{FUS1}	Yeast *FUS1* gene promoter
P_{TEF1}	Yeast *TEF1* gene promoter
P_{TRP1}	Yeast *TRP1* gene promoter
PEG	Polyethylene glycol
PRE	Pheromone Response Element
QconCAT	Quantification concatamer
RT-qPCR	Reverse Transcription Quantitative Polymerase Chain Reaction
rpm	Revolutions per Minute
SBML	Systems Biology Mark-up Language
S. cerevisiae	*Saccharomyces cerevisiae*
SDS	Sodium doedecyl sulphate
sec	Second
SNR	Signal to Noise Ratio
TAE	Tris-acetate-EDTA
Tris	Tris[hydroxymethyl] aminomethane

UAE	Upstream Activator Sequence
URA3	Orotidine-5'-phosphate decarboxylase
V	Volts
wt	Wild-type
YFP	Yellow Fluorescent Protein
YNB	Yeast Nitrogen Base
YPD	Yeast Peptone Dextrose

Abstract

Synthetic biology is an emergent field incorporating aspects of computer science molecular biology-based methodologies in a systems biology context, taking naturally occurring cellular systems, pathways, and molecules, and selectively engineering them for the generation of novel or beneficial synthetic behaviour. This study described the construction of a novel synthetic gene circuit, which utilises the inducible downstream transcriptional activation properties of the pheromone-response pathway in the budding yeast *Saccharomyces cerevisiae* as the basis for initiation. The circuit was composed of three novel yeast expression plasmids; (1) a reporter plasmid in which the luciferase reporter gene was fused to the iron response element (IRE), and expressed under the control of the pheromone-inducible *FUS1* promoter, (2) a repressor plasmid which constitutively expressed the mammalian iron response protein (IRP), which can bind to the IRE in the luciferase mRNA transcript, blocking translation, and (3) a de-repressor plasmid which also utilised the pheromone-inducible *FUS1* promoter to express the bacterial LexA protein that represses transcription of the IRP gene, and thereby de-represses luciferase translation.

Yeast cultures were propagated in media that selected for cells containing all three plasmid components of the gene circuit. In these cells, during vegetative growth conditions, reporter gene translation is constitutively repressed by IRP until addition of pheromone. Upon pheromone-induction, the pheromone response pathway up-regulated the expression of the LexA protein which represses transcription of IRP, enabling the translation of luciferase, which is itself up-regulated by the pheromone response pathway. The combination of the repressors functioned to increase the ratio of induction of the reporter gene between pheromone-induced and un-induced states. Proteins and mRNA species expressed by each plasmid were semi-quantified using SDS-PAGE, Western blot, and RT-qPCR. Luciferase expression was measured using an *in vitro* whole cell luminescence assay, and the data used to define the circuit "output".

Metabolic control analysis was used prior to building the circuit *in silico*, and identified the transcription of IRP, as well as the IRP protein half-life as significant control points for increasing the expression of luciferase *in vivo*. Modelling resulted in the development of multiple variations of the circuit, incorporating strong and weak constitutive promoters for the IRP. For the degradation rate, the IRP was fused with a degradation tag from the PEST rich C-terminal residue of the Cln2 protein, forming IRP_{PEST}, with approximately a 10-fold reduced half-life compared to wild type. By varying the promoter strength and half-life of the IRP, the circuit could be tuned in terms of the amplitude and period of luciferase expression during pheromone induction.

Simulated annealing and Hooke-Jeeves algorithms were used to estimate model parameter values from the experimental luminescence data, refining the modelling such that it produced accurate time course simulation of the circuit output. While further characterisation of the individual components would be advantageous, the construction of the system represents a completed cycle of extensive modelling, experimentation, and further model refinement.

Declaration of Authorship

I confirm that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Acknowledgements

I would like to acknowledge the members of the McCarthy lab for their help throughout this project, particularly Sheona Drummond, John Hughes, Maja Firczuk, Abi Stevenson, John Hildyard, John Duncan, Naglis Malys, Paola Pietroni, Mary Ortmayer, and Zurin Tajul Arifin. I would also like to thank Jean-Marc Schwartz for his help with the final corrections for this document.

I would like to thank Professor Hans Westerhoff for giving me the opportunity to study at the University of Manchester, and thank Professor Gerold Baier for his time and patience in teaching me maths.

Finally I would like to acknowledge my supervisors, Professors John McCarthy and Pedro Mendes, for allowing me the opportunity to work on this project.

Dedicated to Kim,
for all of the years of encouragement and support.

“*What I cannot build, I cannot understand.*”

Richard Feynman

Preface

This work details investigations into synthetic gene circuitry by the parallel employment of both *in vivo* laboratory experiments and *in silico* computational modelling. The two approaches were employed in a simultaneous fashion, and indeed each approach frequently used information gained from the other to validate hypotheses, and aid rational experimental design. While it is hoped that this work thus demonstrates the power of such interdisciplinary methods, for illustrative clarity the author has largely segregated the descriptions of the two aspects of this work. As such, the computational modelling collected together in chapter five actually describes simulations developed over the course of this entire work, and can be cross-referenced with the laboratory experiments depicted in chapter four. The author has made efforts to indicate such cross-referencing in the text, wherever appropriate.

CHAPTER

ONE

INTRODUCTION

1.1 Aims and Objectives

The objective of this project was to build a novel gene circuit in the yeast *Saccharomyces cerevisiae* that could enable cells to respond to environmental stimuli with the expression of a quantifiable reporter gene. Published research has shown that reporter genes can be coupled to promoters that control the expression of genes involved in the yeast pheromone response pathway (or "mating pathway") enabling cells to express a gene of interest in response to the presence of an extracellular stimulus. [1–3]. In this way, the project investigated the construction of a synthetic system that could be used to study features such as amplification, sensitivity, and noise.

Previous research in the McCarthy lab characterized the human iron response protein (IRP) and its interaction with genes containing the iron response element (IRE) as an effective repressor of translation in yeast [4]. Also research by Brent and Ptashne had shown that the LexA repressor from *Escherichia coli* functioned as a repressor of transcription in yeast [5].

In this project, a gene circuit was designed that utilized both of the IRP and LexA repressors, that resulted in repression at both transcription and translation levels, with a pheromone response pathway-inducible reporter gene. This circuit design is unique from previously published work linking reporter genes to the mating response, in that the circuit was designed to suppress the basal expression of the reporter gene in an OFF-state, and then simultaneously de-repress and trigger expression in an ON-state. The design of the circuit therefore reduces the level of background noise from the reporter gene through inhibition of basal expression, enabling a higher relative-fold increase in expression, compared with a pheromone response pathway-inducible reporter gene alone [1, 6, 7]. Also, at the time of writing this approach to boosting the output of a reporter gene, and combining transcriptional and translational inhibition in a gene circuit, had not been attempted.

To achieve these objectives, the project utilized a synthetic biology approach to building a system of interacting plasmids that function in combination as a module. The simultaneous interactions of multiple components produce complex, dynamic behaviours that are impossible to conceptualize without the aid of mathematical modelling and computer simulation [8, 9]. Synthetic biology incorporates aspects of computer science from systems biology to augment molecular biology with computer aided-design and enable the modelling of gene circuits as an engineer would design electronic devices [10–13]. Synthetic biology projects result in rounds of iterative design and development as models are constructed and used to guide experimental design. Parameter values are obtained experimentally and the model evolves alongside the engineered system [11, 14–16].

1.2 The Yeast *Saccharomyces cerevisiae*

The *Saccharomyces cerevisiae* genome was one of the first to be completely sequenced, and is a widely recognized model organism for studying the genetic systems of eukaryotes, providing rapid growth, dispersed cells, simple replica plating and mutant isolation, together with a well-defined genome [17–20]. *S. cerevisiae* is non-pathogenic,

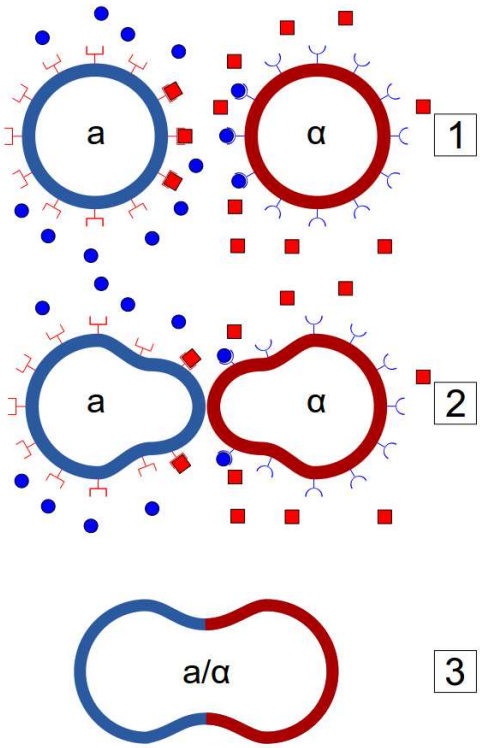

FIGURE 1.1: Diagrammatic representation of the yeast mating process. 1. Yeast cells produce mating pheromone that binds to receptors on cells of the opposite mating type. 2. Cells exhibit chemotaxis, and grow towards the opposite mating type. 3. The haploid cells fuse to form a diploid cell. Reproduced with permission from A. Fijalkowski [21].

can be handled with few precautions, and can be propagated easily and cheaply in large quantities giving rise to an ideal organism for biochemical studies [18].

1.2.1 Yeast Mating

Yeast cells exist in two distinct haploid forms, a and α with genotypes MAT**a** and MATα respectively, which can mate to form a diploid a/α cell (figure 1.1). The mating between Mat**a** and Matα enables genetic transfer within the yeast population and enables the colony to evolve through genetic recombination [20]. The diploid cells can continue to bud in vegetative growth until they encounter starvation conditions, at which point the cells undergo meiosis and sporulation, re-establishing the haploid phase [22–29].

Each cell type produces a 13 amino acid peptide pheromone protein that binds to specific receptors on the surface of the opposite mating type; *MAT*α cells produce α-factor (WHWLQLKPGQPMY) and binds to *MAT***a** cells [26]. *MAT***a** cells produce

a-factor (YIIKGVFWDPAC) that binds *MAT*α cells [22, 25]. Binding of one of these pheromone proteins to a cell of the opposite mating type halts the cell cycle, induces changes in cell morphology, and prepare the cell for cytoplasmic and nuclear fusion [29–34]. This signalling system in yeast has become one of the most well characterized signal transduction and developmental systems, and nearly all of the pathway has now been extensively documented through molecular genetics, cell biology, and biochemistry studies [35–42].

Yeast are non-motile organisms and therefore require some mechanism to orient themselves into close proximity with cells of the opposite mating type [18]. In order to achieve this, yeast exhibit a chemotropic response to pheromone secreted by the opposite mating type through asymmetric cellular organization, directing their growth towards the mating partner [29, 32, 34, 43]. Yeast demonstrate the ability to polarize their actin cytoskeleton in the direction of the site of highest pheromone concentration [34]. The cells elongate towards the mating partner, forming a structure termed a “mating projection”, containing proteins involved in signalling, polarization, cell adhesion, and fusion, causing the cells to take on a “pear-like” morphology known as a “Shmoo” (figure 1.1) [34, 38, 43].

The yeast mating response involves a complex cascade of events that enable yeast to translate changes in environmental conditions into an appropriate genetic and metabolic response [31, 44]. The mating response is an intracellular signal transduction pathway comprising a trans-membrane spanning heterotrimeric G-Protein-coupled receptor, and a mitogen activated protein kinase (MAPK) cascade which activates transcription factors for genes that enable the appropriate genetic response in the nucleus to a particular input stimulus at the cell surface receptor [33, 45–47]. Understanding the interactions and dynamic behaviour of the cascade is important when building gene circuits that use the mating response pathway as a generic “signal processing module” [1, 2, 46].

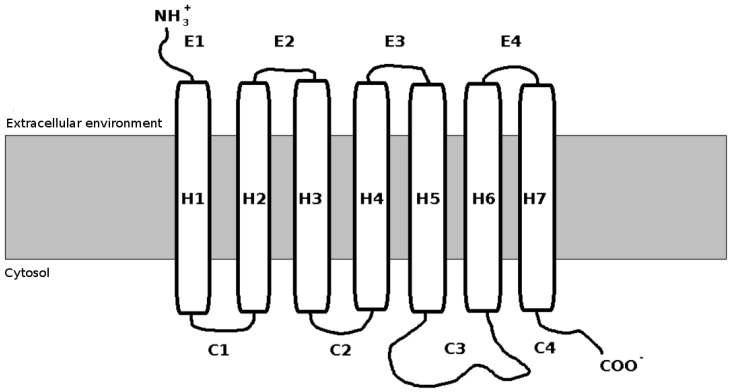

FIGURE 1.2: Diagram of the yeast Ste2 G-protein receptor demonstrating the 7-transmembrane domains (H1-H7,), 4 extracellular domains (E1-E4) and 4 intra-cellular domains (C1-C4), an intracellular loop for G-protein coupling, and a cytoplasmic carboxy-terminal domain.

1.2.2 Pheromone Receptor-G-protein Coupling

*S. cerevisiae MAT***a** cells express the Ste2 α-factor binding receptor, and *MAT*α cells express Ste3 a-factor binding receptor [22, 25, 26, 32]. The pheromone receptors have a structural topology of seven trans-membrane domains, a third intracellular loop that is involved in G-protein coupling, and a cytoplasmic carboxy-terminal domain that mediates ligand-induced endocytosis and desensitization (figure 1.2) [38]. Hundreds of G-protein coupled receptors have been identified in eukaryotic cells, responding to a variety of stimuli such as hormones, neurochemicals, light, odours, and tastes [48, 49]. G-proteins constitute a large proportion of known drug targets, as the released G-proteins elicit biochemical responses, and changes in cellular physiology by stimulating a variety of target (effector) enzymes [50–52]. G-protein receptors share a common design consisting of 7 membrane spanning regions linked to the G-protein [33]. In yeast, the G-proteins are formed from three subunits - Gα (Gpa1), Gβ (Ste4), and Gγ (Ste18) [53, 54]. Gβ and Gγ act as a heterodimer Gβγ [48], and Gα subunit interacts Gβγ to form the inactive Gαβγ trimer (figure 1.3 1.) [33, 55]. A superfamily of G-protein subunits has been identified in eukaryotes comprising 17 distinct Gα, 5 Gβ, and 6 Gγ isoforms, allowing for many combinatorial possibilities for cell receptors [49]. G-proteins are activated when a ligand molecule binds to the linked surface receptor, in this case the mating pheromone from the opposite cell type, causing a conformational change in the receptor that is transmitted to the G-protein causing the Gα subunit to exchange

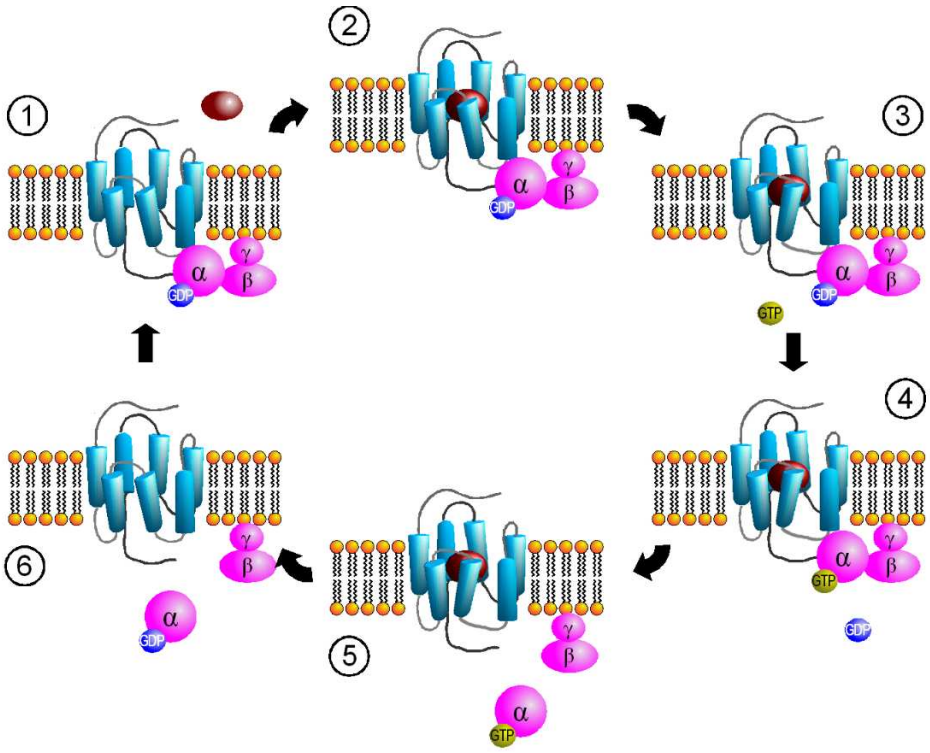

FIGURE 1.3: Diagrammatic representation of the process of G-protein activation. 1 and 2. Pheromone binds to the extracellular Ste2 receptor and is internalized by the receptor. 3 and 4. GDP is exchanged for GTP and the α subunit disassociates from the Gβγ units, resulting in activation of the pheromone response pathway. 5 and 6. During the process of switching off the pheromone response pathway, the Gα subunit binds GDP in place of GTP and re-associates with the Gβγ subunits.

GTP for GDP and disassociate from Gβγ exposing the effector binding regions of Gβγ [38, 42, 56, 57] (figure 1.3). The released Gβγ is then able to participate in a 3 level (MAP) kinase cascade that quickly transmits the pheromone binding signal through the cell to the nucleus [48, 51, 54, 58, 59] (figure 1.3 5, and figure 1.4 right.). The Gα subunit is released from the inner membrane into the cytoplasm [38]. The Gβ subunit has been shown to be most significant in activation of the signal response, while Gγ has been found to contain a conserved cysteine-aliphatic-aliphatic-X motif at the carboxy terminus that is thought to localize the Gβγ subunits to the membrane [33].

1.2.3 Pheromone-Induced G-protein Activation

The free Ste4 Gβγ subunit interacts with three effectors: Ste5/Ste11 complex, Ste20 protein kinase, and Far1/Cdc24 complex via a binding site that was previously buried

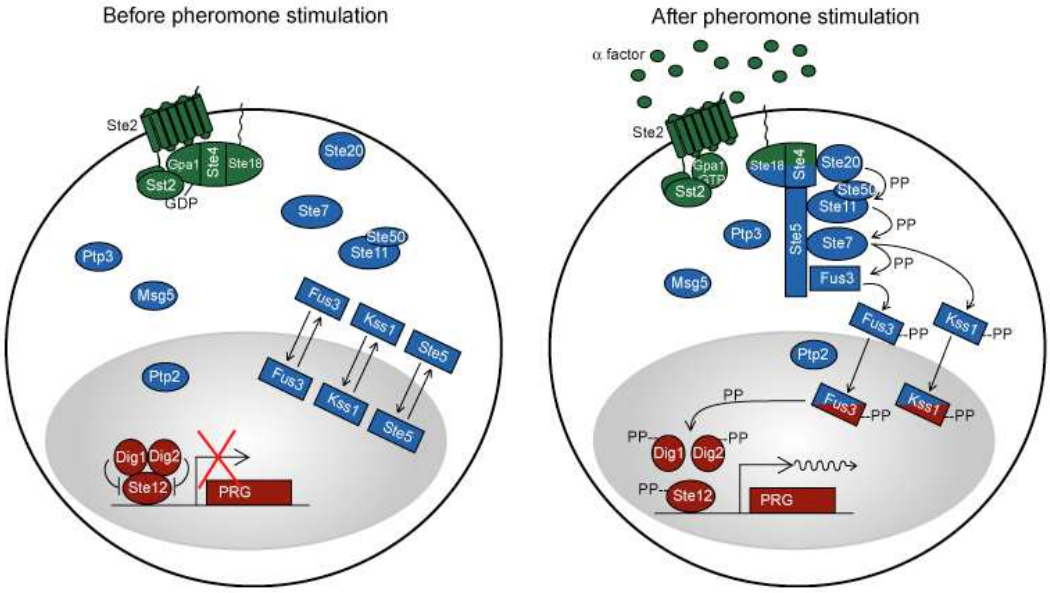

FIGURE 1.4: Diagrammatic representation of the main components of the yeast pheromone response pathway. Left is the inactive pathway prior to pheromone activation. Right is the pheromone stimulated, active pathway. The activated G-protein results in the phosphorylation of Ste20, subsequently resulting in the sequential phosphorylation of Ste11, Ste7, and Fus3, forming the MAPK cascade. The MAPK cascade communicates the pheromone receptor binding event through the cytosol to the nucleus where the appropriate mating response genes are up-regulated via the de-repression of the Ste12 transcription factor by the phosphorylated Fus3. (Image from yeastpheromonemodel.org [61])

within the Gα associated molecule [49, 55]. The Ste18-Gβγ complex anchors the βγ G-protein subunits to the inner cell surface by covalently attached lipid farnesyl and palmitoyl groups [54, 60]. The association of the Gβγ subunit with the inner cell membrane surface localizes the position of the mating response, and assists in orientating the cell towards the pheromone gradient, and the mating partner [25, 34, 54, 60]. Localization of the Gβγ subunit results in Ste20 moving in close proximity to Ste11 and Ste5, forming the initiating step in the signal cascade 1.4 [38]. Ste20 exists in an inactive form in the cytoplasm and is activated by a small 21kD, Rho-like G-protein Cdc42 [38]. Cdc42 in yeast has a similar amino acid sequence to members of the *Ras* super family and is known to be involved in the control of several morphogenetic events during the cell cycle, including the generation of cell polarity, development of normal cell shape, localization of secretion, and deposition of cell-surface material [62]. Cdc42 binds to the CRIB domain of the large N-terminal region of Ste20 that ordinarily sterically occludes and auto-inhibits the active kinase C-terminal

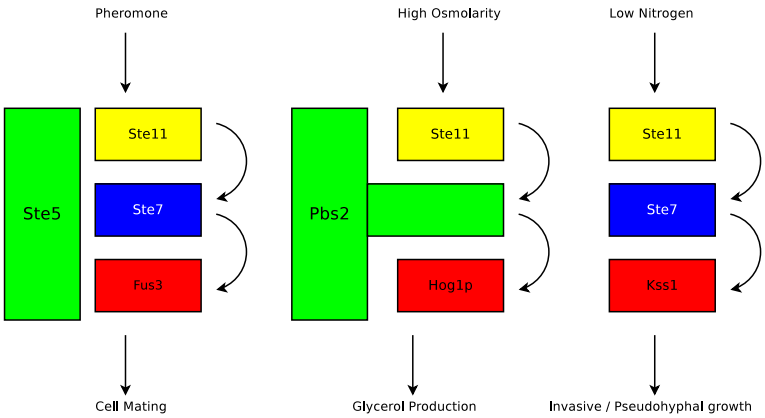

FIGURE 1.5: Schematic overview of yeast MAP kinase modules that share Ste11, adapted from Drogen F. [51]. Yeast signalling pathways are capable of sharing components but maintaining signal specificity through the use of pathway specific scaffold proteins. Ste11 and Ste7 are shared pheromone, high osmolarity, and low nitrogen response pathways but differentiated through the use of the Ste5 and Pbs2 scaffolds that ensure signal specificity and prevent cross-talk between physiological responses [51].

region, thereby activating Ste20 by permitting auto-phosphorylation of the exposed activation loop [63]. Cdc42 is also permanently tacked to the inner leaflet of the plasma membrane, assisting in localizing Ste20 to the membrane [63].

Ste5 is a large, multi-functional scaffold protein that, whilst having no catalytic activity, serves as a scaffold and binding platform for components of the MAP kinase cascade [63–67]. In yeast, the pheromone pathway scaffold Ste5 binds Ste11, Ste7, and Fus3, whilst the high osmolarity glycerol pathway scaffold Pbs2p interacts with Ste11 and Hog1 (see figure 1.5) [51]. There are common components in each pathway (figure 1.5) and the scaffold serves to insulate the signal, preventing cross-activation between signalling pathways [51, 68–71]. Choi, 1994 demonstrated by yeast two-hybrid analysis and co-immunoprecipitation that Ste11, Ste7, and Fus3 associate with different domains of Ste5 implying, that Ste5 simultaneously binds the components of the MAP kinase reaction. Ste5 initially forms an adapter between Gβγ and Ste11, bringing Ste11 into proximity with Cdc42-bound Ste20 at the plasma membrane, resulting in immediate phosphorylation of Ste11 by Ste20 [64–66].

The third effector to bind Gβγ is Far1, complexed with Cdc24 [55]. Far1 moves from the nucleus to the cytoplasm when cells are stimulated by pheromone, and interacts

with Gβγ at the cell membrane transiently via the MAP kinase cascade [34]. The N-terminal domain of Far1 contains a RING H2 domain that interacts with Gβγ, while the C-terminal end binds Cdc24, a guanine nucleotide exchange factor (GEF) that promotes exchange of GTP for Cdc42 [72]. GTP-bound Cdc42 is then able to bind to Ste20 and several other regulators of cell polarity and the actin cytoskeleton [43]. Far1 is a multi-functional regulator of the mating process. One function is to bind to Cdc24 and facilitate growth towards the mating partner, another function is to mediate pheromone-induced cell cycle arrest [32]. Chang demonstrated that Far1 ("factor arrest") is a non-essential gene, induced 4 to 5 fold by pheromone-induced Ste12 which, in turn, interacts with Cdc28 cyclin-dependent kinase, the master regulator of the cycle growth phase [32]. This interaction provides the link between the signal transduction pathway and cell cycle arrest under pheromone stimulation [73]. The exact mechanism of how Far1 inhibits the cell cycle is unclear. However Pi and Gartner determined that Fus3-mediated phosphorylation of Far1 is required for cell cycle arrest [28, 72].

1.2.4 The MAP Kinase Cascade

The MAP kinase cascade is the most prominent signalling mechanism in yeast, facilitating a rapid response to extracellular stimuli [34, 40, 43]. MAP kinase cascades are found ubiquitously in eukaryotic organisms functioning in cell growth, differentiation, tumorigenesis, and stress responses [40, 46, 74]. MAP kinase pathways usually consist of three protein kinases that act in series: a MAP kinase kinase kinase (MAPKKK), a MAP kinase kinase (MAPKK), and a MAP kinase (MAPK) (figure 1.4) [39, 43, 75, 76]. When the cascade is activated, the MAPKKK phosphorylates the MAPKK, which in turn phosphorylates the MAPK [77]. In yeast, the MAPKKK is Ste11 and the MAPKK is Ste7, and there are two MAPKs: Kss1 and Fus3 [47, 74, 78] (figure 1.4 and figure 1.5 left). The MAPK usually serve to regulate transcription factors by MAPK-mediated phosphorylation, and many intracellular and extracellular signals modulate transcription of specific genes through activation or inhibition of MAPK cascades [28, 35, 79].

1.2.4.1 Ste11, Ste7 and Fus3

The MAPK cascade function is facilitated by Ste5 and Ste20 (figure 1.4) [39, 67]. Ste11 bound to Ste5 is activated by Ste20 and subsequently activates Ste7 by phosphorylating a threonine residue in the Ste7 activation loop [47]. Ste7 does not bind strongly to the Ste5 scaffold, but binds with high affinity to Fus3 and Kss1 [64]. Ste7 contains a highly-conserved catalytic domain and a less conserved N-terminal domain, in which the first 20 amino acid residues form the MAPK-binding/docking site (D-site) [27]. Ste7 activates Fus3 and Kss1 by phosphorylating threonine and tyrosine residues in their activation loops [64, 65]. The MAPK's Fus3 and Kss1 are proline-directed kinases and phosphorylate their targets on serine or threonine residues that are immediately followed by a proline and primarily target the Ste12/Dig1/Dig2 transcription factor complex as well as Far1, and both can activate Ste12, demonstrating functional redundancy [74]. Fus3, however can also activate Ste7 and Ste5, and can phosphorylate Far1, whereas Kss1 cannot [65]. Bardwell hypothesizes that this redundancy provides overlapping reinforcing contributions to the activation of the MAPKs so that a loss of the mating response is only observed when multiple links are severed simultaneously ([55]. A recent publication by Malleshaiah *et al.* revealed that a phosphatase Ptc5 competes with Fus3 for phosphorylation sites on Ste5, facilitating a switch-like response in the mating pathway and ultra-sensitivity to pheromone [80].

1.2.4.2 Ste12 and The Pheromone Response Element

Ste12 is a protein consisting of 688 amino acids with an N-terminal DNA-binding region providing its function as a transcriptional activator, enabling it to form a protein-DNA complex specifically with the genes it regulates [28, 81, 82]. Genes up-regulated following pheromone induction all contain a common pheromone response element (PRE) with sequence $5'$-ATGAAACA (or sometimes reported as $5'$TGAAACA) [81]. The PRE sequence is found in over 200 genes associated with cell mating [83, 84], of which over 100 are induced two-fold by the pheromone response pathway [28, 73, 82, 85, 86].

Yuan and Fields partially characterized the DNA binding domain of Ste12, localizing the minimum region to 164 amino acids near the N-terminus between amino acids 41 and 204 [86]. They also found an N-terminal domain can bind cooperatively to two copies of the PRE in a manner independent of the orientation, binding head-to-tail or tail-to-tail with variable spacing between the two elements. [86]. Kirkman-Correia *et al* located the transcriptional activation domain at the C-terminus (residues 384-688), and it has been shown that deletion mutants lack the ability to activate basal and induced transcription of PRE genes, however, only region 255-354 is required for pheromone-induced transcription [87].

Ste12-dependent, pheromone-induced genes include positively-acting components of the mating pathway (Ste2, Fus3, and Far1), together with negative feedback regulators of the pathway (Sst2, Msg5, Ptc1, and Gpa1), as well as genes involved in the process of cell fusion (Fus1, Fus2, Fig1, Fig2, Aga1) [28, 80, 88]. Ste12 has been shown to up-regulate its own transcription during pheromone response and can also work in conjunction with other transcription factors, in particular Tec1p, forming a heteromultimer with Ste12 regulated by Kss1 [68, 89], and this complex guides Ste12 to specific genes in the filamentous growth pathway [85, 90].

1.3 Switching Off The Pheromone Response

A natural property of G-protein signalling systems is the ability to attenuate the response following prolonged stimulation [91]. Haploid cells that do not mate and form diploids must return to the vegetative growth state [92, 93]. It has been observed in many signalling systems, and particularly with G-protein-coupled receptors, that prolonged signal exposure results in desensitization of the response [91, 94–96]. This attenuation of signal response involves a number of complex mechanisms that are activated within minutes of receptor activation, and these mechanisms are thought to be responsible for attenuation in response to light, colours, odours, chemical stimulants and narcotics [91]. Unlike the detailed information that has been accumulated about the activation and response of the mating pathway, there is much less understanding

of the mechanisms involved in switching it off [43, 90, 92, 97]. The yeast α-factor desensitization mechanism is similar to hormone desensitization in animal cells and receptor desensitization has been extensively studied in the vertebrate β-adrenergic and rhodopsin receptors [91]. In yeast, there are a number of negative feedback mechanisms that facilitate control of the mating response.

The four main mechanisms employed to attenuate the pheromone signal are: pheromone degradation, pheromone de-sensitization, phosphorylation of the Gβ subunit, and dephosphorylation of the Fus3 MAPK by a phosphatase encoded by Msg5 [92]. Chan and Otte screened for genes involved in the desensitization and recovery from the mating response [98, 99]; by screening for haploid cells that were hyper-sensitive to pheromone-induced cell-cycle arrest, they discovered two classes of super-sensitive mutants designated *Sst1*Δ and *Sst2*Δ [98]. *Sst1*Δ mutants are allelic for the gene Bar1 which encodes a 587 amino acid endoprotease and cleaves α-factor, inactivating the pheromone and forming a negative feedback loop in the pheromone response pathway [1, 99]. *SST1*Δ mutants demonstrate hyper-sensitivity to pheromone and are slow to recover from G_1 growth arrest [100].

SST2Δ mutants are unable to degrade α-factor pheromone and cannot recover from cell cycle arrest [99]. The SST2Δ mutants defined a novel gene that was the first discovery of the RGS ("regulator of G-protein signalling") factor family. RGS factors are negative regulators of G-proteins, so called because they stimulate the hydrolysis of the GTP bound active form of the Gα subunit, back into inactive GDP bound Gα [49, 101]. Sst2 stability is increased by phosphorylation by Fus3 and has been shown to increase the rate of hydrolysis of active Gα by at least 20-fold [55]. The activity of Sst2 serves to complete the G-protein cycle by sequestering free Gβγ subunits thereby forming a second negative feedback loop, terminating signal response, and restoring the pool of inactive Gαβγ [52].

Chen reported in the absence of ligand, the Ste3 receptor is subject to rapid degradative endocytosis [56]. However, when bound to a-factor pheromone, the receptor transcription is up-regulated and subjected to a process of recycling, whereby the ligand is degraded prompting ligand disassociation and re-utilization at the membrane surface

[56]. This ensures that a suitable quantity of receptors are available throughout the mating response and, more importantly, receptor expression is focused at the point of pheromone contact, facilitating the chemotropic response up the pheromone gradient [56]. This response is not as prominent in the α-factor stimulated Ste2 receptor, where pheromone stimulation increases Ste2 vacuole-directed transport and degradation [96, 102]. Dohlman *et al* reports Ste2 desensitization occurs through binding of the RGS protein Sst2, such that it is positioned in close proximity to Gpa1 [31].

Research suggests that the Gα subunit of the heterotrimeric G-protein has a positive signalling role and is responsible for pheromone desensitization and recovery back to the vegetative haploid growth stage [49, 92, 103, 104]. GTPase-deficient Gpa1 mutants demonstrate constitutive expression of pheromone response elements and morphological changes in the absence of pheromone [104, 105]. Kurjan, 1991 introduced mutations in the SCG1 gene, encoding the Gα subunit and observed defects in mating response, growth and cell morphology [105]. Dohlman and Thorner later found that inactivating mutations in the Gα gene Gpa1 do not block pheromone response, but result in constitutive signalling and it has been concluded that this is due to uncontrolled pathway activation by free Gβγ [49]. It was also found that over-expression of Gα leads to diminished signal transduction due to over-sequestration of Gβγ [33, 101, 106]. Cole *et al* also demonstrated how over expression of Gα subunit leads to suppression of the mating response, and represses the response even when over expressing the Gβ and Gγ subunits [33]. Deletions in either of the Gβ or Gγ genes results in pheromone insensitive sterile cells, whilst over-expression leads to constitutive activation of the mating pathway [33, 49, 106]. It has also been shown by Cole *et al* that over expression of Ste4 (Gβ) with expression of Ste18 (Gγ) promotes constitutive activation of the pheromone signalling pathway [33, 101].

Blackwell *et al* reported that Msg5 works in concert with Gα to down-regulate the mating signal by inhibiting the pheromone-induced increase of Fus3 in the nucleus [107]. Doi *et al* earlier reported that Gα may induce a post translational modification of Msg5 resulting in enhanced protein phosphatase activity or that Gα may induce transcription of Msg5. The MSG5 nucleotide sequence suggests that it functions as a PTPase; an enzyme group that specifically dephosphorylates phosphotyrosyl residues

in selected proteins [108]. Both Fus3 and Kss1 require tyrosine phosphorylation for activation, making them potential targets for Msg5, and a GST-Msg5 fusion protein has been shown to dephosphorylate and deactivate *in vitro* phosphorylated Fus3 [108]. Doi *et al* went on to show that epistatic interactions imply that Msg5 functions between Ste11 and Ste12, disruption of the Msg5 gene enhances Fus3-dependent kinase activity, and over expression of Msg5 suppresses pheromone-induced modification of Fus3 [108].

In addition, studies by Straton *et al* demonstrated the Gα subunit functions as a slow negative feedback function on activation of the signal pathway by Gβγ and interacts with an effector molecule, stimulating an adaptive signal that decreases sensitivity to pheromone over time and eventually shuts off the mating response downstream of the receptor [92]. This signal is delayed relative to the mating signal and through observations using two-hybrid analysis, does not involve sequestration of Gβγ [92]. Zhou *et al* demonstrated how, under low pheromone conditions Gα interacts with the GTPase activating protein Sst2, stimulating Gβγ sequestration [109]; through pheromone concentrations sufficient to halt the cell cycle, Gα functions as an adaptive mechanisms to recover the cell from the mating response [109]. Metodiev *et al* using GST-tagged Gα protein with glutathione-agarose pull-down experiments, 2D-gel electrophoresis, and mass spectrometry, found Gα associated with the phosphorylated form of Fus3 [110]. Histidine-tagged Fus3 applied to a nickel column also demonstrated binding to Gα [110]. A number of hypotheses were presented by Metodiev *et al* to explain the association of Gα with Fus3. Firstly, it is thought that Gα might anchor Fus3 to the membrane and restrict it from transmitting the mating signal to the nucleus, and secondly, the active form of Gα is targeted for degradation and interaction with Fus3 might include the kinase in this degradation process [110]. Finally, Metodiev *et al.* postulated that in cells exposed to a physiological gradient of pheromone, Gα recruits Fus3 to the mating projection site where the kinase phosphorylates Gβγ, which promotes assembly or stabilization of the Gβγ-Far1 complex required for chemotropic growth [110]. The function of Gα presents a paradox, in that Gα acts as both an effector and inhibitor of Fus3 activity in the mating pathway [106, 110, 111].

In conclusion, published research has revealed the mating pathway as not a simple linear chain of events from pheromone stimulation of the membrane receptor through the MAP kinase cascade to the transcription factors. Instead there is a subtle interplay of secondary messengers and auxiliary effectors fine tuning the interaction of the major components, ensuring an appropriate level of response and timely recovery from the mating process. [80, 90, 97, 106, 109, 110].

1.4 Modelling The Mating Pathway

Systems biology is an emerging scientific field that undertakes a holistic approach to understanding biological processes through the interactions of the component parts [14]. Systems biology seeks to gain an understanding of the functions of biological systems, using methods that cannot be described by studying the component macromolecules in isolation, and consequently requires interaction between diverse experimental fields and datasets to arrive at this understanding [9, 14, 112–114]. The physiological response of cells to internal and external stimuli is governed by a complex set of interacting genes and proteins with non-linear reaction kinetics and pathway fluxes [113, 115]. Recent advances in theoretical biology have shown that biological reaction networks can be accurately modelled using mathematics [116–118], and these models can provide understanding of the principles of biological control systems as well as predictions that can be varied experimentally in the laboratory [14, 115, 119]. Figure 1.6 illustrates the cycle of systems biology research, employing an iterative process of computational and experimental science to explore complex biological problems through modelling and systems analysis. The model provides hypotheses for experimental research, which produces data that feeds back into the model, driving new understanding and further hypothesis generation.

This iterative modelling and experimentation approach has been applied to the yeast mating pathway by a number of researchers [46, 75, 76, 120]. As mentioned earlier, the yeast mating pathway is a well-characterized system that is easily modified and has a number of discrete and accessible quantifiable behaviours making it a favourable

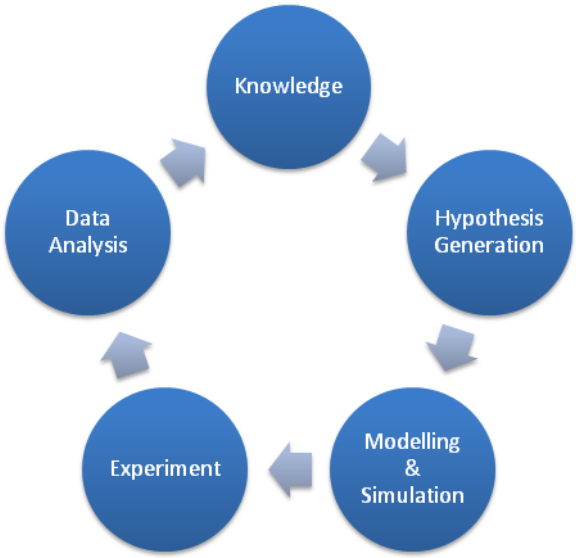

FIGURE 1.6: Hypothesis driven research in systems biology, adapted from Kitano *et al* [14]. Hypotheses can be explored using "dry science" techniques of modelling and systems analysis. Models can be parameterised through experimentation and data analysis which leads to further knowledge and understanding, leading to further hypothesis. [14]

target for mathematical modellers [77, 121–125]. In the following sections the key publications in modelling budding yeast pathways will be discussed.

1.4.1 Chen *et al* (2000): Kinetic Analysis of Budding Yeast Cell Cycle Model

Chen *et al* have developed a kinetic model of the cyclins Cln1-3 and Clb1-6 which have been shown to coordinate the events of the cell cycle: DNA synthesis, bud emergence, spindle formation, nuclear division, and cell separation in yeast [56]. Chen converted the established mechanisms of cyclin synthesis and degradation into a set of differential equations, describing the time courses of three major classes of cyclin-dependant kinase activities [56]. The model was then used to examine the molecular events controlling the initiation of chromosome replication, bud formation, and mitosis (the "start" of cell division) and also the transition through metaphase to anaphase (the "finish" steps of cell division) in both wild-type and a selection of mutants [56]. After refining the model based on laboratory experimental data, the model included 10 non-linear ordinary differential equations for the cyclins, and their associated proteins,

three algebraic functions for transcription factors, three "integrators" to trigger DNA synthesis, budding, and spindle formation, and a simple rule for separating mother and daughter cells at division [56]. The model includes approximately 50 parameters that are fitted to the phenotypic behaviour of yeast and require further optimization, but are sufficient to account for the properties of cell cycle control in yeast [56].

1.4.2 Yi *et al* G-Protein Model

Yi *et al* quantitatively characterized the G-protein cycle in yeast, based on direct *in vivo* measurements using fluorescence resonance energy transfer (FRET) [126]. A cyan fluorescent tagged protein (CFP)-Gα, and yellow fluorescent protein (YFP) tagged G$\beta\gamma$ were used to observe a reduction in FRET when the receptor was stimulated with pheromone, causing the G-protein to disassociate [126]. Time course experiments were performed and data was obtained on how Sst2 and the C-terminal tail of the α-factor receptor, modulates the kinetics of G-protein signalling. The data used to build a quantitative model to estimate the *in vivo* rates of G-protein activation and deactivation in yeast [126].

The model validated existing observations that the majority of the control of the mating pathway resides at the G-protein cycle [126]. The work published by Yi *et al* also found that G-protein activation, transcriptional induction, and cell-cycle arrest responded with the same $K_{0.5}$ value for pheromone dose response, and aligns with observations in mammalian G-proteins, where K_d for receptor-antagonist binding and IC_{50} values for inhibiting the corresponding physiologic downstream processes overlap. The work by Yi *et al* provides quantitative evidence that the overall G-protein cycle determines the dose response of G-protein systems, not just the receptor ligand dynamics.

1.4.3 Hao *et al* RGS Protein Pheromone Desensitization Model

Hao *et al* published a model of the activation, desensitization, and re-sensitization steps of the mating pathway, following pheromone induction [127]. The study used

a combination of experimental and mathematical techniques. Hao *et al* investigated how external signals produce responses inside the cell, specifically G-protein activation and desensitization by the pheromone receptor and the RGS proteins (Sst2 and Gα). Radio-ligand binding measurements were used to measure receptor expression, while quantitative immunoblotting on whole cell extracts was used to quantify Sst2 and Gα. Expression levels were measured for wild-type and also in mutants engineered to over-express Ste2, Gα, Gβγ, Sst2, and Ste18, and changes in protein level were measured with immunoblotting. [127]. The functional changes brought about by altered expression were investigated using a reporter transcription assay comprised of a pheromone-responsive promoter (*FUS1*) fused to *LacZ* (β-galactosidase) [127]. Experimental data was used to build a mathematical model of the pathway using differential equations (equation 1.1 and figure 1.7)

An overview of the model is presented in figure 1.7. Hao *et al*'s model simulates the pathway activation and inactivation with two coupled ordinary differential equations, and assumes that free Gβγ activates the expression of pheromone response genes and RGS protein switches the pathway off by attenuating the amount of free Gβγ through recombination with Gα. The model provided predictions that could be compared with the experimental results and both correlated a sharp rise in Gβγ during pheromone stimulation and a slower increase in Sst2. A mathematical expression for the response of the signalling pathway was derived from the model using response R of the signalling pathway as a function of pheromone concentration $[L]$, where R_{min} is the response in the absence of pheromone and R_{max} is the maximum response and C in terms of $[L]$, steady state RGS concentration [127]. (equation 1.1).

$$R = \frac{R_{min}C + R_{max}[L]}{C + [L]} \tag{1.1}$$

The model did not predict a reduction in the mating response when over-expressing Sst2 however, which prompted a second round of experimentation with a GFP-tagged *LacZ* reporter and individual cells assessed by flow cytometry to assess wild-type and mutant cells with over-expressed Sst2. Following 90 minutes exposure to α-factor, the wild-type displayed a small intensity peak in fluorescence, which diminished

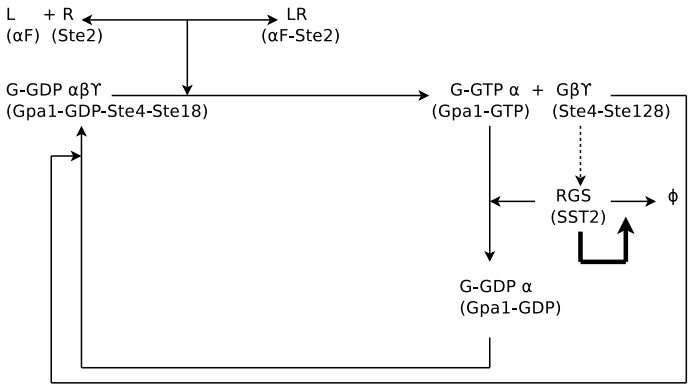

FIGURE 1.7: Pathway regulation by RGS and Gβγ proteins. Upon binding of the ligand (L) α-factor (αF) to its receptor (R, Ste2), the G-protein α subunit (Gpa1) releases GDP, binds to GTP, and liberates the G-protein βγ subunits (Ste4/Ste18). Sustained signalling requires activation of multiple effectors (not shown) by the dissociated Gβγ components. These effectors activate a pathway (dotted line) leading to transcription of several genes including the RGS protein Sst2. GTP hydrolysis is accelerated by the RGS protein, and this leads to subunit re-association and pathway inactivation. The model assumes that GTP hydrolysis is the rate-limiting step of subunit reassembly. A potential positive feedback loop leading to Sst2 degradation is indicated by a darker line. Adapted from Hao *et al*.

and was replaced by a second peak of higher intensity, however the lower intensity peak continued in the Sst2 mutant cells [127]. The authors hypothesize that when Sst2 is over-expressed, the graded response to α-factor is replaced with a binary response through positive feedback regulation where a second feedback loop promotes degradation of Sst2 [127]. Implementing the experimental observations into the model, the author was able to explain the slow initial induction of Sst2, as seen in Sst2 over-expression mutants. The model was used to show that alterations in the expression of Sst2 occur slower than alterations in the active state of the G-protein such that the activation state of the G-protein adjusts rapidly to the slow change in Sst2 expression and is therefore in equilibrium, allowing the state of the system to be determined by Sst2 expression levels [127]. Degradation of Sst2 occurs at a constitutive rate when pheromone is absent and is proportional to expression (a “constitutive steady-state”). Following pheromone induction, the rate of Sst2 production exceeds the constitutive degradation rate and as Sst2 levels increase, expression becomes inhibited by increasing levels of inactive G-protein and the production rate reduces back to the constitutive rate, forming a positive feedback loop. In over-expression mutants the level of Sst2 degradation falls much more rapidly than the production rate and the rate of production

rises until it reaches a new steady state, correlating with the experimental observations [127]. In addition, the model was modified with a stochastic term to account for random fluctuations in protein concentration. 10,000 simulations were run using the new random model, averaged for a mean time course and the results demonstrated the binary behaviour observed in the experimental work for the Sst2 over-expression mutant [127]. To confirm that pheromone stimulation promotes Sst2 degradation, cells were grown for an hour in presence of α-factor, treated with cyclohexamide to block further protein synthesis, and the remaining Sst2 protein was monitored with immunoblotting. Results demonstrated a faster reduction in Sst2 when pretreated with pheromone, as predicted by the model [127].

The work published by Hao *et al* demonstrates the iterative process of mathematical modelling combined with wet lab experimentation described by Kitano *et al*. The modelling component of the work provided unique insights into the biological interactions and hypothesis generation that could not be derived through experimentation alone, such as the positive feedback loop that facilitates re-activation of the pathway that had not been observed prior to the study. The model also developed from a simple mathematical derivation of the activation and de-activation of the signalling pathway, to an *in vivo* representative simulation of the mating response, eventually including the feedback loops that provide the timing and coordination for controlling the pathway response.

1.4.4 The Kofahl and Klipp Yeast Pheromone Pathway Model.

Kofahl and Klipp published a mathematical model of the dynamics of the pheromone pathway in haploid yeast cells of mating type MATa after stimulation with α-factor [128]. Yi *et al* and Hao *et al* modelled specific aspects of the yeast pheromone pathway to augment the specific areas of their research [126, 127]. Prior to Kofahl and Klipp there was no single model that attempted to simulate the complete pheromone pathway and concatenate the research conducted in this field. The Kofahl and Klipp model consists of a set of coupled differential equations that describe the transmission of the mating signal from the surface receptor, through the G-protein, to the MAP kinase

cascade, and activation of the Ste12 transcription factor [128] (see figure 1.8). The model includes:

- activation of the membrane-bound pheromone receptor.
- activation of the G-protein.
- formation and activation of the scaffold-bound MAP kinase cascade.
- activation of transcription factor Ste12.
- downstream effects on gene expression alteration and preparation for mating.
- down regulation of the signal process through Sst2 and Bar1.

The Kofahl and Klipp model was not part of a combined wet and dry experimental project as with the work by Yi *et al* and Hao *et al*, but the authors used parameter values obtained from published literature to fit the model behaviour to experimental observations of the changes in the relative levels of the mating response pathway components over time [126–128]. The model attempts to provide the most complete representation of the yeast mating pathway, in terms of including all of the interactions between the known components and the available kinetic data (figure 1.9). [128]. The model groups the reactions that comprise the yeast mating response into a series of complexes which represent the temporal order of events of the pathway including receptor activation, the G-protein cycle, Ste5 complex formation, and down stream effects of phosphorylated Fus3 and Far1 (figure 1.9). Using the model to investigate mutant phenotypes, Kofahl and Klipp were able to demonstrate the pheromone desensitization response of yeast cells to prolonged pheromone exposure. Cole *et al* showed that over expressing Gα resulted in five times the normal level of α-factor required to induce a mating response and can compensate for super sensitivity to pheromone observed in Sst2 and Ste2 mutant strains, resulting in pheromone desensitization [33]. Cole hypothesized that this was due to increased Gα mopping up available G$\beta\gamma$ and preventing progression of the mating signal through to the MAPK cascade [33]. This observation was also observed in the Kofahl and Klipp model where an increase in Gα creates a decrease in free G$\beta\gamma$, resulting in shortened complex

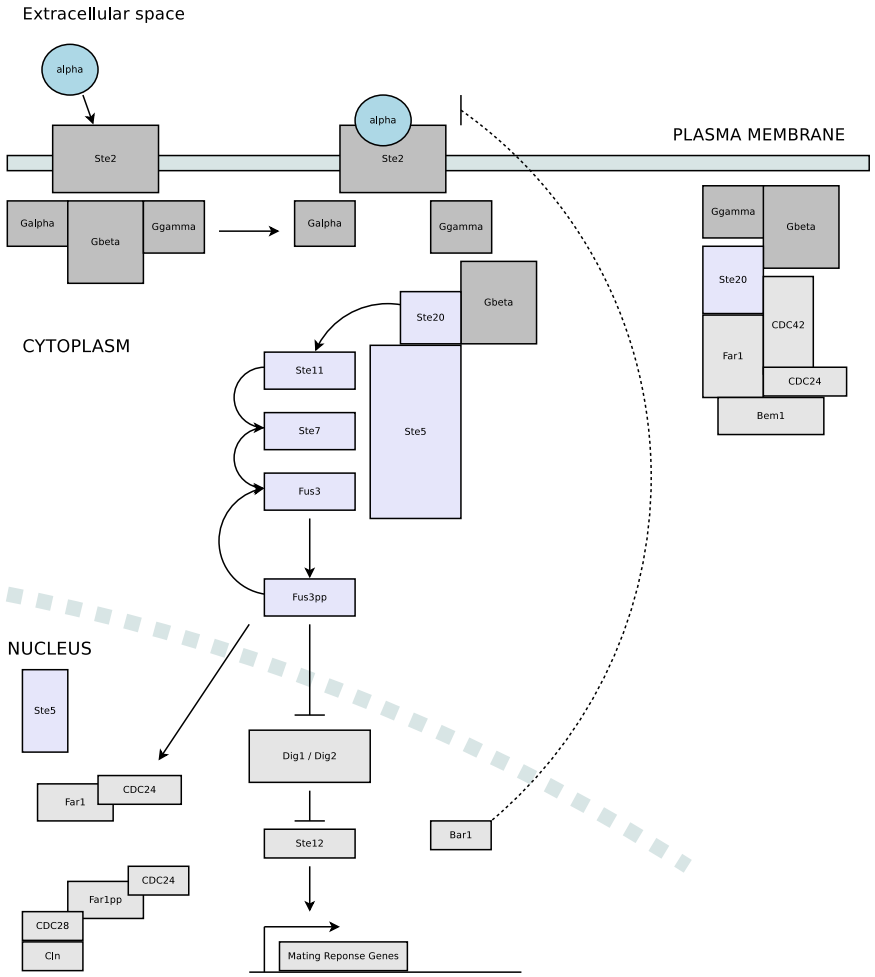

FIGURE 1.8: Spatial diagram of the pheromone pathway in yeast. Adapted from Kofahl *et al* [128]. The α-factor pheromone binds to the Ste2 receptor in the membrane, which is close to the heterotrimeric G-protein (middle). The G α subunit disassociates from the Gβ and Gγ sub-units. The Gβ and γ subunits are bound by Ste20 and to Ste5 which functions as a scaffold for the sequential phosphorylation of the MAPK cascade components Ste11, Ste7, and Fus3, as well as Cdc24 and Bem1 (right). Elements of the MAPK cascade shuttle to and from the nucleus (Fus3 and Bar1). Fus3pp phosphorylates Dig1 and Dig2 resulting in de-repression of the transcription factor Ste12 which initiates transcription of the mating response genes resulting in the up-regulation of over 200 genes (bottom).

formation, reduced Fus3 phosphorylation and Far1PP-Gβγ, and eventually reduced pheromone sensitivity [128]. The published role of the G-protein components Sst2, Ste12, Ste11, Ste20, Msg5, and Far1 were replicated in the model and used to validate its response to observed phenotypic changes [128].

The Kofahl and Klipp model also produces the same quantitative results as the Yi model for the G-protein cycle, however it does not include the feedback loops developed by Yi *et al.*, but replicates the observed behaviour [126, 128]. The model incorporates

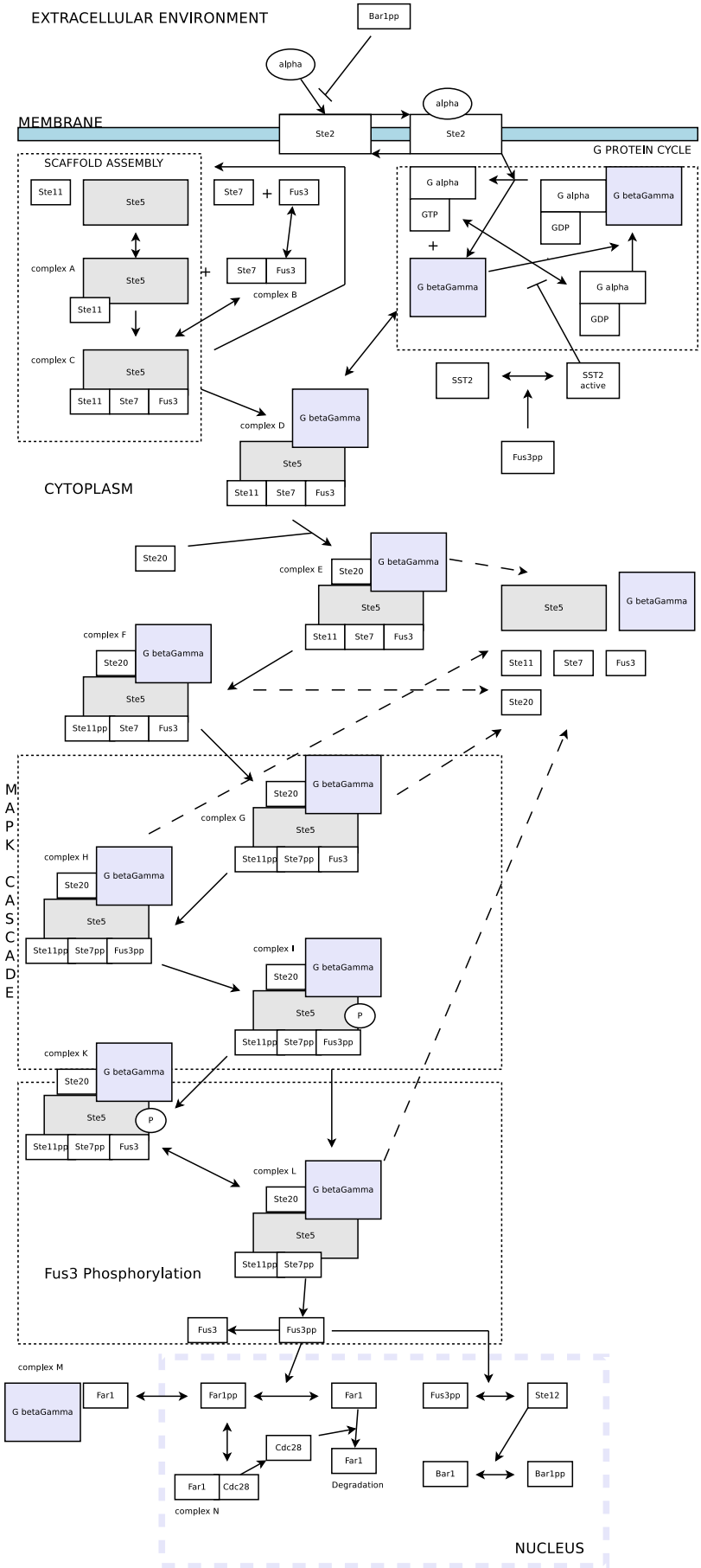

FIGURE 1.9: Diagram of the reactions modelled in the Kofahl and Klipp model. The model includes reactions for G-protein cycling, assembly of the MAPK scaffold, and sequential phosphorylation of the Ste11, Ste7, and Fus3 kinases. The model results in the activation of the transcription factor Ste12, as well as the Sst2 and Bar1 negative feedback components [128].

regulatory control with several feedback loops. Phosphorylated Fus3 activates Sst2 which stimulates hydrolysis of GαGTP, closing the G-protein cycle [128]. Also, the transcription and activation of Bar1 results in the degradation of α-factor resulting in down regulation of the pathway and negative feedback under prolonged pheromone stimulation [128].

Although many models of MAP kinase cascades have been published, they are not parameterised with data from yeast, and do not include yeast mating pathway-specific features such as the Ste5 scaffold and the Dig1/Dig2-Ste12 activation complex. Parameters are often obtained from studies of MAP kinase cascades in Xenopus species [120, 122, 124] and focus on the phosphorylation of the three kinases in isolation. The models also do not incorporate more recently identified components of the pathway such as Ptc1, observed by Malleshaiah *et al* [80]. It was hypothesized that the non-linear behaviour of the yeast mating pathway may influence the experimentally observed behaviour of the gene circuit. The inclusion of an ultra-sensitive cascade relevant to the yeast cells in which the circuit is embedded may be more useful for predicting experimental observations than a simplified version of the Kofahl and Klipp model that does not incorporate any of the dynamic behaviours of the cascade.

The Kofahl and Klipp model disregards a number of important key features of the yeast mating response, particularly the central MAPK cascade that transfers the extra-cellular signal through the cytoplasm to the nucleus (figure 1.8). There has been a great deal of research conducted into MAPK cascades over the past 30 years, both experimentally and mathematically. A number of researchers have investigated why eukaryotic signalling systems are comprised of three sequentially activated kinases with multiple rounds of non-processive phosphorylation [77, 129]. Extensive modelling work by Goldbeter in the early 1970s, Kholodenko, Huang and Ferrell, and Markevich in the 1990s and early 2000's, Xaio Wang, and Fernando Ortega in 2006, and O'Shaughnessey *et al* in 2011 have shown that the signalling cascades produce an ultra-sensitive response to input, converting a graded input signal to a binary output response through the action of nested feedback loops within the cascade and amplification through repeated phosphorylation of the kinases [76, 120, 122–124, 130–132]. The model by Kofahl and Klipp did not replicate any of the behaviours observed by other models of MAP

kinase cascades, and did not build on any of the previous research in this field. While the author has included all of the known components of the pathway, the relationship between them and the dynamic behaviour of the cascade was not replicated. The Kofahl and Klipp model did not produce a sigmoidal increase in the steady-state level of Fus3pp in response increasing initial concentrations of pheromone, characteristic of ultra-sensitivity [133]. The model also did not replicate the chronological order of events observed experimentally (as reported by Yu *et al* during time course simulations, in terms of the activation of the G-protein, followed by sequential phosphorylation of the kinases, and Ste12 [121].

The Kofahl and Klipp model is capable of reproducing the change in the relative amounts of the components in mutants of the mating pathway [128], but cannot be used to study the systems-level behaviour of the signal cascade that underlines the pathway, such as ultra-sensitivity to pheromone, or potentially more complex behaviour, such as bi-stability and oscillation [77, 120, 131].

1.4.5 Modelling tools

Simulation and modelling is becoming a standard approach to understanding biological systems, and this requires software tools that enable researchers to access diverse mathematical modelling and simulation methods [134]. Fortunately there are a range of applications available that enable researchers to access these tools without a mathematics specialism, which will be discussed below [135].

1.4.5.1 Copasi

Copasi is a software application for the simulation and analysis of biological networks. The software is free for non-commercial use and runs on all major operating systems [134]. The Copasi project is an international collaboration between three groups at the Virginia Bioinformatics Institute, the University of Heidelberg, and the University of Manchester. Copasi has a number of unique features, including the criteria to switch between stochastic, deterministic, and hybrid modelling methods; flexible parameter

scans, optimization of arbitrary expressions and parameter estimation using time course and steady-state data [134].

Copasi can be used to build models of systems of biochemical reactions with pre-set rate laws, such as mass-action and Michaelis-Menten, as well as the option to program custom rate laws. Copasi includes a number of simulation functions, including time course simulations, steady state, sensitivity, and metabolic control analysis, and flux balance analysis. Copasi can also perform multiple parameter scans, parameter estimation and fitting of experimental data, and includes a number of optimization algorithms including: genetic algorithms, Levenberg-Marquardt, particle swarm, simulated annealing, and steepest descent. The are also a number of ODE solvers available including the LSODA deterministic method, a hybrid Runga-Kutta method, a Gibson and Bruck stochastic simulator as well as τ-leaping methods. Copasi also includes 2D graphing functions and can output csv format report files containing simulation output data that can be imported into other plotting, modelling, or statistics tools.

Copasi however cannot be used to simulate algebraic equations or differential equations such as partial, difference or delay functions. Copasi also cannot be used to investigate bistable systems or search for multiple steady-states. Steady state analysis can be performed, however Copasi cannot implement features such as Nullclines, or expand around steady states in the graphical interface, such as can be done with XPPAUT. Copasi also lacks the capability to perform statistical analysis on stochastic simulations or output 3D plots of multiple parameters. These functions require the user to process the output data from Copasi with additional tools such as Matlab or AUTO.

1.4.5.2 XPPAUT

XPPAUT is a tool for solving differential, difference, delay, and functional equations as well as solving boundary value problems and stochastic equations [136]. XPPAUT provides an interface to the commonly used bifurcation tool AUTO and can be used to analyse steady states and bifurcation points [136]. XPPAUT is freely available on all major operating systems and uses a graphical user interface enabling easy access

for non-mathematicians. The systems biology format converter project, available from http://sourceforge.net/projects/sbfc/ has developed software to convert SBML files to other common modelling file-formats including XPPAUT. The XPPAUT interface however is difficult to use and the AUTO component is prone to crashing. The Python extensions XPPy enables scripting functions that can be handled by XPPAUT, eliminating the need to use the interface [137], and XPPAUT remains the best tool for bifurcations analysis and has been used in a variety of scientific fields from engineering to biology (Ermentrout B. (2011), Personal communication. Department of Mathematics University of Pittsburgh).

1.4.5.3 Cytoscape

Cytoscape is an open-source bioinformatics tool for visualizing molecular interactions. Cytoscape can import SBML model files and represent them as directed graphs, enabling further bioinformatics study such as network topology, as well as integration with gene expression profiles [138]. Cytoscape provides a useful means of visualizing complex models and contains a number of graph layout features that enable locating functional modules and sub-networks in larger models. Additional features for annotation and interaction with databases of protein-protein, protein-DNA interactions can be augmented through a library of plugins developed by the community of users. Cytoscape can be downloaded from http://www.cytoscape.org.

1.4.5.4 Mathematical Programming Languages

There are a number of commercially-available mathematical programming tools that can be used in engineering, scientific, and mathematical fields to solve sets of complex equations. Two commonly-used tools are MatLab from MathWorks and Mathematica from Wolfram Research. Matlab (meaning “matrix laboratory”) is a commercial, industry standard numerical programming environment, enabling scientists to perform matrix manipulation, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs in other languages [139].

Wolfram Mathematica is a commercial mathematical programming software used in scientific, engineering, and mathematical fields much like Matlab. Matlab focuses on high speed algorithms for numerical computation [140], while Mathematica is designed for symbolic algebra with features such as unlimited precision arithmetic [141].

1.4.5.5 Scripting Languages

Scripting languages are high-level programming languages that are interpreted rather than compiled and use a simpler, more intuitive language than low-level "machine languages" [142] such as C and Assembly language. Fortran, Java, Python, and Perl, are scripting languages suited to processing large data sets and matrices and are consequently used for a wide range of applications by computer scientists [143]. Specific builds of these languages have been developed for use in biosciences, such as BioPerl, and BioPython, and include specialized routines for bioinformatics and modelling. Scripting languages can be used to build programs to automate the manipulation of large data sets and can be adapted to almost any application, however they are slower than machine languages, as they must be run through an interpreter, making them less suited to simulating large-scale computationally intensive models [144]. There are also many interfacing libraries available for Python such as XPPy, Matplotlib, and the Copasi language bindings that enable calling the functions of 3rd party software packages such as XPPAUT, Matlab, and Copasi from within a Python script, combining their features into custom modelling and data handling tools.

1.4.5.6 SBML

The systems biology mark-up language (SBML) is an XML-based computer language designed for representing and exchanging models between different simulation and analysis tools [145]. SBML is an effort to standardize a common file-format, providing compatibility between systems biology researcher's using different modelling tools. SBML can store the details of systems of reactions such as rate laws, parameter values, species numbers and concentrations, and compartment volumes and can be annotated

with the author's notes and publications [146]. Models created in SBML can also be uploaded to public access databases such as biomodels.org. Models can then be imported into the researchers choice of modelling software for simulation or further development, and data produced by a model are reproducible in any lab with software that can read SBML. SBML is an open project and the file-format is free to use from http://www.sbml.org. There are also a number of interfaces available for programming languages such as Python, Perl, Java, Mathematica and Matlab that enable the import and export of SBML model files that can then be simulated or modified using their own native tools [140, 141, 147].

1.4.6 Metabolic Control Analysis

Metabolic control analysis (MCA), or metabolic control theory is a sensitivity analysis of metabolic systems [148]. MCA is a method of analysing how the control of fluxes (J) or metabolite concentration (S) in a metabolic pathway is distributed among the different enzymes that constitute the pathway [149]. MCA can be applied to synthetic biology to determine the control coefficients of the parameters in a network of interactions [150]. The control coefficients generated by metabolic control analysis are related to sensitivity analysis used in engineering, and measure the relative steady state change in a system variable in response to a relative change in a parameter. The two main control coefficients are the flux (equation 1.2) and concentration control coefficients (equation 1.3) [119].

$$C_{vi}^{J} = \left(\frac{\delta J}{\delta p}\frac{p}{J}\right) / \left(\frac{\delta vi}{\delta p}\frac{p}{vi}\right) = \frac{\delta \ln J}{\delta \ln vi} \tag{1.2}$$

$$C_{vi}^{s} = \left(\frac{\delta S}{\delta p}\frac{p}{S}\right) / \left(\frac{\delta vi}{\delta p}\frac{p}{vi}\right) = \frac{\delta \ln S}{\delta \ln vi} \tag{1.3}$$

Where J is the flux through the pathway, S is the concentration of an intermediate in the pathway, i is the reaction step in the pathway, p is the parameter, and vi is the steady state rate of the reaction that is perturbed. Any variable in a system can be analysed with MCA. An important property of the steady state of a network is that for a given flux the

sum of all of the flux control coefficients in a pathway is equal to unity [119, 149–151]. For a small increase n in the rates of all reactions in a pathway, the relative rates of production of the metabolites in the pathway increase by the same amount n, as does the relative rates of consumption. Therefore the metabolite concentrations remain unchanged and the flux of the system increases exactly by n [150]. Mathematically, this means that the flux is a homogenous function of degree one and the metabolite concentration of degree zero. The summation theorem was developed by Gierch and employs the Euler theorem for homogeneous functions for flux control coefficients (equation 1.4) and concentration control efficient (equation 1.5) [152].

$$\sum_i C_{vi}^J = 1 \tag{1.4}$$

$$\sum_i C_{vi}^{[S_j]} = 0 \tag{1.5}$$

The summation theorem can be applied over all the steps in a pathway, and connected pathways therefore MCA and the summation theorem could, in principle be applied over all of the metabolic steps in a cell [119, 152].

1.4.7 Parameter Estimation

Parameter estimation is the process of attempting to calculate model parameter values based on a dataset [153, 154]. A number of mathematical algorithms can be used to estimate a given set of parameter values based on experimental data. For this project, the Hooke and Jeeves, and simulated annealing algorithms were used, as it is good practise to apply multiple parameter estimation algorithms for parameter fitting to compare results (Mendes P. (2009)). Personal communication. Manchester Centre for Integrative Systems Biology).

The method of Hooke and Jeeves is a direct search algorithm that searches for the minimum of a non-linear function without requiring derivatives of the function [155]. At each iteration, this method first defines a pattern of points by moving each parameter one by one, so as to optimize the current loss function. The entire pattern of points is

then shifted or moved to a new location; this new location is determined by extrapolating the line from the old base point in the m dimensional parameter space to the new base point. The step sizes in this process are constantly adjusted to "zero in" on the respective optimum [155]. The Hooke and Jeeve's algorithm is considered to be one of the oldest and simplest of the parameter estimation algorithms, but provides a fast and simple method of fitting experimental data [153].

The method of simulated annealing was developed by Kirkpatrick *et al* using statistical mechanics applied to the way in which perfect crystals are formed. Perfect crystals are formed by first melting a substance, and then allowing it to cool very slowly over a long period of time. At high temperature, the molecules in the crystal vibrate with a wide amplitude, which decreases as the temperature lowers until the molecules settle into the optimum configuration, forming a crystal [153, 156]. The simulated annealing optimization algorithm uses a similar concept: the objective function is considered a measure of the energy of the system and this is maintained constant for a certain number of iterations, called a temperature cycle. During each iteration, the parameters of the model are changed by a small amount and the new objective function is calculated. If the value has decreased then the new state is accepted. If the value increased then the state is accepted with a probability that follows a Boltzmann distribution, therefore, a lower temperature means a higher probability of accepting the new state. After a fixed number of iterations, the stopping criterion is checked; if it is not time to stop, then the system's temperature is reduced and the algorithm continues [156].

Simulated annealing is one of the most robust global optimization algorithms, and although it is also one of the slowest it is guaranteed to converge if run for an infinite number of iterations. [134, 153].

1.4.8 Signal to Noise Ratio

The signal to noise ratio (SNR) is a method of differentiating between the level of a desired signal and the level of a signal from the background noise of the system [157]. A high SNR results from a high signal detection and low signal from the background,

and a low SNR results from a low signal detection and a high background signal. The SNR is calculated as the reciprocal of the coefficient of variation, or the ratio of the mean to the standard deviation of a signal (equation 1.6, where μ is the mean of the data set and σ is the standard deviation of the data set). [158–160].

$$SNR = \frac{\mu}{\sigma} \tag{1.6}$$

Noise in a biological system can originate from extrinsic noise, in which the cellular capacity to produce proteins, and the regulatory mechanisms of the cell fluctuate over time, or from intrinsic noise which is due to stochastic variation in the transcription and translation events in the cell [157].

In relation to this project, the gene circuit is designed to repress the background signal from the reporter gene, increasing the fold change in expression when expression of the reporter gene is induced by the yeast pheromone response. Therefore, the expression of the reporter gene in the gene circuit should have a higher SNR than cells expressing only the reporter gene, as the background signal should be higher from the control than the circuit.

1.5 Synthetic Biology

Synthetic biology is an emergent scientific field, developing from advances in molecular biology and new collaborations between biological and computer sciences [161–164], and utilizes a pragmatic approach of designing interacting components previously studied in isolation by classical fields such as molecular biology [165]. Synthetic biology projects currently attempt to construct genetic circuits in the same way as an engineer combines electrical components, building circuits in living organisms with pre-determined, predictable, and robust behaviours [1, 8, 15, 161, 166]. The term "synthetic biology" first appeared in the Journal Gene and was used by the Polish geneticist Waclaw Szybalski in 1974 [167], describing *in vitro* transcription, and was later used by Hobum *et al* in 1980 describing bacteria that had been genetically engineered using recombinant DNA technology, and the field has more recently been

synonymous with “bioengineering” [162]. The term was used again at the 2000 Annual American Chemical Society in San Francisco to describe work in the field of bio-mimetic chemistry where organic synthesis was used to produce artificial molecules that reproduce biological components generating artificial, life-emulating systems [162, 168]. Therefore, there are two different fields each claiming the term “synthetic biology” - chemists mimicking biology and biologists engineering synthetic phenotypes. The term “synthetic biology”, much like “systems biology” is not new, and the work done by synthetic biologists has originated in classical fields of molecular biology and genetic engineering, leading to scepticism of the term used in research and in what it can deliver over existing established fields [169, 170]. Synthetic biology differs from molecular biology, in that it focuses on building novel functions and behaviours from molecular interactions, and apply them in biotechnology [169]. Also, synthetic biology enables the “reverse engineering” of biology by attempting to construct biological systems from the bottom-up using component parts of gene and protein interactions [161, 171]. Biological processes such as oscillations and switches that pervade biological processes such as the cell cycle are not fully understood [172], and by engineering such systems through synthetic biology we gain a greater understanding of how they function in native systems [171].

Synthetic biology integrates computer science by building models to predict and optimize the behaviour of these interactions prior to construction in the laboratory [10, 173]. The key to building biological systems is not in how the individual components function (as studied in classical fields of molecular biology), but on how they interact [169].

Currently, synthetic biologists share the holistic philosophy of systems biology in that rather than studying individual genes or pathways, they assemble systems of genes and gene products into interacting biological devices that can be combined into modular components, conferring new functions on the systems into which they are embedded [8, 55, 170, 174, 175]. However, the synthetic biologist also attempts to rationalise the emergent behaviour into a set of rules for the modular construction of these behaviours [8, 11, 176].

In using this approach, synthetic biologists digress from classical molecular biology towards the field of engineering through the construction of interchangeable parts lists of "biological circuitry" [8, 170, 177]. The engineering perspective presents biology as a tool set of parts that can be used to achieve a specific task [11, 168], whilst testing modularization concepts in biology, and exploring the challenge of artificial reconstruction [10, 168]. The evolution of synthetic biology through combination with systems biology and engineering is facilitating the application of *in silico* design and testing of cellular circuitry prior to fabrication, allowing for "design-based engineering" of biological systems [168, 170, 173].

Synthetic biology modules cover a wide range of applications and tend to originate from naturally occurring systems, modified by the synthetic biologist towards a desired function [8, 168]. However, it can be difficult to transplant wild-type genetic circuits that have evolved and been optimized over millions of years in their host environment into an artificial context, this requires rational redesign based on modelling and directed evolution to help them interface together and function alongside their host [8]. Thus-far in the development of synthetic biology, synthetic transcriptional regulation networks are the most widely implemented and characterized modules, and have been used to build cascades, feed-forward, and feedback loops, forming switching and oscillating responses as well as rudimentary information processing tasks [46, 118, 178–184].

1.5.1 Transcription Cascades

Regulatory cascades are sequences of genes that activate each other in a step-wise manner through the cascade, passing a signalling input from the top of the cascade through to an output response at the bottom [8, 46, 168, 185]. Regulatory cascades are ubiquitous in biological systems and can be found in signal transduction and protein kinase pathways, such as the MAP kinase cascade in budding yeast and flagellar motion in *E. coli* [8, 185]. Regulatory cascades provide an "all or nothing" response to graded input signals, where small changes in input concentration switch the output between high and low levels. Hooshangi *et al* demonstrated that the longer the transcriptional cascade, the higher the sensitivity and the faster it switches between steady-states, and

also the higher the amplification of noise at each step, disrupting synchronisation within cell populations [8, 185].

1.5.2 Synthetic Oscillators

Oscillations of protein levels are vital to coordinating cellular events such as the cell cycle or circadian rhythms [168]. Elowitz and Leibler designed a cyclic negative-feedback loop using three transcriptional repressor systems in *Escherichia coli*, forming an oscillating network which they termed the "represillator" (see figure 1.10) [179].

The repressilator was constructed using a low-copy number plasmid containing the *lacI* gene from *E. coli* which inhibits transcription of the second repressor gene, *tetR* from the tetracycline-resistance transposon Tn*10*, whose protein product in turn inhibits the expression of a third gene, *cI* from λ phage, and the *cI* gene product inhibits *lacI* expression, completing the cycle [179] (see figure 1.10). The repressilator activity was observed using a high-copy number reporter plasmid containing a tet-repressible promoter fused to an intermediate-stability GFP gene producing oscillating fluorescence [179]. The represillator was simulated using a set of coupled differential equations

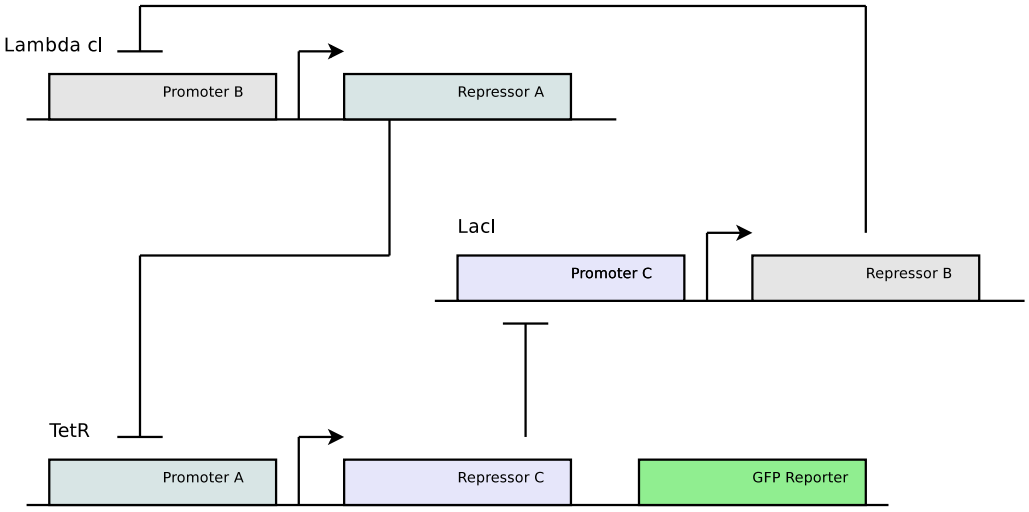

FIGURE 1.10: Diagrammatic representation of an oscillating gene circuit. Adapted from Elowitz and Liebler [179].

with repressor-protein concentrations p_i and the corresponding mRNA concentrations, m_i (where i is *lacI*, *tetR*, or *cI*) were treated as continuous dynamic variables. Each

of the six molecular species participated in transcription, translation, and degradation reactions. All three repressors were treated identically, except for their DNA-binding specificities. The kinetics of the system were determined by two coupled first-order differential equations (equations 1.7 and 1.8) where the number of protein copies per cell produced from a given promoter type during continuous growth is α_0 in the presence of saturating amounts of repressor (owing to the "leakiness" of the promoter), and $\alpha + \alpha_0$ in its absence; β denotes the ratio of the protein decay rate to the mRNA decay rate; and n is a Hill coefficient [179].

$$\frac{dm_i}{dt} = -m_i \frac{\alpha}{(1+P_i^n)} + \alpha_0 \tag{1.7}$$

$$\frac{dp_i}{dt} = -\beta(p_i - m_i) \tag{1.8}$$

The represillator demonstrated it was possible to design and construct an artificial genetic network with new functional properties from generic components that naturally occur in other contexts [179]. However, the represillator was noisy and subject to variation in the amplitude of oscillation which the researchers attributed to possible stochastic effects inherent in natural gene-expression systems, particularly when there are a small number of reactants [186]. The variation was replicated in the Elowitz model when these stochastic interactions were simulated [179].

Barkai *et al*, using the Monte Carlo algorithm also simulated stochastic reaction events in an oscillating gene network and demonstrated variation in the amplitude of oscillations when changing the rates of transcription and translation [186]. It has been suggested that the presence of both positive *and* negative feedback control elements in natural oscillatory systems enables bistability and hysteresis, and therefore noise resistance [178, 179, 186].

Later, in 2002 Hasty *et al* published a theoretical oscillator comprising positive and negative feedback controls as a set of ordinary differential equations. Stricker *et al* utilized the theoretical synthetic oscillator circuit proposed by Hasty *et al*, and constructed the circuitry in *E. coli* [180]. Stricker built on previous published work and assembled the circuitry using linked positive and negative feedback loops (figure

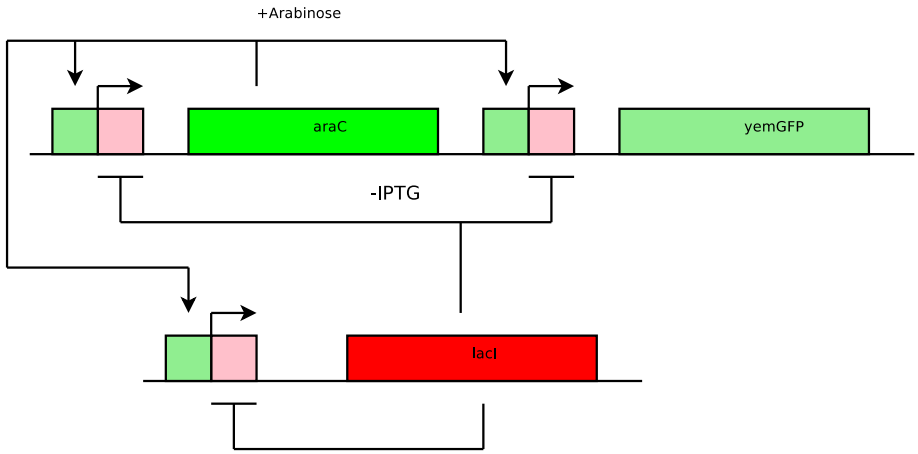

FIGURE 1.11: Network diagram of the dual-feedback oscillator. A hybrid promoter $P_{lac/era-1}$ drives transcription of *araC* and *lacI*, forming positive and negative feedback loops. Adapted from Stricker *et al* [178].

1.11) [178]. Stricker's oscillator circuit was comprised of a hybrid $P_{lac/ara-1}$ promoter with the activation operator site from the *araBAD* promoter placed upstream of the transcription start site, and repression operator sites from the *lacZYA* promoter placed both upstream and immediately downstream of the transcription start site [178] (figure 1.11). The circuit is activated by the AraC protein in the presence of arabinose and repressed by the LacI protein in the absence of isopropyl β-D-1-thio-galactopyranoside (IPTG) [178]. Stricker placed *araC*, *lacI* and *yemGFP* (yeast-enhanced green fluorescent protein) genes under the control of three identical copies of $P_{lac/ara-1}$ to form three co-regulated transcription modules. Activation of the promoters by the addition of arabinose and IPTG to the growth medium results in transcription of each component of the circuit, and increased production of AraC in the presence of arabinose results in a positive feedback loop that increases promoter activity [178]. The concurrent increase in production of *lacI* results in a linked negative feedback loop that decreases promoter activity, and the differential activity of the two feedback loops drives the oscillatory behaviour [178].

The author was able to modulate the oscillator response by varying the level of arabinose in the medium and "tune" the oscillatory period. At a fixed value of 2mM IPTG and at 37°C, the oscillatory period can be tuned from 13min to 58min by varying the arabinose concentration [178]. Stricker *et al* however found that the original model by Hasty *et.al* predicted a small parameter space for the inducer that would facilitate oscillation, and

did not predict the experimental observation [178]. The model required fine tuning of the parameter values to achieve an oscillatory response whereas the circuits *in vivo* behaviour was robust through a range of parameter values [178]. Stricker *et al* further developed the model to include both the positive and negative feedback loops and it was found that directly modelling protein-DNA binding, multimerization, translation, DNA looping, enzymatic degradation and protein folding greatly increased the accuracy of the model [178]. The result was a computational model that was very robust to parameter variations and correctly describes the dynamics of the oscillator for a large range of IPTG and arabinose concentrations [178].

This observation was also investigated by Tsai *et al* who researched the significance of feed-forward and backward loops in regulatory circuits using a modelling approach [187]. Tsai *et al* analysed a large number of established oscillatory models with various loop back systems including negative, negative plus negative, and negative plus forward feedback loops (figure 1.12). Tsai *et al* constructed a set of three hypothetical oscillator circuits, consisting of a three-variable triple negative feedback loop, one with no additional feedback, another with added positive feedback, and a third with an added negative feedback loop (figure 1.12) [187]. A random parameter set for each of the models was generated, which were then observed for limit cycles. The parameter sets were generated until 500 oscillating sets had been found for each model. The study found that the inclusion of a feed-forward with a feedback loop produces a wider parameter space to vary the frequency of oscillation, whilst maintaining a constant amplitude, compared with a feedback loop alone [187]. Feedback loops that include an additional feed-forward loop are more robust than feedback alone, maintaining constant oscillatory period and amplitude over a wider parameter range, and provided insight into the reasoning behind the natural design of cell signal regulation [187].

The work published by Stricker *et al* demonstrated the advantage of a synthetic biology approach, adopting computer modelling with more classical fields such a molecular biology. Using wet lab molecular biology alone resulted in oscillatory circuits with high variability and noise, and could not provide insight into developing a persistent and robust oscillator [179]. Pure theoretical studies provided a hypothesis for the design of an oscillatory circuit, however the response predicted by the model was not replicated

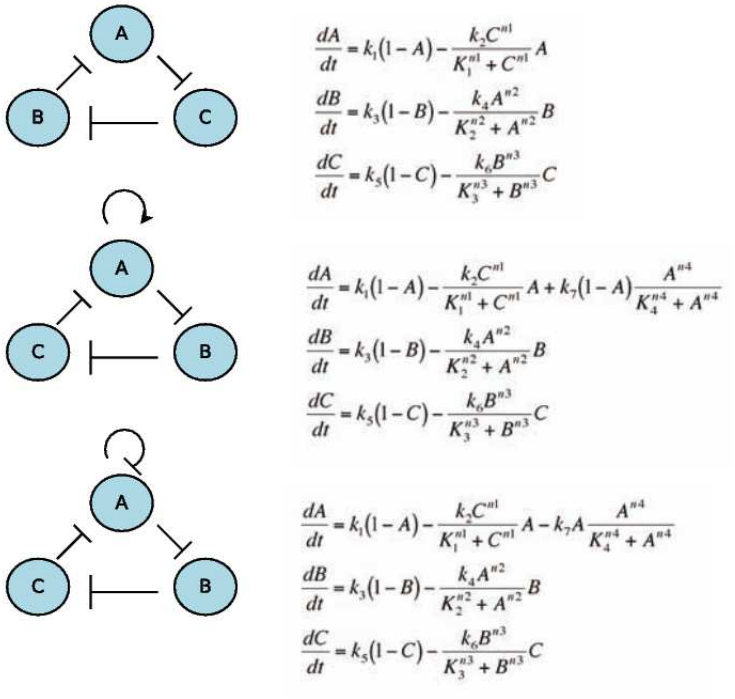

FIGURE 1.12: Tsai *et al* negative-feedback models parameterised with random values to find oscillatory behaviour. A. Negative feedback only. B. Positive-plus-negative feedback. C. Negative-plus-negative feedback. Adapted from Tsai *et al* [187].

in the laboratory, and modelling alone was not able to anticipate all the parameters affecting the *in vivo* response [180]. The approach of combining modelling with laboratory experimentation by Stricker *et al* and later Tsai *et al* enabled the refinement of both the modelling and the *in vitro* circuit, which would not have been possible from the individual approaches [178, 187].

1.5.3 Synthetic Switches

Gardner *et al* designed a genetic "toggle switch", and created a gene circuit enabling a bistable state and "memory" in a biological system [188]. The toggle switch was constructed in *E. coli* and consisted of a bistable gene-regulatory network, using integrated theoretical modelling and laboratory experimentation[188]. The switch was comprised of two repressors and two constitutive promoters, where each promoter is inhibited by the repressor that is transcribed by the opposite promoter (see figure 1.13). Gardner cites this design as the most efficient configuration for a switch, as it requires the fewest genes and *cis*-regulatory elements to achieve a robust behaviour, in terms

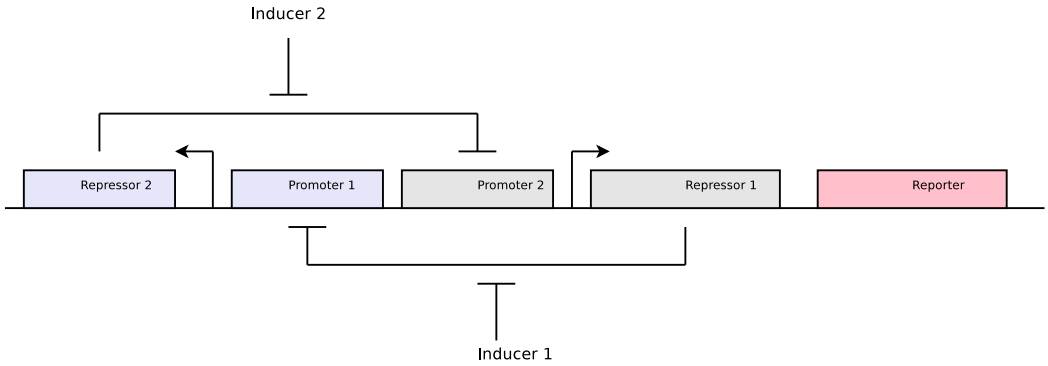

FIGURE 1.13: Gardner *et al* toggle switch design. Repressor 1 inhibits transcription from promoter 1 and is induced by inducer 1. Repressor 2 inhibits transcription from promoter 2 and is induced by inducer 2. Adapted from Gardner *et al* [188].

of tolerance to fluctuations in gene expression [188]. The toggle switch was modelled using a set of coupled differential equations (see equations 1.9 and 1.10), where U is the concentration of repressor 1, V is the concentration of repressor 2, α_1 is the effective rate of synthesis of repressor 1, α_2 is the effective rate of synthesis of repressor 2, β is the cooperativity of repression of promoter 2 and γ is the cooperativity of repression of promoter 1 [188].

$$\frac{dU}{dt} = \frac{\alpha_1}{1+V^{\beta}} - U \tag{1.9}$$

$$\frac{dV}{dt} = \frac{\alpha_2}{1+U^{\gamma}} - V \tag{1.10}$$

The bi-stability arises from the mutually-inhibitory arrangement of the repressors. In the absence of inducers the switch either transcribes repressor 2 from promoter 1 or transcribes repressor 1 from promoter 2. Introducing an inducer of the current active repressor activates the switch through maximal transcription of the opposite repressor, until it stably represses the original active promoter [188]. The Gardner toggle switch was the first form of "synthetic biology" that was significantly different from genetic engineering because it utilized network architecture for the switching mechanism, rather than proteins and other regulatory elements to achieve the required behaviour [188]. The toggle switch provided a simple self-contained programmable circuit, requiring transient rather than sustained induction, and retained its new stable state after induction. Later, in 2001 Becskei published an alternative method of producing a genetic switch circuit using positive feedback through a tetracycline-dependant activator that up-regulated its own expression (see figure 1.14) [189]. Becskei *et al* constructed

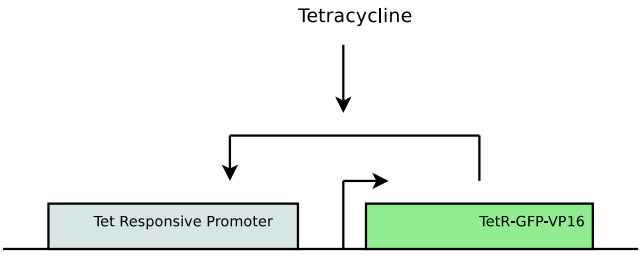

FIGURE 1.14: Diagrammatic representation of the Becskei *et al* positive feedback genetic switch circuit. Adapted from Becskei *et al* [189]

.

the circuit from well-defined promoters and transcriptional activators, and the circuit operated autonomously of the existing genetic machinery of the cell. The circuit is comprised of a plasmid containing the tetracycline-responsive transactivator (rtTA): rtTA produces a graded response in constitutive systems, making possible the analysis of positive feedback by the model [189]. The circuit can convert a graded response to the inducer (tetracycline) or gene copy number into a binary response from the reporter (GFP) [189]. The results of the *in vitro* work demonstrated that under conditions where rtTA is expressed with positive feedback the cell population is divided into populations of "on-cells" and "off-cells" (GFP expression or no expression), and the number of cells in each state is proportional to inducer concentration [189]. Off-cells were also capable of switching to the on-state in a stochastic manner, and the model suggested that on-cells could also switch to the off-state, however this was not observed experimentally [189]. The population of off-cells became a population of mixed on/off cells over time, independent of the level of basal expression indicating that a high basal expression rate was not required for activation of the circuit after induction [189]. Becskei suggests the switch could be defined as a "noise-based" switch rather than a toggle switch (as designed by Hasty *et al*, 2000) as the population continues to switch from the off to the on-state over a range of inducer concentrations rather than the entire population switching at once, as seen with the Gardner *et al* switch [188].

Ajo-Franklin *et al* built on the work by Gardner and Becskei, by designing and constructing a memory circuit in yeast [188–190]. The circuit was constructed from an activator gene and a reporter gene (figure 1.15). The activator gene consists of a DNA binding domain (DBD), two copies of the red fluorescence protein gene (RFP) (mCherry), the viral activation domain VP64, and the SV40 nuclear localization

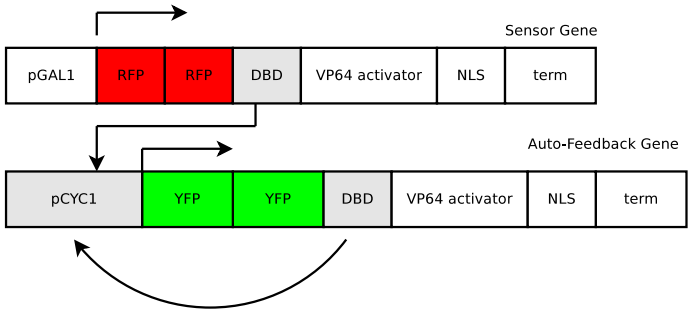

FIGURE 1.15: Ajo-Franklin memory circuit diagram. The circuit consists of a sensor gene and an auto-feedback gene. Upon addition of galactose the growth medium, the sensor gene promotes the expression of RFP and up-regulates the expression of the auto-feedback gene. The auto-feedback gene expresses YFP and up-regulates its own expression, maintaining expression of YFP after removal of galactose from the growth medium. Adapted from Ajo-Franklin [190].

sequence (NLS), under the control of the galactose-inducible promoter (p*GAL1*). The reporter gene contained the *CYC1* promoter and two copies of the yellow fluorescence protein (YFP). The DBD's used were LexA, an engineered version of the murine zinc-finger Zif268 (ZifH), and the human zinc finger Gli1. The components were arranged into a positive auto-feedback loop that switches between two steady states following activation (figure 1.15). During vegetative growth the cells express neither reporter gene. Upon stimulation with galactose the cells express both RFP and YFP as the sensor gene expresses the activator for *pCYC1*. The *CYC1* promoter up-regulates expression of its own activator, maintaining expression of YFP. Therefore, upon returning to raffinose the cells continue to express YFP in the absence of galactose. The circuit therefore moves to a second steady state and demonstrates a memory characteristic [190]. The project incorporated construction of the memory circuit as well as modelling that enabled prediction of relationship between the activator concentration and the production rate of either the reporter or auto-feedback activator. Using Michaelis-Menten kinetics, the experimental data for the expression of the fluorescent reporter genes was used to fit a hill coefficient in the model rate equations to build an *in vivo* relevant model of the system that could predict the expression level of the reporter genes and the switch point between the steady states of the system [190].

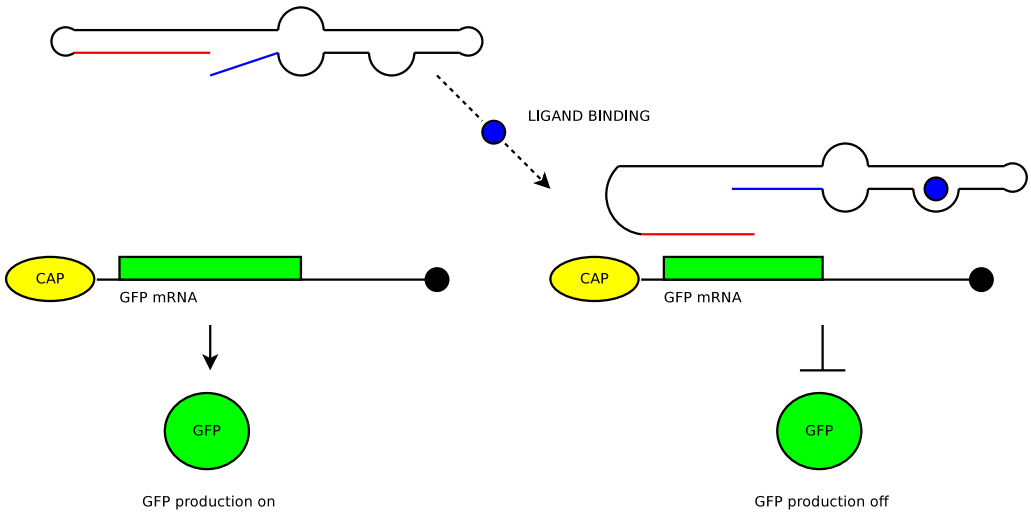

FIGURE 1.16: Diagrammatic representation of the mechanism of translational regulation by the Bayer and Smolke riboswitch. Adapted from Bayer and Smolke [182].

1.5.4 Riboswitches

A third type of genetic switch circuit was published by Bayer and Smolke, utilizing an anti-switch molecule that regulates gene expression *in vivo* [182]. The authors designed a set of non-coding small *trans*-acting RNA riboregulators, termed "antiswitches" that regulate gene expression in a ligand dependant manner. The riboregulators consist of two distinct domains. One end of the molecule contains an antisense domain that is specific to the gene that is to have its expression controlled, and the other end has an aptamer domain that recognizes a specific effector ligand [182]. When a ligand binds to the aptamer domain it induces a conformational change in the riboregulator that exposes the antisense domain that can bind with the mRNA transcript of the target gene and block translation. In the absence of ligand, the antisense domain is sequestered in an "antisense" stem and is not available for mRNA binding, and the antiswitch is in the "off" state. [182]. In this state the antiswitch is unable to bind to the target transcript, which has a green fluorescence protein coding region, and as a result, GFP production is on. In the presence of effector, the antiswitch binds to the switch molecule, forcing the aptamer stem to form, switching its confirmation to the "on" state. In this state the antisense domain of the antiswitch will bind it the target transcript and through and antisense mechanism turn the production of GFP off (figure 1.16) [182]. The work demonstrated that engineered ligand controlled, antisense RNAs can be used as

allosteric regulators of gene expression [182]. In the absence of the ligand the free energy of the anti-sense stem is lower than that of the aptamer stem and binding of the ligand stabilizes the formation of the aptamer stem facilitating binding to the antisense domain of the target mRNA transcript [55]. The anti-switch can be used to enable both positive and negative regulation, with "on" switches designed such that the absence of ligand destabilizes the aptamer stem binds to the anti-sense domain and disables the molecule from binding to the transcript [55]. The riboregulators can also be tuned by engineering the free energy of the antisense domain, altering the conformation of the riboregulator with a predictable functional effect [55]. The study demonstrated that this approach can be used to decrease the stability of the antisense stem, decreasing the ligand concentration necessary to switch the riboregulator conformation and conversely, increasing the stability of the antisense stem increases the amount of ligand required and shifts the system's dynamics to favour the "off" state at low ligand concentration [55].

The riboregulator "antiswitch" mechanism is also modular in that ligand response and transcript targeting can be engineered by swapping domains within the aptamer and anti-sense molecule and both domains operate independently. The customizable generic nature of riboswitches provides a potential wide range of applications in both prokaryotes and eukaryotes. Specific transcripts could be targeted in gene therapy or cell specific targeting to complement existing therapies [55]. The technology could also be applied in synthetic biology in the design of regulatory pathways and control loops for synthetic circuit design by enabling the cell to sense and respond to intracellular metabolite levels and environmental signals, potentially providing "smart regulators" capable of targeting any gene with any ligand [182].

1.5.5 Application of Synthetic Biology

The growing range of tools being developed in synthetic biology are enabling researchers to construct increasingly complex synthetic behaviours through the modular combination of genetic components [12, 190]. These synthetic circuits can be implemented in practical applications to solve biotechnology problems. For example, bacteria and yeast are already in widespread use in the biotechnology and fermentation

industries. The application of multi-cellular fermentation systems comprising separate "sender" and "receiver" cells presents the opportunity to eliminate the requirement for expensive inducers, as cultures will maintain their own gene expression levels, eliminating the need for monitoring batch cultures [191]. The production of expensive small molecules can also be replaced by synthetic biology processes, as demonstrated by Du Pont and Tate & Lyle to produce chemicals commonly used in textiles from corn sugar [192]. Artemisinin, a naturally occurring anti-malarial drug produced through an expensive and low efficiency plant extraction process can now be obtained from yeast, which produce a precursor to the active drug with a high yield [193].

Coupling gene regulatory networks to external stimuli can be used to produce new biosensor cells for the detection of a variety of compounds with industrial and medical applications [1]. Programmed behavioural interactions between prokaryotes and eukaryotes could provide new disease treatments, as demonstrated by Anderson *et al* who engineered *E. coli* cells to invade specific mammalian cells exhibiting a tumourogenic phenotype under specific inducer conditions, providing potential new cancer treatments [194].

This project builds on the concept of engineering modules of genetic interactions that confer synthetic phenotypes on the host cell. The project utilized the yeast mating response pathway as a generic input and signal processing module, exploiting the signal amplification and noise attenuating features of the MAPK cascade [131, 195, 196]. The circuit was constructed from three independent plasmids, which by themselves do not produce an effect on the phenotype of the host cell; however, when in combination they form a discrete network or system, and it is the emergent property of this system that produces the synthetic behaviour in the host [197, 198]. This systems-level interaction also presents a complexity in understanding and designing the circuit, which requires new modelling and simulation tools from systems biology to understand and predict its behaviour [112, 113].

Founding research in synthetic biology has so far focused around building interacting genetic "parts lists" in *E. coli*; through efforts such as the Biobricks Foundation and iGem there are registries of parts that can be combined together to build synthetic

behaviours [11, 199, 200]. This project is one of a limited number of studies that have attempted to construct, and characterize such synthetic parts in more complex eukaryotic yeast cells [1, 2, 190], allowing for the creation of more complex circuits and provide a stepping stone towards building more complex synthetic systems and even organisms [8, 12, 161, 201, 202].

1.5.6 Project Overview

Building on the concept of combining defined genetic components into circuits, the project aims to utilize three components previously characterised by published research. The iron response protein (IRP) by Koloteva *et al*, the LexA repressor by Brent and Ptashne, and the luciferase reporter gene, assayed with a published *in vivo* whole cell assay [4, 203, 204]. The combination of these components enables repressors to be combined that function separately at the transcription and translation levels [4, 203]. This enables the hierarchical and time-scale separation of the repressors, and also the further characterisation of the two components in a synthetic biology application, using a well characterized and sensitive reporter assay [204]. The IRP repressor is well-defined in mammalian systems [205–207], and has been shown to function in yeast where it can repress translation of genes containing the target IRE stem-loop structure [4]. The circuit is designed to express the luciferase reporter gene, which has been well characterized and used extensively in both prokaryotic and eukaryotic cells [208–210]. Expression of the luciferase gene is linked with the yeast pheromone response pathway via P_{FUS1}, the promoter for the *FUS1* gene which is known to be up-regulated by the Ste12 transcription factor and involved in initiation filamentous growth and formation of the Schmoo tip [211]. This system forms the sensor component of the circuit, using the pheromone response pathway as a signal processing molecule and expression of the luciferase reporter gene as the output.

The two repressors, IRP and LexA, are designed to tune the circuit's luciferase output. The IRP is designed to be expressed constitutively using known constitutive yeast promoters that can be interchanged to express different levels of IRP concentration within the cell [212]. The luciferase reporter gene includes the IRE sequence, encoding

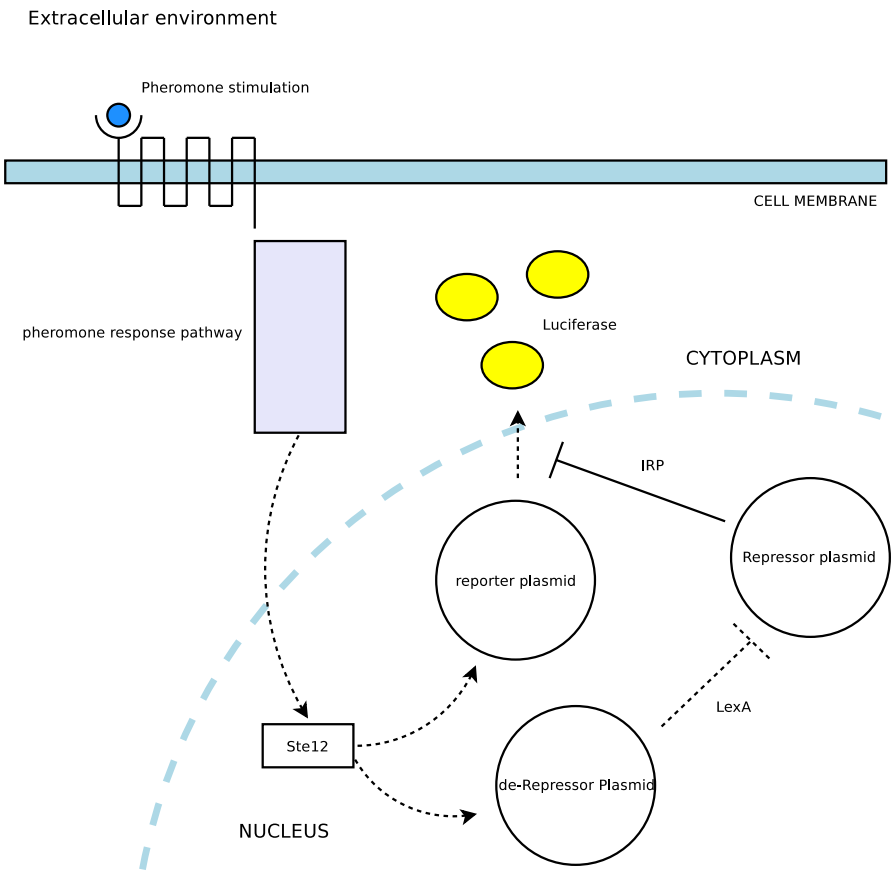

FIGURE 1.17: Schematic diagram of the gene circuit demonstrating the interaction of the components. The yeast pheromone response pathway functions as a signal processing module translating pheromone binding at the cell receptor to Ste12 transcription factor activation. Ste12 up-regulates expression of the reporter and de-repressor plasmids, resulting in repression of IRP transcription and de-repression of luciferase translation. Solid line denotes constitutive repression by the IRP. Dotted lines denote pheromone response pathway induced reactions.

a stem-loop structure in the luciferase mRNA molecule to which the IRP can bind and block translation by the ribosome [4]. Constitutive repression of luciferase mRNA translation is designed to reduce basal expression of the reporter gene during vegetative growth and reduce noise in the circuit output. The second repressor gene; LexA is expressed from the same P_{FUS1} pheromone inducible promoter as luciferase, and LexA operator sequences (DNA binding domains for the LexA protein) were placed upstream of the IRP constitutive promoter. This design enables pheromone inducible repression of IRP transcription, and simultaneous up-regulation of luciferase expression during pheromone induction (figure 1.17).

The hypothesis for the circuit is that through repression of basal gene expression, the

circuit achieves a higher fold change increase in luciferase expression compared with cells that would contain a pheromone-induced luciferase reporter gene alone. The circuit is unique in the application of both transcriptional and translational repression, and enables the characterisation of these two systems in a synthetic biology circuit. The project aims to characterize the expression of each of the components of the circuit through quantification of mRNA and protein expression levels, as well as investigating the interaction of the components as a circuit using luminescence.

There are a number of interactions within the circuit, at a number of hierarchical levels within the cell, which leads to complex non-linear dynamics in the behaviour of the circuit. The project therefore uses mathematical modelling to build a predictive model of the circuit that can be further refined through parameterisation with experimental data. The combination of experimental and theoretical fields is a novel approach that differentiates synthetic biology from molecular biology. Modelling enables faster design and development cycles of gene circuits, through the application of systems analysis to identify key control points in the pathways under construction. Metabolic control and sensitivity analysis will be employed at the start of the project to understand the role of each interaction in the circuit, and then further parameterisation with experimental data will enable predictive time course simulations to be constructed. The model will then become a tool for tuning the behaviour of the circuit and adding additional features, driving hypothesis generation for the experimental work. The interplay of each of the component fields therefore, forms a novel approach compared to classical genetic engineering of yeast.

CHAPTER

TWO

MATERIALS AND METHODS

2.1 Plasmids

The plasmids used in this study are listed in table 2.1.

2.2 Primers

The primers used in this study are listed in table 2.2.

2.3 Yeast & Bacterial Strains

The yeast strains used in this study are listed in table 2.3. The bacterial strain used in this study was *E. coli* TOP10. The strain details are as follows: *E. coli* TOP 10F' [F'*lac*^{q}Tn10 (*Tet*^{r})*mcrA* Δ *(mrr-hsdRMS-mcrBC)* F80 *lacZ*Δ M15 *Dlac74 deo*R *rec*A1*ara*D139 Δ (*ara-leu*) 7697 *galU galK rpsL endA1 nupG*.

Plasmid	*Notes*	*Source*
pRS315	single copy yeast ARS-CEN plasmid (*LEU2*)	Sikorski and Hieter. [213]
pRS313	single copy yeast ARS-CEN plasmid (*HIS3*)	Sikorski and Hieter. [213]
pRS315-P_{FUS1}-IRE-luciferase	single copy yeast ARS-CEN plasmid (*LEU2*). pFUS1 cloned BamHI and NdeI, luciferase cloned with NdeI and HindIII, and IRE cloned with NdeI.	Based on pRS315.
pRS313-P_{FUS1}-LexA	single copy yeast ARS-CEN plasmid (*HIS3*) with pFUS1 cloned with BamHI and NdeI and LexA cloned with NdeI and HindIII.	Based on pRS313.
pRS315-*pFUS1Pax3*-luciferase	single copy yeast ARS-CEN plasmid (*LEU2*) with pFUS1-Pax3 cloned with BamHI and NdeI.	Based on pRS315-pFUS1-IRE-luciferase.
YCp33-Supex2	single copy yeast ARS-CEN plasmid (*URA3*).	Oliveira *et al* [214].
LexAop-P_{DCD1}-IRP	single copy yeast ARS-CEN plasmid (*URA3*). LexA operator cloned with HindIII, IRP cloned with SalI and XhoI.	Based on pDCDex, Oliveira *et al* [214].
LexAop-P_{TEF1}-IRP	single copy yeast ARS-CEN plasmid (*URA3*). LexA operator cloned with HindIII, IRP cloned with SalI and XhoI.	Based on pTEFex, Oliveira *et al* [214].
pJM4	single copy yeast ARS-CEN plasmid based on YCp22-FL with luciferase (*URA3*),	Oliveira *et al* [215].
pJM6	single copy yeast ARS-CEN plasmid based on YCp33-Supex2 with IRP (*URA3*))	Oliveira *et al* [215].
pTRPEX	single copy yeast ARS-CEN plasmid based on YCp33-Supex2 with *TRP1* promoter (*URA3*).	Oliveira *et al* [215].
pDCDex	single copy yeast ARS-CEN plasmid based on YCp33-Supex2 with *DCD1* promoter (*URA3*).	Oliveira *et al* [215].
pTEFex	single copy yeast ARS-CEN plasmid based on YCp33-Supex2 with *TEF1* promoter (*URA3*)	Oliveira *et al* [215].
pSVa17	Cln2 PEST tagged *yEGFP* vector plasmid	Gift from Simon Avery [216].
TOPO TA Cloning Vector		Invitrogen TOPO cloning kit.

TABLE 2.1: List of plasmids used in this study.

Name	*Nucleotide Sequence*
PCR primers and Oligonucleotides	
LexAOpCasLeft	AGCTTCGAGTACTGTATGTACATACAGTACTCGAGTACTGTATGTACATACAGTACTTAATTAA
LexAOpCasRight	AGCTTTAATTAAGTACTGTATGTACATACAGTACTCGAGTACTGTATGTACATACAGTACTCGA
Luciferase Fwd	CTAGCTTAGTCGACGAAGACGCCAAAAACATAAAG
Luciferase Rev	CTAGCTTACTCGAGTTACACAATTTGGACTTTCCG
LexA Fwd	ATCGCTAGCATATGATGAAAGCGTTAACGGCCAGG
LexA Rev	ATCGCTAGAAGCTTTTACAGCCAGTCGCCGTTGCG
IRE Left	TACCAATTATCTACTTAAGCTTCAACAGTGCTTGAACTTAAGAACACAAAACTCGAGAAGA
IRE Right	TATCTTCTCGAGTTTTGTGTTCTTAAGTTCAAGCACTGTTGAAGCTTAAGTAGATAATTGG
PEST Fwd	ACAATCGATGGCCATCGCGAAAGCATCCAACTTGAACATTTCG
PEST Rev	TAATTAGTTGGCCA TCGCGA CTATATTACTTGGGTATTGCC
IRP SalI	CGTGTAACGTCGACATGAGCAACCCATTCGCA
IRP XhoI	CGCGTCACCTCGAGGCTTGGTTCTCTCTTTCTGGC
Sequencing Primers	
DCD1-SEQ-FWD	GCGGTACGCAGTTATGAG
DCD1-SEQ-REV	ATTCACACCTTTAATGTGCCAA
IRP1-SEQ-FWD	CAACCCATTCGCACACCTTG
IRP2-SEQ-REV	GAATGCCCAAGCCATCAATC
IRP3-SEQ-FWD	GGGAGATTCGGTAACAACTG
IRP4-SEQ-REV	GCGGATCATGTAGTTGAG
IRP5-SEQ-REV	CGCTGAAGGGTAACATAG
IRP6-SEQ-FWD	ATGACGCCAGATGGCAGTAG
PGG-Terminator	CGCTGAAGGGTAACATAG
IRP-SEQ-REV	GTAAGCGTGTGGAACGACT
TEF1-SEQ-REV	CTTTCCTAGGCAGCTGAGCT
M13-Fwd (-20)	GTAAAACGACGGCCAGT
M13-Rev (-27)	CAGGAAACAGCTATGAC
qPCR Primers	
qALG9-Fwd	CACGGATAGTGGCTTTGGTGAACAATTAC
qALG9-Rev	TATGATTATCTGGCAGCAGGAAAGAACTTGGG
qHEM2-Fwd	TTCCGCTATTCATCTCCGATAATCCAG
qHEM2-Rev	ACAGACATCGCAAATAATATACAGTTCAGG
qALG9-Fwd	CACGGATAGTGGCTTTGGTGAACAATTAC
qALG9-Rev	TATGATTATCTGGCAGCAGGAAAGAACTTGGG
qIRP-Fwd	AACCCATTCGCACACCTTG
qIRP-Rev	ATGGTAAGCGCCCATATCTTG
qLexA-Fwd	CAGGAAGAGGAAGAAGGGTTG
qLexA-Rev	TCGGCTTGAATAAGGAAGGA
qLucif-Fwd	TACTGGGACGAAGACGAACA
qLucif-Rev	TTCCGTGCTCCAAAACAAC

TABLE 2.2: List of primers used in this study.

Strain	*Genotype*	*Source*
pTC5 (BY5741)	*Mat**a**;his3Δ leu2Δ0 met15Δ0 ura3Δ0*	Euroscarf, Brachman *et al*[217]
*sst2*Δ (BY4741)	*Mat**a**; his3 Δ1; leu2Δ0; met15Δ0; ura3Δ0; YLR452c::kanMX4*	Euroscarf

TABLE 2.3: List of yeast strains used in this study.

2.4 Yeast Growth Conditions

Yeast cells were grown in liquid broth media consisting of yeast peptone dextrose (YPD) (ForMedium^{tm}) (1% yeast extract, 2% peptone, 2% glucose) or on agar plates containing 2% agar. Yeast transformed with plasmids containing auxotrophic markers *LEU2*, *URA3*, or *HIS3* were grown in liquid yeast nitrogen base (YNB) without amino acids, and supplemented with the appropriate (ForMedium^{tm}) drop out medium: -*LEU*, -*URA3*, or -*HIS3* respectively with 2% glucose, or on plates containing 2% agar. Cells containing combinations of plasmids were grown in the appropriate combinations of drop out media for the auxotrophic marker combinations. Yeast cells containing the circuit were grown in in liquid yeast nitrogen base (YNB) without amino acids, supplemented with (ForMedium^{tm}) drop out: -*LEU*, -*URA3*, -*HIS3*. Luciferase control strains were grown in liquid yeast nitrogen base (YNB) without amino acids (ForMedium^{tm}) drop out: -*LEU*. No de-repressor plasmid controls containing the repressor plasmid and reporter plasmid were grown in liquid yeast nitrogen base (YNB) without amino acids (ForMedium^{tm}) drop out: -*LEU*, -*URA3*. All media contained 2% glucose as carbon source. Cultures were also grown in the equivalent agar plates containing 2% agar.

Broth cultures were grown in 25ml volumes of liquid broth in 100ml baffled shake flasks, incubated at 30°C with shaking at 250rpm. Plate cultures were incubated at 30°C. Stocks strains were stored in the appropriate broth media containing 25% glycerol, and stored at -70°C.

2.5 Bacterial Growth Conditions

E. coli TOP10 cells were cultured in 2ml LB (Luria-Bertani) broth (Formedium) (1% bacto-typtone, 0.5% bacto-yeast extract and 0.5% NaCl) or LB agar plates containing 1% agar. *E. coli* TOP10 in liquid broth were incubated at 37°C with shaking at 250rpm. Plates were incubated at 37°C overnight, and then stored at 4°C.

All circuit plasmids contained the ampicillin resistance marker, which was added to the LB broth medium and agar plates at a concentration of 100µg per ml (ForMedium^{tm}).

2.6 Transformation of competent *E. coli* TOP10 cells

0.5µl of plasmid was transferred to 100µl of competent *E. coli* TOP10 cells and incubated on ice for 30 minutes. The cells were heat shocked at 42°C for 1 minute and then incubated on ice for 5 minutes. 1ml of LB broth was added, and the cells incubated for 1 hour at 37°C. 200µl of cells was plated onto LB agar containing the appropriate antibiotic resistance marker for the plasmid being transformed. For the gene circuit plasmids all cells were plated onto LB agar containing 100µg per µl ampicillin.

2.7 MINIPrep Plasmid Purification

Plasmid DNA was obtained from 2ml overnight cultures of *E. coli* TOP10 cells in liquid LB broth (containing 100µg per ml) ampicillin), using the QIAgen QIAprep Spin miniprep kit (QIAgen catalogue number 27106). A 2ml culture of *E. coli* cells harbouring the plasmid for purification was grown in suitable selective liquid broth culture overnight (for a maximum of 16 hours). 1ml of culture was transferred to a 2ml centrifuge tube, and centrifuged at 4,000g (13,000rpm) for 5 minutes. The supernatant was discarded and the pellet re-suspended in 250µl of buffer P1. The protocol continued as directed in the QIAgen QIAprep Spin Miniprep Kit handbook for

plasmid purification with a micro-centrifuge. The DNA bound to the miniprep column was eluted using 50µl of sterile distilled water.

2.8 Manual Miniprep Plasmid Purification Protocol

When large numbers of plasmid purifications were required, manual purification of plasmid DNA was performed to save using commercial equipment and reagents. The protocol was adapted from the cleared lysate procedure of D. B. Clewell and D. R. Helinski [218]. The advantages of this method are; high yield (as opposed to alkaline lysis), high purity; suitable for restriction, ligation, sequencing, and transformation of *E. coli* and *S. cerevisiae*; and after additional phenol-chloroform extraction, *in vitro* transcription. The method is also simple and fast, enabling DNA purification of 24 samples using a single microfuge tube, without organic extractions, in under one hour (Hughes J. (2009), Personal Communication. (McCarthy lab).

2.8.1 Reagents

2.8.1.1 25% sucrose

Sucrose solution consisted of a 25% w/v of solution of sucrose in 50mM Tris pH 7.5

2.8.1.2 Lysozyme

Lysozyme solution consisted of 4mg/ml Lysozyme (Sigma), 0.1mg/ml RNase A in 50mM Tris, pH 7.5, 50% glycerol. Lysozyme was stored at -20°C.

2.8.1.3 Triton Lytic Mix

Triton lytic mix consisted of 50mM Tris, pH 7.5, 20mM EDTA, pH8.0, and 0.1% Triton X-100.

A single colony of transformed *E. coli* TOP10 was inoculated into 1.5ml LB growth media containing appropriate antibiotic for the transformed plasmid, and grown overnight at 37°C. The culture was transferred to an 2ml centrifuge tube and centrifuged at 3,000rpm for 1 minute. The supernatant was discarded, and the pellet re-suspended in 200µl of 25% sucrose solution. 20 µl of lysozyme solution was added and the suspension was vortexed briefly, then incubated for 5 minutes at room temperature. Following incubation 400µl of triton lytic mix was added followed immediately with the addition of 80µl of 8M potassium acetate. The cell solution was mixed by inversion several times and the cells incubated on ice for 5 minutes. Following incubation on ice, the tube was centrifuged for 20 minutes at maximum speed at room temperature. After centrifugation the supernatant was transferred to a clean 1.5ml centrifuge tube and the pellet was discarded. 0.5ml isopropanol was added to the supernatant and mixed several times by inversion. The solution was then centrifuged for 10 minutes at maximum speed to sediment the plasmid DNA. Following centrifugation the supernatant was discarded and the pellet re-suspended in 0.5ml 70% ethanol, and centrifuged at maximum speed for 1 minute. The supernatant was discarded and the pellet allowed to air dry for 10 minutes. Following air-drying, the pellet was re-suspended in 10µl sterile distilled water and incubated at 70°C to inactivate remaining DNase activity. The plasmid suspension was then stored at -20°C until required.

2.9 Plasmid DNA Restriction digest

Plasmid DNA was digested with restriction enzymes purchased from Fermentas and New England Biolabs. T4 DNA ligase was purchased from New England Biolabs. Standard recombinant DNA techniques were used, as in Sambrook *et al* [219]. For ligations of DNA with incompatible end structures, the ends were first made flush with Klenow DNA polymerase (for 5′ extensions) or T4 DNA polymerase (for 3′ extensions).

2.9.1 Analytical Plasmid DNA Digest

Analytical digests were performed to check for correct restriction sites, complete digestions, and DNA miniprep yields. In a 20µl total volume, 5µl of DNA for analysis was added to 2µl of appropriate restriction enzyme buffer. 0.5µl of appropriate restriction enzyme(s) was added, and the volume made up to 20µl with distilled water. The digest was incubated at 37°C for 1-3 hours. Following incubation appropriate volume of loading dye was added and the analytical digest observed using gel electrophoresis.

2.9.2 Preparative Digest

Preparative digests were performed to digest plasmid DNA in preparation for gel extraction. A larger total volume of 200µl was used to maximize DNA yield from the digest. 40µl DNA was added to 20µl of appropriate enzyme buffer. 3µl of restriction enzyme(s) was added, and the volume made up to 200µl with distilled water. The digest was incubated at 37°C overnight. Following incubation, the digest was phenol extracted and ethanol precipitated, and the resulting DNA pellet re-suspended in 15µl distilled water. A microlitre of the DNA preparation was observed using gel electrophoresis to check the yield. The remaining DNA was loaded onto an agarose gel and run at 100V for 30 minutes before extraction of the appropriate bands.

2.10 Cranenburgh Ligation Method

DNA and RNA was quantified using a Thermo Scientific Nanodrop 2000 spectrophotometer. The method detailed by Cranenburgh, 2004 was employed for ligating DNA fragments. The method utilizes equations 2.1 and 2.2.

$$V_v = \frac{T}{(\frac{V_c \cdot I_i \cdot I_r}{I_c \cdot V_i}) + 1} \tag{2.1}$$

$$I_v = T - V_v \tag{2.2}$$

The equations above enable the determination of volumes required in a ligation reaction. Insert and vector parameters must both be in the same units (e.g. kilobases for length and micrograms per microlitre for concentration). I_l Insert length, V_l vector length, I_c insert concentration, V_c vector concentration, I_r required insert-to-vector concentration, T volume of DNA solution component, V_v vector volume, I_v insert volume. The I_r should be inserted as insert/vector (e.g 2 for a two-fold excess, 0.5 for a two-fold vector excess) [220].

2.11 Primer Design

Primers for amplification of DNA were designed by obtaining the nucleotide sequence from the EBI online database and isolating the ATG start codon for the gene of interest and cross checked against the corresponding protein sequence in the Pubmed online database. The primer sequence was extended 20 bases in the $3'$ direction from the start codon with the required DNA restriction site added to the $5'$ end with 8 nucleotides upstream. The $3'$ end of the primer contained a guanidine and cytosine nucleotides, forming a "GC clamp" to stabilize the association of the primer with the DNA to be amplified. The same procedure was used for the $3'$ to $5'$ antisense strand primer, starting at the stop codon and adding the appropriate restriction enzyme recognition sites and 8 nucleotide overhand added to the start of the primer. The melting temperatures of the primers were calculated using the Eurogentec online "melt temp calculator" during purchase of the primers and checked to be of a similar temperature and the primers ordered from Eurogentec.com.

2.12 PCR

PCR amplification was performed using the Expand Long Template PCR System kit from Roche (catalogue number 11681834001), using a BioRad C-1000 PCR machine. Using approximately 100ng of template DNA, the reaction mix was formulated by adding 4µl of deoxyribonucleotide mixture (containing 2.5mM of each nucleotide), 5µl of 10x buffer 2, and 2.5µl of a 2.5mM primer solution was added for each primer. 0.5µl Expand DNA polymerase (2.5 units) was added, and the reaction mix made up to 50µl with sterile distilled water.

The PCR reaction was performed using a program consisting of an initial heating period of 94°C for 5 minutes, followed by a cycle of 94°C for 30 seconds, the primer melting temperature for 1 minute per kb, 72°C for 1 minute. The cycle was repeated 25-30 times, then 72°C for 7 minutes before holding at 4°C.

2.13 Colony PCR Protocol

Colony PCR was performed using a BioRad C-1000 PCR machine and the REDTaq DNA polymerase kit (Sigma Aldrich catalogue number D4309). Colony PCR was used to confirm the insertion of cloned DNA into plasmids prior to sequencing. The PCR reaction mix was formulated with 0.2µl of each primer (from a primer stock of 2.5mM) and 10µl of REDTaq DNA polymerase mix. A single colony of transformed *E. coli* TOP10 was transferred from the transformation plate and inoculated into the colony PCR reaction mix, and into a 2ml culture of LB medium containing the approporate antibiotic for the selection marker on the plasmid. The PCR run was performed as above.

2.14 Genomic DNA Extraction

E. coli TOP10 genomic DNA was extracted by centrifuging an overnight culture of cells at 5000rpm for 5 minutes. The supernatant was discarded and 200µl of extraction buffer added to the pellet. 200µl of phenol:chlorophorm isoamyl 25:24:1 was added to the solution with a volume of glass beads equivalent to the size of the pellet. The mixture was vortexed for 2.5 minutes and then centrifuged at 13,000rpm (or maximum speed of the benchtop centrifuge) for 5 minutes. The top layer of the supernatant was extracted into a fresh 1.5ml centrifuge tube and phenol-chlorophorm extracted, before ethanol precipitation. The DNA was re-suspended in 50µl sterile distilled water.

2.14.1 Extraction Buffer

The genomic DNA extraction buffer consists of 2% Triton X-100, 1% SDS, 100nM NaCl, 10mM TrisCl (pH 8.0), and 1mM EDTA (pH 8.0).

2.15 Site Directed Mutagenesis Protocol

Site directed mutagenesis was performed using the Westlab Quickchange II site-directed mutagenesis kit (catalogue number 200524). A set of forward and reverse primers containing the modified sequence for insertion were manually designed and ordered from Eurogentec. The PCR reaction mix was formulated with 5µl 10x buffer, 0.2µl double stranded DNA template (5-50ng final concentration), 1µl of forward and reverse mutagenic primer (120ng final concentration), 1µl mixture of deoxyribonucleotides containing 2.5mM of each base, and 1µl PFU Ultra enzyme (2.5U per µl). The reaction mix was made up to a final volume of 50µl with sterile de-ionized water.

2.15.1 Site Directed Mutagenesis PCR Reaction Program

Site directed mutagenesis was performed using the BioRad C-1000 PCR machine. The program was 95°C for 30 seconds, followed by 18 cycles of 95°C for 30 seconds, 55°C for 1 minute, 68°C for 30 seconds. Following the PCR amplification cycles, 1µl of DPNI enzyme (10 Units per µl) was added to digest the parent plasmid. The PCR mix was incubated at 37°C for 1 hour before 1µl was used to transform 100µl of *E. coli* XL Gold ultra-competent cells, provided with the mutagenesis kit.

2.16 Phosphorylation and Annealing of Synthetic Oligonucleotides

Olignonucleotides for annealing, such as the iron response element and LexA operators were ordered from Eurogentec as a set of complementary sequences that require phosphorylating and annealing prior to ligation into a plasmid construct. Phosphorylation was carried out using 50pmol of each oligonucleotide (0.5µl of 100µmol stock solution), 200pmol of ATP (0.5µl of 100 µmol stock solution), and 5µl of 10x PNK. The phosphorylation reaction was incubated at 37°C for 30 minutes, and then 95°C for 3 minutes. The oligonucleotides were then placed in a beaker of boiling water and allowed to cool to room temperature to anneal.

2.17 Agarose Gel Electrophoresis

1% agarose gels were used for visualizing and obtaining plasmids and DNA fragments. Agarose was melted in 1x TAE buffer using a microwave. 100ml of molten agarose was transfered to a casting tray and allowed to cool slightly before the addition of 3µl of 2mg/ml ethidium bromide. The ethidium bromide was mixed into the gel and allowed to cool and solidify with a comb of appropriate size for the number of samples to be loaded. DNA and plasmid preparations were mixed with DNA loading dye. Samples

were run for 30 minutes at 100V with the gel immersed in TAE buffer in a gel running tank (BioRad). 5µl of Fermentas GeneRuler or New England Biolabs DNA ladder was added next to the samples for size comparison and estimation. 1kb and 100bp ladders were used depending on the sample DNA length. Agarose gels were visualized using an ultra-violet light transilluminator and a Kodac Gel Logic 100 imaging system.

2.17.1 TAE buffer - 5 Litre, 10x stock

TAE buffer was made as a 10x stock and diluted to 1x for each experiment in distilled water. The 10x stock was prepared with 242g Tris base, 57ml glacial acetic acid, 100ml 0.5M EDTA pH8.0, made up to 5L with distilled water.

2.17.2 Preparation of DNA loading dye

DNA Loading dye was prepared as a 6x concentrated stock and diluted in the sample to a final concentration of 1x. The 6x stock was prepared with 10mM Tris-HCl (pH 7.6), 0.03% bromophenol blue, 0.03% xylene cyanol, 60% glycerol, and 60mM EDTA.

2.18 Yeast Transformation

The yeast transformation protocol is based on the method by Guldener *et al* [221].

2.18.1 Preparation of Solutions and Growth Media for Yeast Transformation

Yeast cells were grown overnight at 30°C in a baffled shake flash containing 5ml YP growth medium with 2% glucose (YPD).

2.18.1.1 Preparation of 10x LiAc and 10X TE solution for yeast transformation

10x LiAc consisted of 1M LiAc pH7.5, 10x TE consisted of 0.1M Tris, 0.01M EDTA pH7.5. 200ml volumes of each were prepared and sterilized by autoclaving.

2.18.1.2 Preparation of 20ml PEG/LiAc/TE solution

PEG/LiAc/TE solution was prepared by mixing 8g PEG 2000, 2ml 10x LiAc, 2ml 10X TE, 9.75ml distilled water and filter sterilized with a 0.2µl syringe filter. The solution was stored at room temperature for a maximum of 1 month.

2.18.1.3 Preparation of YP agar

500ml of 2xYNB and 200ml of 2x CSM knock out amino acid media were prepared and stored separately until required, at which time they could be diluted together forming 1xYNB. Glucose was added after autoclaving by filter sterilization, forming YPD medium.

2.18.2 Yeast transformation protocol

Optical density measurements were collected using a Shimadzu UVMini 1240 spectrophotometer. Yeast cells were transformed by diluting the 5ml overnight culture to a starting OD_{600nm} of 0.2 and then growing the cells to 0.7-1.0 OD_{600nm} (approximately 3-4 hours). Salmon sperm carrier DNA was boiled at 95°C for 10 minutes and cooled to room temperature. Cells were harvested by centrifugation at 4,000rpm for 5 minutes and re-suspended in 10ml sterile distilled water. The centrifugation was repeated and the pellet re-suspended in 1ml sterile water. The cell suspension was centrifuged at 5000rpm for 1 minute and the pellet re-suspended in 1.5ml TE/LiAc solution. The cells were centrifuged at 5,000rpm for 5 minutes and the pellet re-suspended in 1ml TE/LiAc. 10 µl of plasmid DNA was added to 5µl (50µg) of carrier DNA, and 50µl of cells. 300µl of PEG/TE/LiAc was added to the mixture and incubated at 30°C in a shaking

incubator for 30 minutes, and then transferred to 40°C for 15 minutes. Following the 40°C heat shock step, 800µl of sterile distilled water was added and the cells harvested by centrifugation at 13,000rpm for 10 seconds. The pellet was re-suspended in 1ml YPD broth and incubated at 30°C for 2-3 hours in a shaking incubator. Following incubation, the cells were centrifuged at 13,000 rpm for 10 seconds and the pellet re-suspended in 200µl of YPD and plated onto YP knock out media lacking the appropriate auxotrophic marker for the plasmid being transformed. Agar plates were incubated for 2-3 days at 30°C.

2.19 Yeast Protein Extraction

Total cell protein was extracted from *S. cerevisiae* cells using the methodology from Von der Haar [222]. Cells were collected at an equivalent optical density of 4x OD_{600nm} of 1.0. The cells were harvested by centrifugation at 4,500rpm for 5 minutes and the pellet washed in 1ml sterile distilled water. The cells were centrifuged at 13,000rpm for 10 seconds, the supernatant discarded and the pellet snap frozen at -80°C. The frozen pellet was re-suspended in 100µl lysis buffer, and boiled at 95°C for 10 minutes. 3µl of 4M acetic acid was added to the cells and boiled for 95°C for 10 minutes. 25µl of SDS sample buffer was then added to the cell suspension and used immediately for western blot, or stored at -20°C until required.

2.19.1 Lysis buffer

Yeast cell lysis buffer consists of 0.1M NaOH, 0.05M EDTA, 2% SDS, and 2% β-mercaptoethanol.

2.19.2 SDS Sample buffer

SDS sample buffer consists of 0.06M Tris-HCl pH6.8, 5% glycerol, 2% SDS, 4% β-mercaptoethanol and 0.0025% bromophenol blue.

2.19.3 Preparation of SDS PAGE Protein Gels

NuSep density gradient (4-20%) pre-cast protein gels were purchased from Generon (http:www.generon.com) (catalogue number NH31-420, 15 sample wells per gel).

2.20 Western blotting

2.20.1 Polyacrylamide gel electrophoresis protocol

10µl of protein extracts was loaded on to 15% pre-cast NuSep density gradient pre-cast gels. Samples were loaded alongside 10µl of Fermentas PageRuler pre-stained protein ladder plus marker for size estimation. Protein gels were run at 80V for approximately 1 hour.

2.20.2 Western Blot Transfer protocol

A semi-dry pierce fast-transfer deck (from Thermo Fisher) was used for all western blots. Nitrocellulose filter (Hybond C from Amersham Biosciences) was used for increased resolution over PVDF film. Blotting paper and film were cut to the appropriate size to cover the gel during the transfer, and soaked in transfer buffer. Protein gels had the stacking gel removed and were washed in 10ml transfer buffer. 2 pieces of blotting paper were transferred to the transfer deck, followed by the nitrocellulose film, the gel, and 2 further pieces of blotting paper. A clean stripette was used to roll across the surface of the blotting stack to remove any air bubbles. The transfer deck was wetted with transfer buffer and then closed over the blotting stack. The transfer was run at 400mA constant amps with a limit of 25V for 1 hour.

Following the transfer, the blotting stack was discarded and the film incubated in Ponceau S staining solution (Thermo Fisher) to check for efficient transfer of protein to the membrane. The Ponceau S stain was removed by briefly rinsing the membrane in distilled water. The membrane was then incubated for 10 minutes in miser antibody

extender solution (Thermo Scientific) to enhance primary antibody binding. The membrane was then washed 5 times in distilled water before blocking with PBST+5% milk (phosphate buffered saline + 0.1% Tween20 + 5% milk) for a minimum of 1 hour at room temperature, or overnight at (4°C) on a rocking platform.

2.20.3 Antibody binding

Primary antibody concentration was calculated from the accompanying data sheet for the relevant antibody. Luciferase, LexA, and IRP antibodies were purchased from Abcam and used at a concentration of 1:200. Primary antibody was prepared in 5 % milk PBS + 1:100 10% sodium azide solution. FITC and horseradish peroxidase conjugated secondary antibody were prepared at a concentration of 1:20,000 in 5% milk PBS without sodium azide. Alkaline phosphatase conjugated secondary antibody was prepared at a concentration of 1:5000 in 5% milk PBS solution. All antibody solutions were stored at 4°C in the dark.

The membrane was sealed in plastic sheeting using a heat sealer and 1ml of the primary antibody added to the membrane. The membrane was then incubated overnight at 4°C. Following incubation, the primary antibody was collected (for re-use) and the membrane rinsed 3 times for 10 minutes in 5ml PBST solution. 5ml of secondary antibody solution was then added to the membrane and incubated for a minimum of 1 hour at room temperature of a rocking platform. Following incubation the secondary antibody was collected for re-use, and the membrane washed 3 times for 10 minutes in 5ml PBST, or 10 times if using the alkaline phosphatase secondary antibodies.

2.20.4 Western Blot Imaging

The membrane was transferred to a clear plastic sleeve and 1ml ECL reagent applied to the surface of the membrane. The film was left at room temperature for 1 minute in the dark for the membrane to react with the ECL. The excess ECL was removed and the membrane placed in a photographic cassette (Kodac) and transferred to the dark

room for developing. Photographic film was applied to the membrane and exposed for 5 minutes before being transferred to the developer. Longer time periods can be used depending on the signal from the membrane.

2.20.5 Alkaline Phosphatase Protocol

For alkaline phosphatase reaction, the alkaline phosphatase buffer was prepared in a 50ml falcon tube and consisted of 1ml 1M Tris-HCl pH 9.5, 0.25ml 4M NaCl, 50µl 1M $MgCl_2$, 8.7ml deionized water for a 10ml final volume. 66µl of NBT was added and mixed, followed by 33µl of BCIP. The solution was mixed and 5ml applied to the membrane. The membrane was incubated in the dark at room temperature on a rocking platform and regularly observed for the development of a colour reaction on the membrane. Once sufficient bands have developed the reaction can be stopped by washing the membrane in distilled water.

2.20.6 Quantification of Western Blot Images

For quantification of western blots, the membrane was scanned on a flatbed scanner (Canon) and the image saved as an 8bit (greyscale) uncompressed TIFF image. Images were then quantified using the Gel-Pro Analyzer software (version 3.1) from Media Cybernetics. Data was then transferred to Microsoft Excel and OriginLab Origin for plotting and further analysis.

2.21 DNA Sequence Alignment

Sequence alignment was performed using the ClustalW software package, version 2.1 compiled for Linux from source code available from http://www.clustal.org. [223]

Analysis of plasmid constructs was performed using Sequencher 4.9 from Gene Codes.

2.22 DNA Primer Design

Primers were designed manually by identifying the transcription start site from the literature concerning the gene of interest, finding the nucleotide sequence in online databases such as NCBI and matching the nucleotide sequence with the protein database sequence to locate the open reading frame. The primer was then constructed using 20 bases upstream of the 5′ to 3′ start site and 20 bases downstream of the 3′ to 5′ start site. Suitable restriction enzyme sites were added to the 5′ and 3′ primers together with 8 random nucleotides at each end to facilitate binding to the target sequence. Primers were ordered online from Eurogentec.

2.23 Pheromone Induction of Yeast Cells for Luminescence Assay

A 10ml overnight culture of yeast cells were grown in a 50ml baffled shake flask (incubated in a shaker incubator at 30°C, 250rpm), and diluted the following morning to an optical density of 0.2 OD_{600nm} in 25ml fresh YP broth medium (containing appropriate amino acid drop out media, and 2% glucose) in a 100ml baffled shake flask (incubated in a shaker incubator at 30°C, 250rpm). Cells were then grown to an optical density of approximately 0.6-0.8 OD_{600nm} and then stimulated with 100nM of alpha-factor pheromone (Zymo Research Y1001).

2.24 Optical Density Measurements

Optical density measurements were prepared with a 1:10 dilution of the yeast culture in a 1ml volume of YP both, in a 1.5ml SemiMicro Cuvette (Starlab catalog number E1412-4150). Optical density Measurements were made using a Shimadzu UVmini-1240 UV-VIS spectrophotometer. Optical density measurements were used to estimate

the cell growth phase for pheromone induction, and the cellometer was used to calculate cell number for luminescence data normalisation.

2.25 Cellometer Cell Measurements

In addition to measuring biomass using optical density, cell numbers were calculated using a cellometer Auto T4, purchased from Nexcelom Bioscience. 20µl of cells was transferred to a cellometer counting chamber. The chamber was left on the bench for approximately 10 minutes to settle before reading in the cellometer instrument. Counts were obtained using the cellometer software in cells per ml. Cell images were exported along with the raw data in Microsoft Excel format. Relative luminescence units measured by the luminometer were divided by the cell count to convert the measured luminescence to relative luminescence units per cell.

2.26 Yeast Growth Rate Measurements

Yeast cells were grown overnight on YPD complete with 2% glucose, for wild-type cells, or YP *-LEU, -HIS, -URA* with 2% glucose for cells transformed with the circuit plasmids. The culture was diluted to an optical density of $0.001OD_{600nm}$, and the optical density of the cultures measured every hour for 16 hours. The Log base 2 optical density measurements were plotted against time to identify the logarithmic growth phase (log phase), and the growth rate was then determined by calculating 1/slope, and compared across strains.

2.27 Yeast *in situ* Luciferase Assay

The following assay is based on the publication by Vieites J.M. *et al.*, and enables the *in vivo* measurement of luciferase activity without requiring cell lysis [204]. 50µl of each yeast culture was transferred to a luminometer tube containing 150µl m-citrate buffer,

pH3 (0.2M sodium citrate (28.8g/l) and 0.2M citric acid (42.2g/l)) (m-citrate buffer was aerated prior to use by vortexing for 15 seconds prior to each measurement). The buffer enhances the permeability of the membrane, enabling exposure of intra-cellular luciferase to extra-cellular luciferin.

The luciferin stock was prepared at 10mM in 10mM sodium bicarbonate. 10µl of luciferin solution was added to 50µl of cells in 150µl of sodium citrate buffer, making a final concentration of Luciferin of 0.5mM n the reaction mixture.

The luminometer (Berthold Lumat LB 9507) was configured to sample 25µl of the reaction mixture, and measure the average luminescence over 30 seconds. The luminescence measurement (measured as relative luminescence units) was then converted to relative luminescence units per cell using the cell count made using the cellomoter. Three reaction mixtures were prepared for each time point and each mixture measured once. The average of the three luminescence measurements was calculated, and converted to relative luminescence units per cell using the corresponding cellometer cell count.

2.28 Real-time Quantitative PCR (RT-qPCR)

2.28.1 RT-qPCR Primer Design

Primers for RT-qPCR were designed using the PrimerPy software from http://code.google.com/p/oligobench.

2.28.2 mRNA extraction and purification

Frozen cell pellets were thawed on ice for approximately 5 minutes and re-suspended in 750µl chilled TES (10mM Tris-HCl pH7.5, 10mM EDTA, 0.5% SDS). 750µl of phenol-chloroform 5:1 pH 7.4 (Sigma P1944-400ML) was added, and vortexed for 5 seconds. The solution was incubated for 1 hour at 65°C with shaking. Following

incubation the cells were chilled on ice for 1 minute and mixed by vortexing for 20 seconds before being centrifuged at 13,000rpm for 5 minutes. 700µl of the aqueous phase was transferred to a new centrifuge tube and 700µl of phenol-chloroform and mixed by inversion. The suspension was centrifuged at 13,000rpm for 1 minute and the aqueous phase transferred to a new centrifuge tube. 650µl of the aqueous phase was transferred to 650µl of chloroform-asoamyl alcohol (25:1) and mixed by inversion. The solution was centrifuged at 13,000rpm for 1 minute and 500µl of the aqueous phase transferred to a new centrifuge tube. 1.5ml of 100% ethanol was added with 50µl of 3M sodium acetate, pH5.2 and the mRNA precipitated at -80°C for 30 minutes. The solution was then centrifuged at 13,000rpm for 10 minutes and the supernatant discarded. The pellet was washed in 500µl 70% ethanol and centrifuged at 13,000rpm for 1 minute. The supernatant was discarded and the pellet air dried. The pellet was re-suspended in 100µl DEPC water and incubated for 1 minute at 65°C to dissolve the pellet. The mRNA yield was quantified using a Nanodrop spectrophotometer and stored at -80°C.

2.28.3 Turbo DNase protocol

40µg of mRNA was transferred to a new centrifuge tube and 300µl (6 units) of Turbo DNase added. The reaction mixture was incubated for 30 minute at 37°C. The mix was then phenol-chloroform purified and ethanol precipitated as in the above section and the pellet re-suspended in 200µl DEPC water. The mRNA yield was quantified using the Nanodrop spectrophotometer. The DNase treated mRNA was stored at -80°C.

2.28.4 Reverse Transcriptase protocol

The reverse transcription reaction was carried out using the PrimerDesign Precision Reverse Transcription Kit (catalogue number RT-nanoScript). To anneal the RT primers, mRNA template was added at approximatley 1µg (recommended 2ng - 2µg), with 1µl of RT Primer, and made up to a final volume of 10µl with RNase/DNase free water. The reverse transcriptase reaction mixture was then heated to 65°C for 5 minutes using the BioRad C-1000 thermocycler and then placed on ice. Following the annealing

step the reverse transcriptase reaction mix was prepared by adding 2µl of nanoScript 10x buffer, with 1µl of dNTP mix (containing 10mM of each dNTP), 2µl 100mM DTT, 4µl of RNase/DNase free water, and 1µl of nanoScript reverse transcriptase enzyme. The reverse transcriptase reaction mix was then added to the annealed mRNA reaction mix forming a final reaction mixture volume of 20µl.

2.28.5 RT-qPCR protocol

The mRNA quantification was used to calculate an equivalent volume of approximately 25ng cDNA (assuming all of the mRNA in the reverse transcription reaction had been transcribed to cDNA) for the qPCR reaction mix. 5µl of 2x Precision master mix containing SYBR green (PrimerDesign catalogue number Precision-SY) was added to 6µl of the reverse transcriptase reaction and made up to a final volume of 10µl with DNase/RNase free water. 10µmol solutions of each primer were prepared in DNase/RNase free water and 1µl of each primer pair was mixed in a Qiagen PCR tube. 8µl of the qPCR reaction mix containing the cDNA template was mixed with the primer making a 10µl final reaction volume. The tubes were loaded into a Qiagen Rotorgene Q qPCR machine, programmed with a protocol consisting of 10 minutes enzyme activation at 95°C (hot start), 50x cycles of 15 second denaturing at 95°C followed by 60 seconds at 60°C for data collection. The qPCR cycles were then completed with a melt curve, pre-programmed by QiaGen. The amplification efficiency wash checked to be above 1.6 and the data was then analysed using the RotorGene Q series software provided by Qiagen.

2.29 Mathematical Modelling

Mathematical modelling was performed using the Copasi version 4.7 (build 34) available from http://www.copasi.org, and referenced in Hoops *et. al*. The Kofahl and Klipp model of the yeast pheromone response pathway was obtained in SBML format

Parameter	Value
Modulation Factor	$1x10^{-9}$
Resolution	$1x10^{-9}$
Derivation Factor	0.001
Newton method	1
Integration	1
Back Integration	1
Accept Negative Concentrations	0
Iteration Limit	50
Maximum duration for forward integration	$1x10^{-9}$
Maximum duration for backward integration	$1x10^{-9}$

TABLE 2.4: Copasi metabolic control analysis parameter values.

from http://www.biomodels.net [224]. Additional parameter values were obtained from http://www.bionumbers.org [225].

2.29.1 Metabolic Control Analysis

Metabolic control analysis was performed using the Copasi software package, with steady state analysis and the parameters described in table 2.4.

2.29.2 Sensitivity Analysis

Sensitivity analysis concentration control coefficients were calculated using Copasi, on all non-constant concentrations of species, and all parameter values.

Sensitivity analysis is performed by making a change to all parameters in the reactions of the model by a value delta (Δ) and measuring the change in the steady state values of all the concentrations of the reactants. Copasi uses the current value of the parameter times the delta factor as variation. If this number is smaller than the delta value then a delta minimum value is used.

Copasi sensitivity analysis was performed using the following parameter values: Δ factor of 0.001 and Δ minimum of $1x10^{-12}$.

2.29.3 Metabolic Control Analysis

Metabolic control analysis concentration control coefficients were calculated using Copasi. Metabolic control analysis was performed using the default Copasi parameters with a modulation factor of 1x10^{-9}, resolution of 1x10^{-9}, deviation factor of 0.001, with Newton, integration, and back integration, and an iteration limit of 50. Maximum duration of forward integration was 1x10^{-9} and maximum duration of backward integration of 1x10^{-6}.

2.29.4 Signal to Noise Ratio

The signal to noise ratio was calculated using the formula:

$$SNR = \frac{\mu}{\sigma} \tag{2.3}$$

The mean, standard deviation, and variance were calculated for each time point, and the confidence interval for the mean and the variance used to calculate the error propagation for each time point. Using a confidence interval of 95%, the following formula was used to calculate the confidence limit for the mean:

$$CI_m = x \pm t_{\alpha/2} \cdot (\sigma/\sqrt{n}) \tag{2.4}$$

where CI_m is the confidence limit for the mean, x is the mean of the sample population, σ is the standard deviation, α is the confidence level divided by 100, and $t_{\alpha/2}$ is the t-distribution value, and n is the sample size. For the luminescence assays reported in this work, n=9.

The standard deviation confidence interval for variance was calculated for a confidence interval of 95% using the formula:

$$CI_v = [(n-1) \cdot S^2/X^2_{\alpha/2,n-1}] \leq \sigma^2 \leq [(n-1) \cdot S^2/X^2_{\alpha/2,n-1}] \tag{2.5}$$

where CI_v is the confidence interval for variance, n is the sample size, S is the variance, α is 1-(confidence level divided by 100), and $X^2_{\alpha/2,n-1}$ is the Chi-square table value.

Using CI as $\delta\mu$, CI_v as $\delta\sigma$, together with the mean (μ) and standard deviation (σ), calculated from equations 2.4 and 2.5, the error propagation for the signal to noise ratio at each time point was calculated using the formula:

$$R = SNR \cdot \sqrt{\left(\frac{\delta\mu}{\mu}\right)^2 + \left(\frac{\delta\sigma}{\sigma}\right)^2} \tag{2.6}$$

Where R is the error, SNR is the signal to noise ratio (calculated in equation 2.3), σ is the standard deviation of the mean, and μ is the mean of the sample set for each time point.

2.29.5 Parameter Estimation

For the final models in chapter 5, section 5.6 fitting was performed using the luminescence time course data. The data from each strain was used to fit to the generic circuit model, to create a new model specific to the strain from which the luminescence data was fitted.

Parameter estimation was performed using Copasi. The parameters V_{max}, K_M, K_i, V, k, and s for each of the reactions in the model were allowed to vary between +/- 100%. The promoter strength parameter P was fixed with the corresponding value of zero or sixty, depending on whether the data was from a TEF1 or DCD1 promoter strain. The Hooke and Jeeves optimisation algorithm was run first to obtain a fit to the experimental data, followed by the simulated annealing algorithm to confirm and/or improve on the Hooke and Jeeves algorithm. The fitted model was then saved as a separate model file, producing four models corresponding with each of the four circuit strains.

The parameters for the Hooke and Jeeve's algorithm were: an iteration limit of 50, a tolerance of $1x10^{-5}$, and a Rho value of 0.2. The simulated annealing algorithm was used with a start temperature of 1, cooling factor of 0.85, and tolerance of $1x10^{-6}$.

2.30 Dissertation

Graphs were produced using LibreOffice version 3.4.5 available from http://www.libreoffice.org, and Origin 8.5, available from http://www.originlab.com. Statistical analysis was performed using IBM SPSS version 19, available from https://www-01.ibm.com/software/analytics/spss/.

Additional stochastic data analysis was performed using the Python programming language, version 2.6 available from http://www.python.org.

The dissertation was written using an IBM compatible personal computer with LaTeX running Ubuntu 11.10, available from http://www.ubuntu.org.

CHAPTER

THREE

RESULTS - CIRCUIT CONSTRUCTION

3.1 Introduction

Three plasmids were designed to exploit the cellular machinery associated with the yeast mating response. The presence of, and thus the cellular response to, α-mating factor drives the simultaneous expression of plasmid-derived proteins that regulate the experimental gene circuit *in vivo* (figure 3.4). The circuit was constructed from three individually well characterized genetic components; the luciferase reporter gene [209], the *E. coli* LexA repressor [203], and the mammalian IRP repressor [4]. The P_{FUS1} promoter from the yeast mating response pathway has been used in a number of publishes studies to enable reporter genes to be expressed in response to pheromone stimulation [1, 3, 6, 7, 211, 226, 227].

3.1.1 The Iron Responsive Element-Binding Protein

The human iron response protein (IRP) is a cytosolic, RNA binding protein that has been well characterised in its ability to repress the expression of iron metabolism genes [206, 207, 228–231]. The IRP regulates the translation of encoding proteins for iron

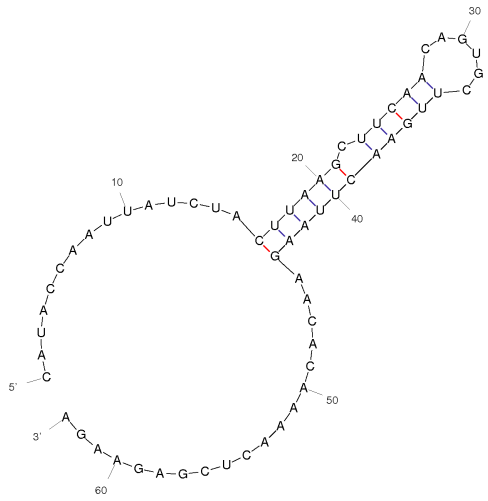

FIGURE 3.1: The iron response element nucleotide sequence was converted into an RNA sequence and a hypothetical secondary structure was calculated using the m-fold software package with an initial ΔG = -8.40 kcal/mol. A single structure was predicted and demonstrates the hair-pin loop structure that binds the IRP to block translation [4].

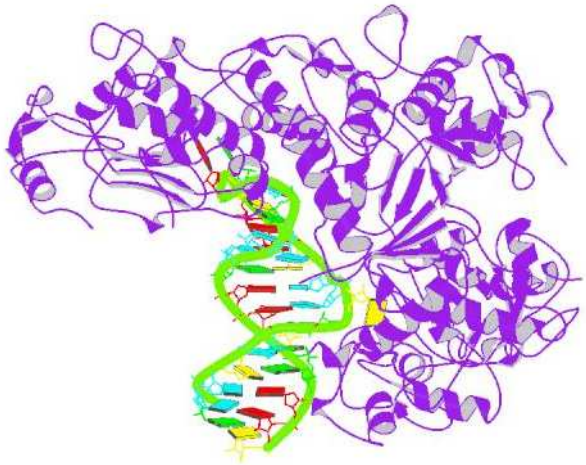

FIGURE 3.2: Crystal structure of the iron response protein in complex with ferritin mRNA. Adapted from Walden *et al* [232].

transport and storage in mammalian cells by blocking the initiation of translation by the 40S ribosomal subunit [4, 231, 232]. It has been shown that the IRP binds to a 62 nucleotide stem-loop structure(figure 3.1) in the mRNA template of iron response genes (figure 3.2) with high affinity (K_d 10^{-10} - 10^{-11}) and is sufficient to block translation [4, 205, 232]. Koloteva *et al* have shown that the IRP functions in yeast, and can bind to the IRE placed upstream of reporter genes [4]. Koloteva *et al* also investigated the

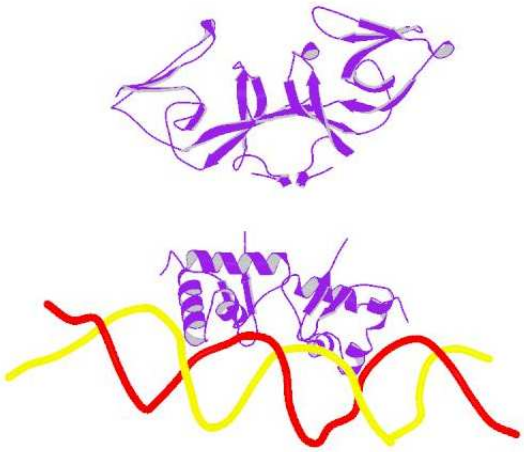

FIGURE 3.3: Unrefined crystal structure of a LexA-DNA complex, adapted from Zhang *et al* [240]

position of the IRE relative to the start codon on the mRNA molecule and found the IRP is able to inhibit translation if it is located at the 5′ end of the molecule and within 50 nucleotides of the mRNA CAP structure. When the IRP is not bound to the IRE, the IRE does not interfere with translation as the 40S ribosomal subunit is able to overcome the structural resistance of the stem-loop structure during the scanning process and initiate translation at the start codon [4].

3.1.2 The LexA DNA Binding Protein

LexA is an *E. coli* transcriptional repressor that represses the SOS response genes coding for DNA polymerases required for repairing DNA damage [233]. LexA binds with high affinity to a specific DNA recognition sequence called the “LexA operator”, and blocks RNA polymerase-mediated transcription of downstream genes. [233–235] (figure 3.3). LexA has been well characterised in bacteria [233, 235, 236] and has been shown to function as a transcriptional repressor in mammalian cells [237] as well as yeast [5, 203, 238, 239].

3.1.3 Yeast Promoters

There are a number of varying strength promoters available in yeast that have been well characterized and provide the capability to vary the expression levels of the genes that they regulate [241]. The *DCD1* promoter (P_{DCD1}) from the yeast dCMP deaminase gene, provides a weak promoter for the gene circuit and was designed to express a low number of mRNA transcripts from the genes it regulates [212, 242]. The TRP1 promoter (from the yeast tryptophan biosynthesis pathway) (P_{TRP1}) is a similar strength promoter to P_{DCD1} that has also been well studied in yeast [243]. The *TEF1* promoter is a strong promoter in that it expresses a high number of mRNA transcripts from the genes it regulates [212, 244]. The combination of these different strength promoters enables differentiating the expression levels of the components of the circuit and tuning the expression of the reporter gene.

Individually, these components provide a tool-kit of parts that can be combined to form synthetic gene circuits. The pheromone response pathway itself forms a generic input module that interfaces these parts with the external environment and enables the circuit to be activated by the addition of an extra-cellular stimulus.

One of the problems encountered when assaying reporter genes and building synthetic circuits with discrete behaviours is discriminating between gene activation and background expression [245–247]. To this end, the gene circuit attempts to reduce noise from background (basal) expression, and attempts to increase the ratio between the inactive and active state of the reporter gene. The circuit also enables the investigation of two separately well characterised genetic control elements (the IRP and LexA repressors) in a synthetic biology application.

3.2 Circuit Overview

3.2.1 Design overview

Figure 3.4 outlines the design of the gene circuit constructed for this project. The luciferase reporter gene was placed downstream of the P_{FUS1} pheromone response pathway promoter, interfacing the reporter gene with the pheromone response, enabling the cells to express luciferase in response to pheromone. The IRE was placed upstream of the luciferase reporter gene, interfacing the reporter plasmid with the repressor plasmid. The repressor plasmid contained yeast constitutive promoters controlling the expression of the IRP gene, facilitating constitutive expression of IRP and repression of luciferase mRNA translation. A third plasmid de-represses luciferase by expressing LexA, also from the pheromone response pathway P_{FUS1} promoter. Upon pheromone stimulation, the cells express LexA which binds to a set of LexA binding domains on the repressor plasmid up-stream of the constitutive promoter. De-repression occurs simultaneously with expression of luciferase, lifting repression on the reporter gene, which is itself up-regulated. This mechanism of controlled repression of the reporter gene enables repression of basal expression that is de-repressed upon activation, boosting the induction ratio through reduced background activity.

3.2.2 Component Interactions

The reporter plasmid expresses a reporter gene that can be measured experimentally providing a means of observing the behaviour of the circuit experimentally (figure 3.4 right). The repressor protein functions at the level of translation, inhibiting the progression of the ribosome translocation along the mRNA molecule (figure 3.4 middle) [4]. The de-repressor plasmid expresses the LexA inhibitor that binds to a specific DNA recognition sequence (the LexA operator) cloned upstream of the promoter in the repressor plasmid, blocking transcription of the IRP gene (figure 3.4 left). This approach produces repression on two different time scales. Repression of the luciferase reporter gene is fast as the IRP blocks the translation of the reporter gene mRNA.

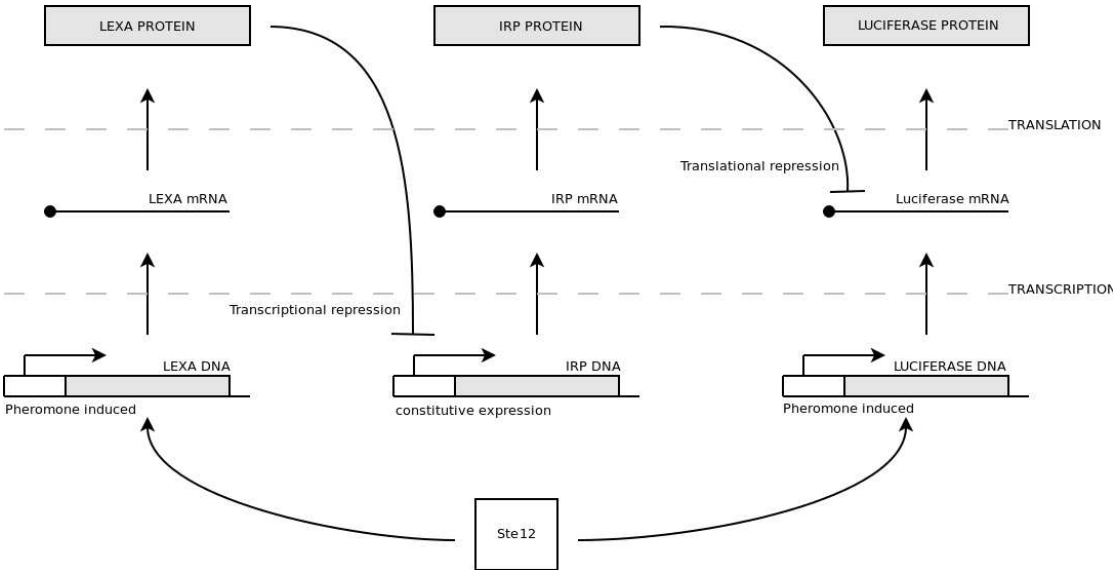

FIGURE 3.4: The circuit is comprised of three plasmids that express proteins that interact to form an discrete circuit in the host cell. The circuit uses the yeast mating response to induce a quantifiable reporter gene (luciferase) (right), the expression of which is controlled through the interactions of the products of the other two plasmids (left and middle). The repressor plasmid (middle) contains a constitutive yeast promoter, expressing a repressor protein (IRP) that continuously represses mRNA translation of the luciferase reporter gene, forming the circuit OFF-state. Upon pheromone stimulation of the cells, the yeast mating response is activated and the Ste12 transcription factor up-regulates the expression of the de-repressor plasmid that expresses its own repressor protein (LexA) which represses transcription of the IRP gene (right). The reporter plasmid is also up-regulated by Ste12 and expresses the luciferase reporter gene, forming the circuit ON-state. The interaction of the three components ensures repression of basal (non-induced) expression of the reporter gene and maximizes a switch-like response when the circuit moves from the OFF-state to the ON-state.

Depending on the half life of the reporter gene, this blocks the accumulation of the reporter protein directly. Repression of IRP by LexA is slower however as LexA inhibits progression of RNA polymerase during transcription. Residual IRP mRNA molecules that were transcribed before up-regulation of LexA (activation of the circuit) circuit will continue to be translated into functional repressor proteins that can continue to repress reporter gene translation. De-repression is therefore a function of the repressor mRNA *and* protein degradation rates. Repression of the reporter is only a function of the degradation rate of the luciferase protein. The circuit can therefore inhibit expression of the reporter gene under normal growth conditions, and upon pheromone activation, the circuit can begin simultaneously de-repressing and up-regulating the reporter gene.

3.2.3 Overview of Luciferase Gene Expression Tuning

Understanding the interaction of the components is key to understanding the performance characteristics of the circuit, and the ability to "tune" the controlling elements enable the optimization of the circuit towards maximal reporter gene expression. Understanding the effect of simultaneous interactions however is difficult to achieve analytically, therefore systems biology techniques of mathematical modelling were employed to provide insight into the interactions prior to construction of the gene circuit, when there is an absence of experimental data.

The first round of modelling indicated the *IRP* gene's rate of transcription, and degradation rate of the IRP protein exerts the most control over the output of the circuit (chapter 5, and figure 5.7). Therefore, additional circuit design variations were incorporated into the laboratory construction phase that included strong and weak promoters, enabling high and low expression levels of the IRP.

In addition to tuning the expression level of the IRP repressor, modelling indicated that the degradation rate of the IRP is also a key control point of the circuit. A degradation tag was also designed and added to the IRP to provide a short and long half-life variant of the repressor protein. Perturbing the half-life of the IRP was predicted to be an effective method of controlling the rate of repression of the luciferase reporter gene. A literature review of protein degradation tags revealed previous work by Mateus and Avery, where the yeast-optimized GFP (yEGFP3) was fused to the C-terminus of the constitutively unstable yeast C_1 cyclin, Cln2 protein [216]. The C-terminus residue contains PEST motifs of Cln2 that are thought to target the protein for ubiquitin (Ub)-dependent degradation [248–250]. This form of degradation is constitutive and therefore does not require heat induction or activation by a ligand molecule, as found in other degradation tags [181]. The PEST tag has been shown to reduce the half-life of human thymidine kinase from 2 hours to 12 minutes, and reduce the half-life of eGFP from 7 hours to approximately 30 minutes [216].

There is no published half-life for the IRP, but some researchers have recorded a half-life of greater than 12 hours [205, 251, 252]. If a 10-fold reduction in the half-life of

IRP can be achieved with the addition of the C-terminus PEST tag it would result in IRP with a half-life of approximately 1 hour which is within the time-scale of the yeast pheromone response.

3.3 Construction of the Reporter Plasmid

The reporter plasmid was constructed using the pRS315 plasmid backbone which is one of a series of pBluescript-based centromere vectors (NCBI accession number U03441), created by Sikorski and Hieter [213]. The plasmid contains the ampicillin antibiotic resistance marker for selection in bacteria and the *LEU2* auxotrophic marker for selection in yeast (table 2.1). The pheromone response pathway promoter P_{FUS1}, from yeast was cloned into the pRS315 multiple-cloning site (MCS) between BamHI and NdeI restriction sites, prior to the start of this project by the McCarthy lab. The luciferase gene was cloned from plasmid pJM4 from the McCarthy lab plasmid collection, utilizing the HindIII and NdeI sites downstream of the P_{FUS1} promoter so that expression of luciferase was pheromone-inducible via the yeast mating response (see figure 3.6). The P_{FUS1} promoter links expression of the reporter gene to the pheromone response pathway which provides a signal input module for the circuit.

3.3.1 The Luciferase Reporter Gene

The luciferase/luciferin bioluminescent system is found in the firefly (*Photinus pyralis*) [253] (figure 3.5) *In vitro*, the activity of the luciferase enzyme is assayed by the addition of its substrate firefly Luciferin, ATP and magnesium ions. Luciferase oxidizes ATP-activated luciferin through a dioxetanone intermediate, and produces carbon dioxide and oxyluciferin in an excited state which decays quickly, emitting a yellow-green light with a high quantum yield [204, 254]. The luciferase reporter gene provides a method of assaying the behaviour of the circuit with high sensitivity compared with other common methods such as fluorescent proteins (McCarthy J. (2009) Personal communication (McCarthy Lab). The firefly luciferase gene was obtained from plasmid pJM4 from

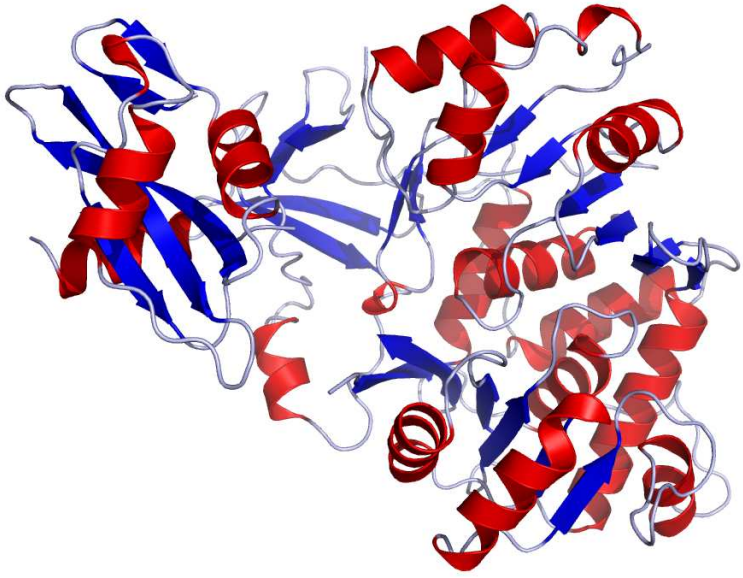

FIGURE 3.5: Crystal structure of firefly luciferase at 2.0Å resolution. The protein is folded into two compact domains. The large N-terminal domain consists of a beta-barrel and two beta-sheets. The sheets are flanked by alpha-helices to form an alphabetaalphabetaalpha five-layered structure. The C-terminal portion of the molecule forms a distinct domain, which is separated from the N-terminal domain by a wide cleft. Image reproduced with permission from Conti [254].

the McCarthy lab plasmid collection. The luciferase gene was located between NdeI and HindIII restriction sites in plasmid pJM4 making it compatible with the pRS315-P_{FUS1}-eGFP plasmid (figure 3.6). The luciferase gene was cloned in place of the eGFP reporter gene in the plasmid pRS315-P_{FUS1}-eGFP, creating the pRS315-P_{FUS1}-luciferase plasmid, and the cloning was confirmed by sequencing. The reporter plasmid was then further modified with the ligation of the IRE sequences enabling interaction with the repressor plasmid. The expression of the luciferase reporter gene needs to be reduced to a minimum level during the circuit OFF-state, reducing noise from the circuit and maximizing the switch to the ON-state. To achieve this repression of basal expression levels the repressor plasmid was further modified with the incorporation of the iron response element, described by Koloteva *et al* [4].

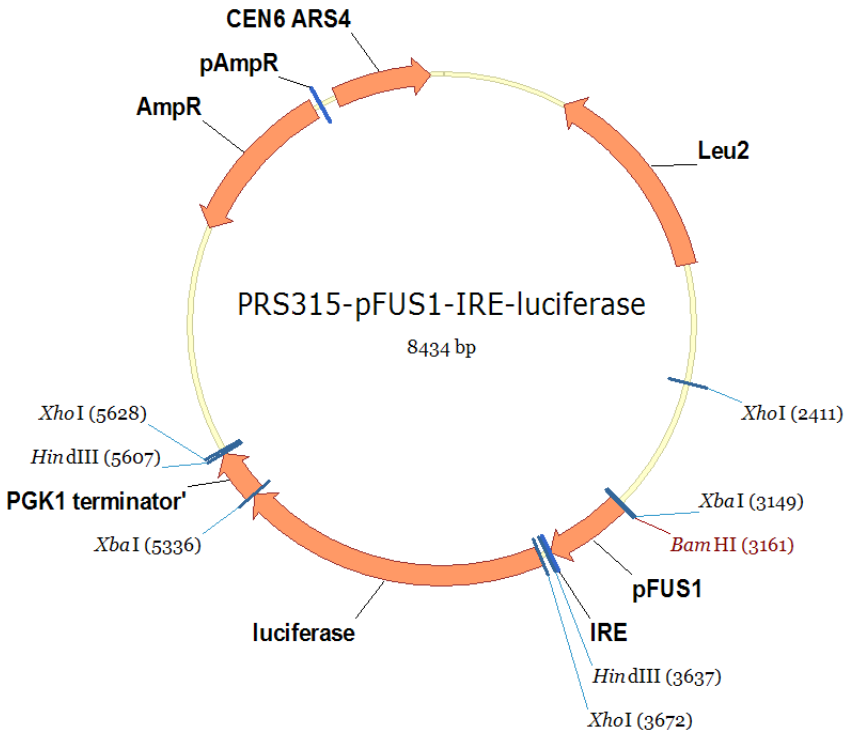

FIGURE 3.6: Map of the pRS315-P_{FUS1}-IRE-luciferase reporter plasmid.

3.4 Insertion of the Iron Response Element

A set of synthetic oligonucleotides was designed containing the IRE nucleotide sequence from Koloteva *et al* between overhanging CA nucleotides (figures 3.7 and 3.8), enabling the IRE to be ligated into the NdeI restriction site on the reporter plasmid (figure 3.6) [4]. The IRE oligonucleotides were synthesized by Eurogentec, annealed together into a double stranded DNA fragment duplex, and ligated into the reporter plasmid. Using this cloning strategy with a single restriction site and a small insert presents a high probability of the vector plasmid re-annealing during ligation resulting in a large number of false positive transformation colonies. However, the IRE contains an additional HindIII restriction site and the elimination of the reporter plasmid's single NdeI restriction site during ligation provided a screening method for identifying transformed colonies using restriction enzyme digest. Following transformation of *E. coli* TOP10, a number of colonies were selected for miniprep DNA amplification and screened by restriction enzyme digest (figure 3.9). The plasmid was digested with restriction enzymes, and agarose gel electrophoresis was used to determine the sizes of the bands following enzymatic digestion, and to confirm the IRE had been correctly

```
5′-CACCAATTATCTACTTAAGCTTCAACAGTGCTTGAACTTAAGAACACAAAACTCGAGAAGA-3′
   3′-GGTTAATAGATGAATTCGAAGTTGTCACGAACTTGAATTCTTGTGTTTTGAGCTCTTCTAC-5′
```

FIGURE 3.7: Nucleotide sequence of the iron response element (IRE), constructed as a synthetic oligonucleotide by Eurogentec for annealing into double stranded DNA, and ligating into the reporter plasmid. The CA overhangs (marked in bold) enable the insertion of the IRE into an NdeI restriction site. The underlined region forms the stem-loop hairpin structure that is bound by the IRP, blocking translation.

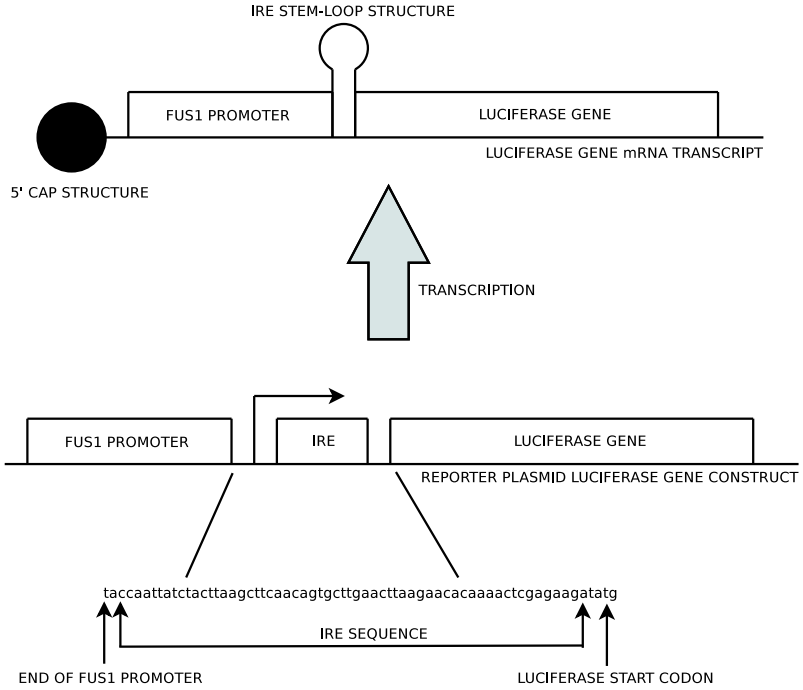

FIGURE 3.8: Schematic diagram of the IRE position, in relation to the FUS1 promoter and IRE gene.

ligated (figure 3.9). The plasmids were digested with HindIII, and BamHI. The IRE ligated reporter plasmid produces 6kb, 2kb fragments and a 500bp fragment due to an additional HindIII site located in the IRE. The pRS315-P_{FUS1}-luciferase without the IRE was used as a control and produces 6kb 2.6kb fragments.

The IRE ligated plasmid was sent for sequencing to confirm the IRE had inserted in the correct orientation, and transformed into *S. cerevisiae sst2*Δ strain.

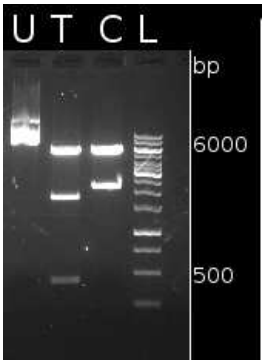

FIGURE 3.9: Insertion of the IRE into the plasmid was confirmed by digestion with restriction enzymes. L is NEB 1kb ladder. U is undigested pRS315-P_{FUS1}-IRE-luciferase plasmid. T is the pRS315-pFUS1-IRE-luciferase plasmid digested with BamHI and HindIII restriction enzymes. C is the control pRS315-P_{FUS1}-luciferase enzyme digested with BamHI and HindIII. The additional HindIII site in the IRE produces 6kb, 2kb and 500bp fragments while the single HindIII site in the control cuts 6kb and 2.6kb fragments.

3.5 Construction of the Repressor Plasmid

The repressor plasmid was initially constructed using the pTRPex plasmid, containing the yeast *TRP1* constitutive promoter (see figure 3.11), and has been used as a low strength constitutive promoter in previous work in the lab (Firczuk, M (2009), personal communication. University of Warwick). In wild type yeast P_{TRP1} controls expression of the *TRP1* gene of *S. cerevisiae* which codes for N(5′phosphoribosyl)-anthranilate isomerise which catalyses the third step in the tryptophan biosynthetic pathway [255]. The P_{TRP1} promoter has been found to generate two mRNA groups with different 5′ ends named “transcript I” and “transcript II” [255]. One group of mRNA transcripts is transcribed with leader sequences which are 60 to 200bp longer than the other [256]. Each group has been found to be transcribed from a region of the promoter with a sequence homologous to the consensus sequence of the TATA box [255]. The P_{TRP1} promoter region in the pTRPex plasmid was identified between positions 571 and 2185bp. The second TATA box was identified between 1864 and 2190bp, and the TATA box identified at 1929bp. It was hypothesized the *TRP1* promoter would provide an appropriate level of IRP expression and efficient inhibition of basal expression of the reporter plasmid, and the published studies of the *TRP1* promoter would provide sufficient understanding to design effective control elements to control IRP expression. Site directed mutagenesis was used to insert a NotI restriction site into

the *TRP1* promoter at 1864bp to enable the insertion of LexA operator sequences into the promoter that would prevent the second TATA box from promoting the transcription of downstream genes (see section 3.4.2).

3.5.1 Cloning the Iron Response Protein Gene

3.5.1.1 *TRP1* promoter strategy

The iron response protein gene was obtained from the laboratory plasmid library pJM6 plasmid (figure 3.10). The pJM6 plasmid and pTRPex plasmids are both based on the same pSupex plasmid, published by Oliveria *et al* [214], and the *TRP1* promoter and IRP gene have been cloned between compatible restriction enzymes sites. A problem was encountered whilst attempting to clone the IRP gene from pJM6. The IRP gene repeatedly failed to ligate into the pTRPex plasmid with none of the transformed colonies producing fragments with XhoI and XbaI restriction enzyme digests. The IRP was also amplified from the plasmid pJM6 using PCR and primers designed to introduce SalI and XhoI restriction sites at the $5'$ and $3'$ ends of the IRP to enable cloning into the pTRPex plasmid, however this strategy also failed to ligate successfully. A large number of cloning attempts were made by the author, along with assistance from senior post doctoral researchers however a successful ligation of the IRP gene into the pTRPex plasmid could not be achieved. It is believed that the problems encountered cloning the PCR amplified IRP gene may have been due to a technical problem with the SalI enzyme not cutting PCR products. New England Biolabs have acknowledged this problem and are investigating at the time of writing. The pTRPex plasmid is a large 8kb plasmid, and the P_{TRP1} promoter is a large promoter, 3kb in length. It was possible that a secondary structure motif in the IRP DNA fragment was interacting with the plasmid during the ligation reaction, or attempting to construct a plasmid of 11kb was also resulting in a problematic ligation. The shorter *DCD1* promoter was available in the McCarthy lab and provides the same promoter strength as *pTRP1* (Firczuk M. (2010). Personal communication (McCarthy Lab), [212, 242]). The *DCD1* promoter was also shorter than *pTRP1* and does not produce a heterogeneous mRNA population or consist of

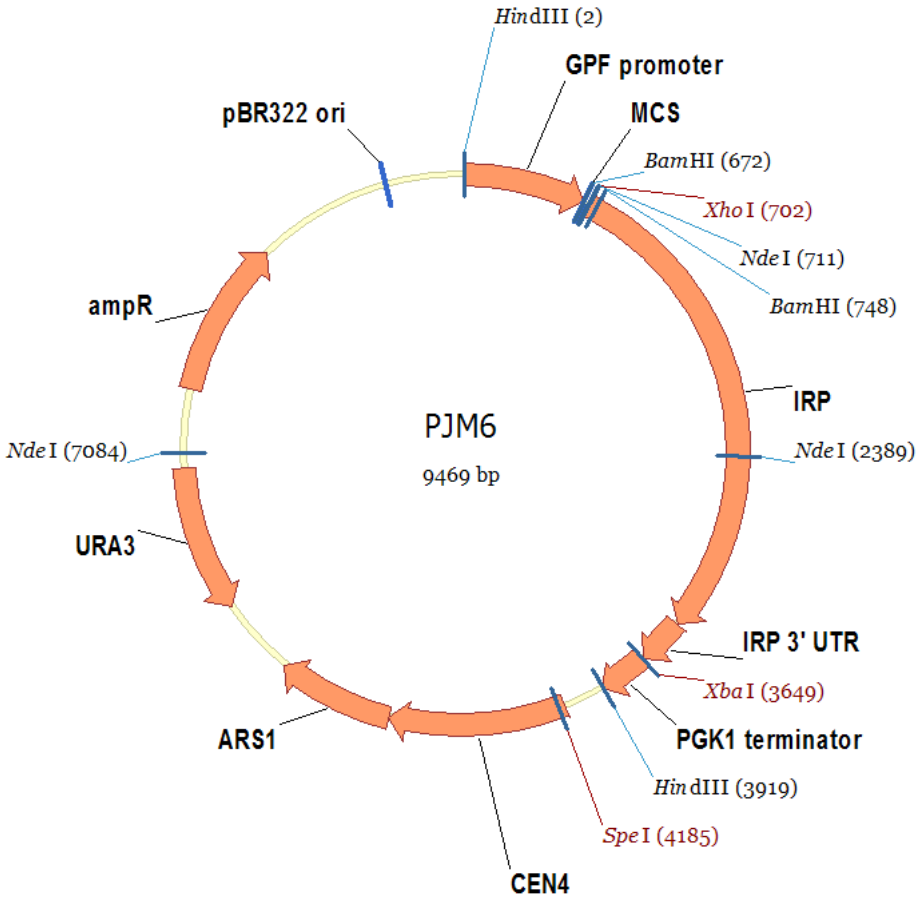

FIGURE 3.10: Map of the pJM6 hIRP plasmid used for cloning the IRP gene into the repressor plasmids (from the McCarthy lab plasmid library). pJM6 is a yeast expression plasmid based on the pSupex plasmid, and contains the *URA3* auxotrophic marker for selection in yeast and the ampicillin resistance marker for selection in bacteria. pJM6 contains the IRP gene under the control of the galactose inducible promoter, P_{GAL}.

multiple transcription start sites, as was reported for the *TRP1* promoter. It was decided therefore, to exchange P_{TRP1} for P_{DCD1} in the circuit.

3.5.1.2 *DCD1* promoter strategy

The *DCD1* promoter is a low strength promoter, and has been studied previously in the McCarthy lab (Firczuk, M. (2010). Personal communication. University of Warwick). P_{DCD1} is the promoter for the gene Deoxycytidine monophosphate (dCMP) deaminase required for dCTP and dTTP synthesis and is expressed constitutively in *S. cerevisiae* [212, 242]. The *DCD1* promoter is 209bp in length and has not been shown in the

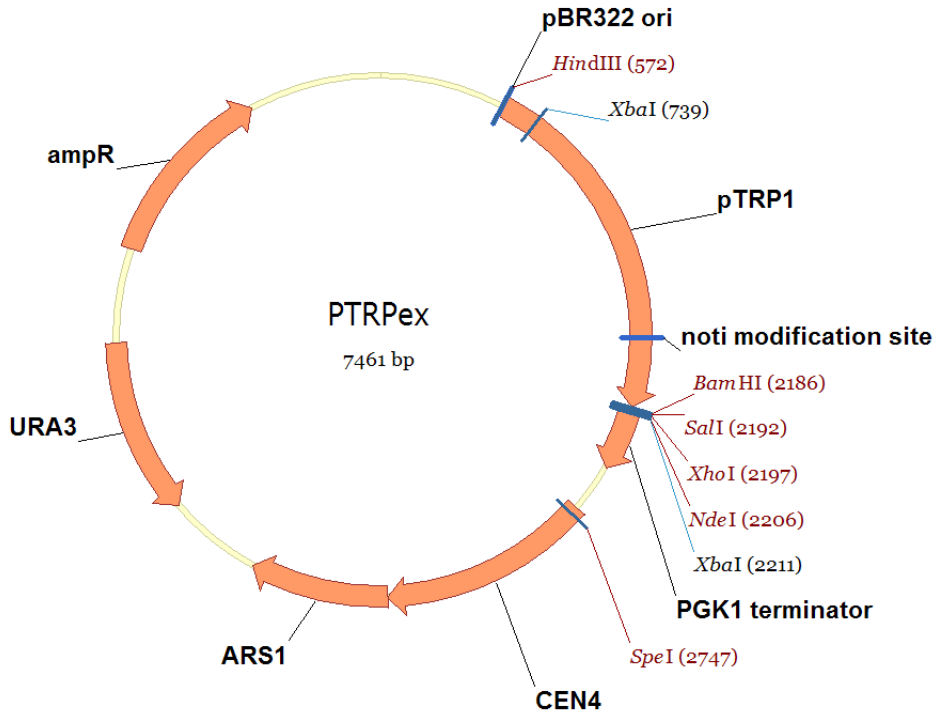

FIGURE 3.11: Map of the pTRPex plasmid from the McCarthy lab plasmid library. The plasmid pTRex is a yeast expression plasmid based on the pSupex plasmid [214]. The plasmid contains the *TRP1* yeast constitutive promoter, modified with a NotI restriction site to enable insertion of additional control elements within the promoter. The plasmid is based on the pSupex plasmid, and contains the *URA3* auxotrophic marker for selection in yeast and ampicillin resistance marker for selection in bacteria. pTRPex contains the constitutive medium strength yeast promoter, P_{TRP1} and a downstream MCS for the insertion of additional genes.

literature to produce heterogeneous mRNA populations from multiple transcription start sites, as observed from P_{TRP1} [242, 255]. P_{DCD1} therefore provides a more compact promoter than P_{TRP1}, and can be more easily modified with up-stream control elements, rather than internal control elements as in the previous design with P_{TRP1}. The *DCD1* promoter was obtained from the pDCDex plasmid, constructed by Maja Firczuk in the McCarthy lab (figure 3.12), and based on pSupex published by Oliviera *et al* [214]. Plasmid P_{DCD1} and pJM6 were digested using the SalI and SpeI restriction enzymes and the IRP ligated into the pDCDex plasmid, downstream of the *DCD1* promoter. Colony PCR was performed using the P_{DCD1} forward primer and the internal IRP primer#2. The P_{DCD1} primer binds to the 5′ end of the promoter, and the IRP primer #2 binds to the complementary strand of the IRP gene 700bp downstream from the 5′ end of the IRP. A PCR product therefore can only be produced by both the *DCD1* promoter and

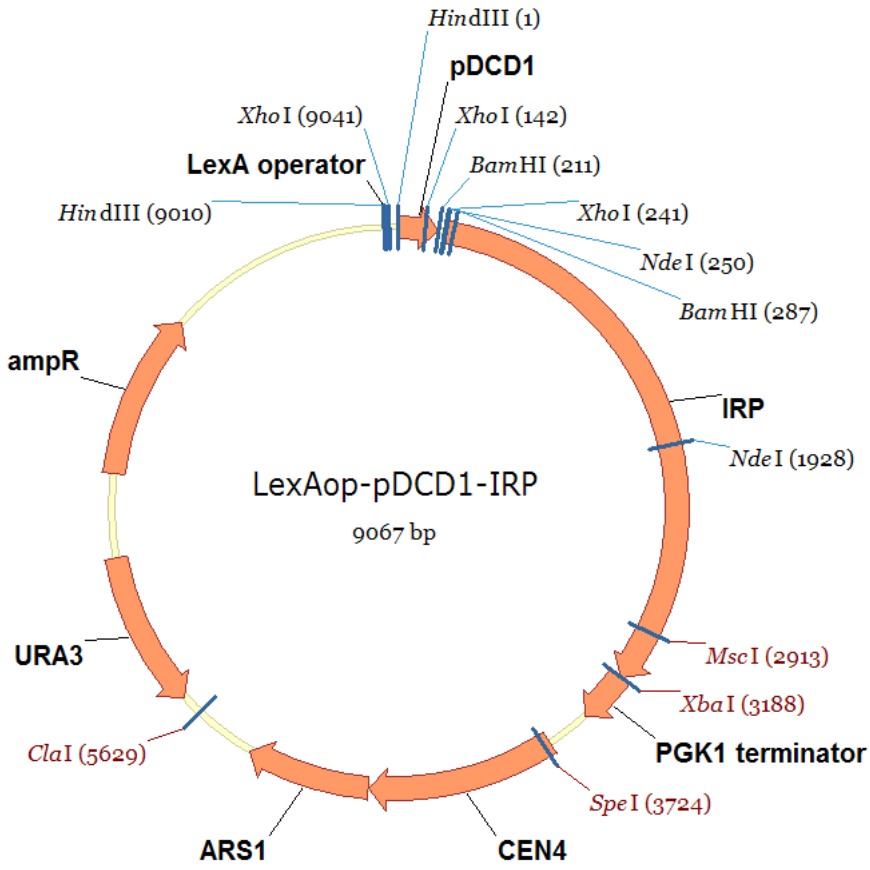

FIGURE 3.12: LexAop-pDCD1-IRP repressor plasmid map.

the IRP gene being present in the correct orientation. The colony PCR results provided a high ratio of colonies with PCR product of the correct size (figure 3.13). Plasmids capable of generating correctly-sized PCR products were sequenced to confirm the correct sequence and orientation of the IRP gene in the pDCDex plasmid.

3.5.1.3 *TEF1* promoter strategy

The *TEF1* promoter is a strong constitutive yeast promoter, promoting mRNA expression levels higher than P_{TRP1} and P_{DCD1} promoters [244]. P_{TEF1} is a 590bp promoter that is also shorter than P_{TRP1} and can potentially be modified with upstream control elements using the same strategy as P_{DCD1}. The IRP gene was cloned from the plasmid pJM6 into the vector plasmid PTRPex (also based on the pSupex plasmid) using the SalI and SpeI restriction enzymes as a double digested. Following the ligation of the IRP gene into the pTEFex vector plasmid, transformed colonies were screened using colony

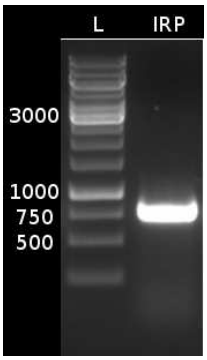

FIGURE 3.13: Agarose gel electrophoresis of colony PCR products from bacterial colonies transformed after ligation of the IRP gene with the pDCDex plasmid. The left lane contains the New England Biolabs 1kb ladder for size verification. The right hand lane contains the colony pcr product from the P_{DCD1}-IRP plasmid construct. The colony PCR experiment produced a DNA fragment of approximately 800bp which is the expected size produced by the polymerase reaction using the P_{DCD1} and IRP#2 combination of primers.

PCR with the internal IRP primer #2, which binds to the complementary strand of the IRP gene 700bp upstream of the 5′ end of the gene, and also the P_{TEF1} primer binding to the 5′ end of the *TEF1* promoter. Only plasmids containing the P_{TEF1}-IRP construct in the correct orientation can produce a fragment of approximately 1.2kb in size. Bacterial colonies that were harbouring plasmids that were capable of generating correctly-sized PCR products were used for plasmid miniprep purification and the plasmids sent for sequencing to confirm the correct sequence and orientation of the IRP gene.

3.5.2 Insertion of LexA Operator Control Sequences

The *E. coli* LexA protein expressed by the de-repressor plasmid binds to a specific DNA recognition sequence called the “LexA operator” [236, 237]. The LexA operator is comprised of the nucleotide sequence: TCGAGTACTGTATGTACATACAGTAC.

The LexA protein has been expressed in yeast in previous work published by Brent and Ptashne and has been shown to inhibit expression of galactose inducible LacZ genes when placed between the upstream activator sequence and transcription start point [203]. It has been suggested that multiple repeats of the LexA operator are required to block downstream transcription [203, 235], however there is no agreed consensus in the literature as to how many repeats are optimum for repressing gene expression [235].

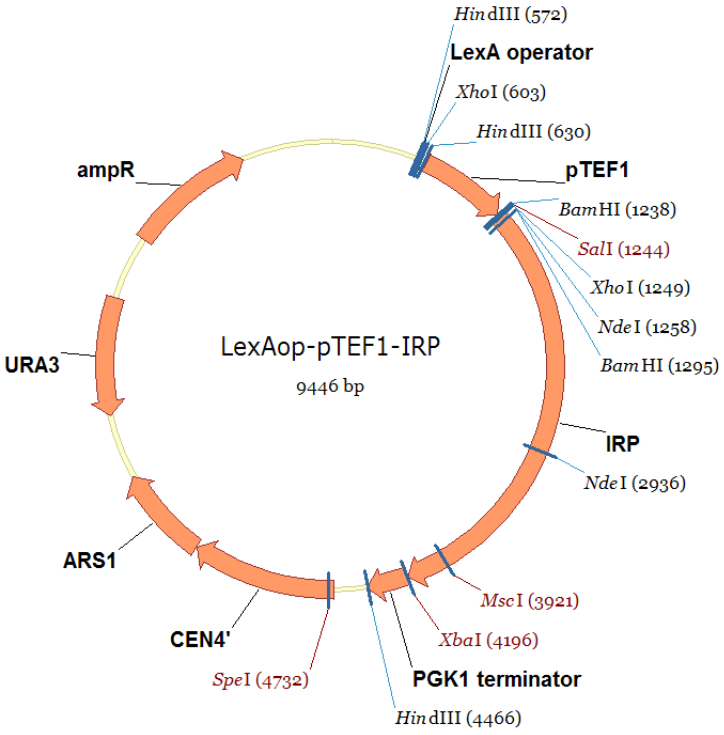

FIGURE 3.14: Map of the LexAop-pTEF1-IRP repressor plasmid. This version of the repressor plasmid contained the IRP gene expressed by the *TEF1* strong constitutive promoter [244]. The plasmid vector was based on pSupex, published by Oliveria *et al* [214].

Also, research published by Brent indicates positioning the LexA operators within 60bp upstream of the start codon is sufficient to inhibit transcription by up to 10 fold [203].

Therefore, a set of two tandem repeats of the LexA operator was designed based on the sequence published by Brent and Ptashne that could be inserted into the HindIII restriction site that is found immediately upstream of the 5′ end of the *DCD1* and *TEF1* promoters [203]. In addition, a unique PacI restriction site was also designed into the 3′ end of the operator sequences to enable screening by restriction enzyme digest. The LexA operator sequence was ordered from Eurogentec as a custom set of complementary oligonucleotides. The oligonucleotide strands were annealed together into a double stranded DNA molecule and ligated into the repressor plasmid, linearised by HindIII restriction enzyme digest. A number of plasmids were purified by miniprep from transformed bacterial colonies, and screened by restriction enzyme digest using the PacI enzyme. The digests were visualized by agarose gel electrophoresis and plasmids that were linearised by the PacI enzyme were sent for sequencing to confirm

```
5′-AGCTTCGAGTACTGTATGTACATACAGTACTCGAGTACTGTATGTACATACAGTACTTAATTAAA-3′
    3′-AGCTCATGACATACATGTATGTCATGAGCTCATGACATACATGTATGTCATGAATTAATTTTCGA-5′
```

FIGURE 3.15: The LexA operator sequence containing flanking HindIII restriction sites (bold) and internal PacI restriction site (underlined). The LexA operator sequences were ordered as a set of oligonucleotides, annealed, and ligated into the HindIII restriction site in the P_{DCD1}-IRP and P_{TEF1}-IRP plasmids enabling interaction of the repressor plasmid with the LexA protein expressed from the de-repressor plasmid during pheromone induction of the gene circuit.

the insertion and location of the LexA operators.

3.5.3 Cloning the IRP PEST Degradation Tag

The plasmid pSVA17 containing the gene yEGFP3-$CLN2_{PEST}$ was obtained from Simon Avery (University of Nottingham) who (with Carolina Mateus) had constructed a short half-life GFP by the addition of the degradation signal from the 3-terminal 534 nucleotides of *CLN2* to the 3′ terminus of yEGFP3 [216]. In order to obtain the PEST tag in a form suitable for tagging the IRP gene, the primers used by Mateus and Avery were modified to exchange the restriction sites in yEGP3 for the MscI restriction sites compatible with the IRP gene, and an additional unique NruI site for screening potential clones after ligation (figures 3.18 and 3.19) [216]. The *CLN2* degron tag was inserted, in-frame at the 3′ terminus of the IRP upstream of the stop codon, creating the degron sequence tagged IRP_{PEST} gene (see appendix A).

The 534bp PEST tag was amplified using PCR with the PEST Fwd and PEST Rev primers containing the MscI restriction sites for insertion into the IRP gene on the repressor plasmids (figure 3.16), and ligated into LexAop-P_{DCD1}-IRP (figure 3.18) and LexAop-pTEF1-IRP (figure 3.19), linearised by restriction enzyme digest with the enzyme MscI. The ligation was transformed into *E. coli* TOP10 and the transformed colonies screened by colony PCR using the PEST Fwd and PGK1 terminator primers. The colony PCR product was expected to be approximately 750bp from the 534bp PEST sequence and 273bp PGK1 terminator sequence (figure 3.17). Plasmid preparations were made from bacterial colonies harbouring plasmids that produced colony PCR

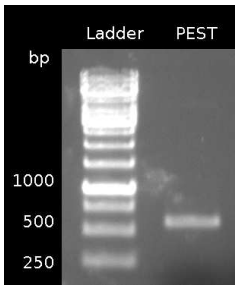

FIGURE 3.16: PCR amplification of the PEST degron tag from plasmid pSVA17 from Mateus and Avery [216].

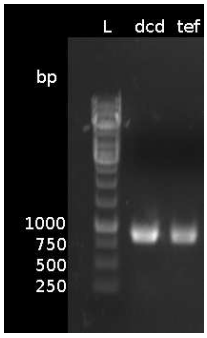

FIGURE 3.17: Colony PCR agarose gel electrophoresis performed on transformed bacterial colonies containing the *TEF1* and *DCD1* promoter repressor plasmids following ligation of the plasmids with the PEST degradation tag. L is the (Fermentas) 1kb ladder for size determination, dcd is the repressor plasmid with the *DCD1* promoter, and tef is the repressor plasmid with the *TEF1* promoter. Colony PCR was performed with the PEST Fwd and PGK1 terminator Rev primer.

fragments of the correct size, and the plasmids were sent for sequencing to confirm the correct sequence and orientation of the PEST tag.

3.6 Construction of the De-Repressor Plasmid

The de-repressor plasmid was initially constructed from the pRS315 plasmid. The LexA protein from *E. coli* was cloned downstream of the P_{FUS1} promoter. Expression of the LexA protein is therefore up-regulated by the yeast mating response. The LexA protein gene was amplified from *E. coli* TOP10 genomic DNA by PCR using primers containing 5′ NdeI and 3′ HindIII restriction enzyme sites for compatibility with the pRS315 plasmid (see table 2.2) (figure 3.21). The amplified LexA gene was cloned into pRS315-P_{FUS1}-eGFP in place of eGFP, forming pRS315-P_{FUS1}-LexA (figure 3.20). Using the pRS315 plasmid for both the de-repressor and reporter plasmid creates a

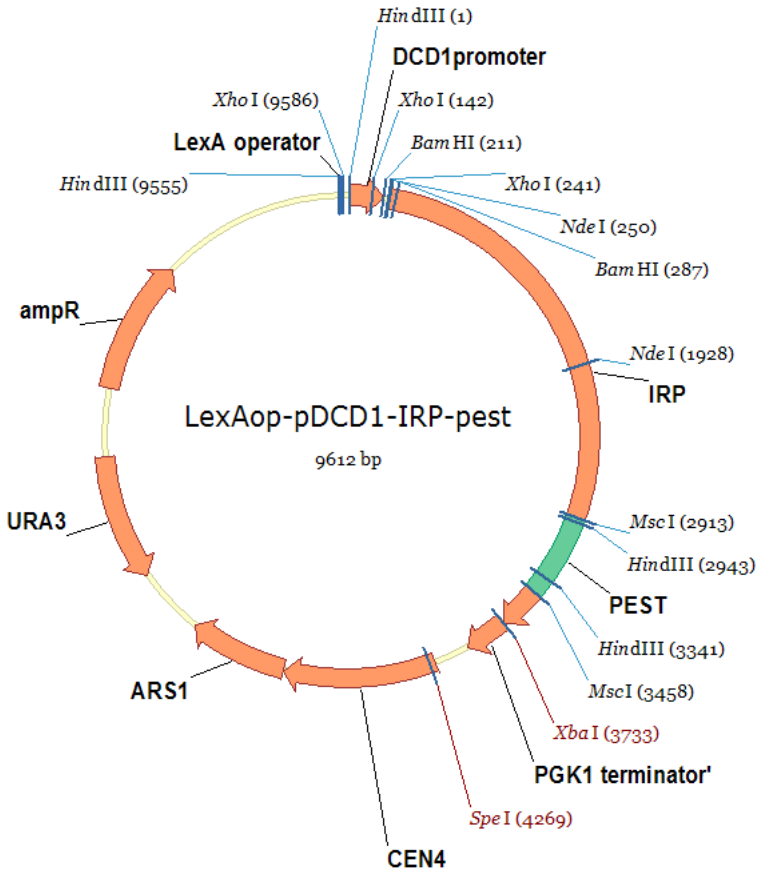

FIGURE 3.18: Map of the repressor plasmid containing the *DCD1* promoter and IRP, modified with the PEST degradation tag.

selection issue as both plasmids share the same LEU2 auxotrophic marker and therefore cannot be co-expressed. The prs313 plasmid backbone, a pBluescript-based centromere vector (NCBI accession number U03439), created by Sikorski and Hieter [213] contains the His3 auxotrophic marker, however the His3 gene did not contain any compatible restriction sites to clone into the pRS315 plasmid. The P_{FUS1}-LexA elements from the pRS315 plasmid were cloned as a cassette into the pRS313 plasmid using the BamHI and SalI restriction enzyme sites, to utilize the His3 auxotrophic marker. Following ligation, transformed colonies were grown overnight in LB with amicillin and plasmid DNA purified by miniprep. The plasmids were digested with SalI and BamHI restriction enzymes to confirm the ligation of P_{FUS1}-LexA into the pRS313 vector plasmid (see figure 3.22 A and B). The plasmid was digested alongside empty pRS313 plasmid (figure 3.22 B). Digestion with SalI and BamHI produced fragment of approximately 1.2kb from the pRS313-P_{FUS1}-LexA plasmid and no fragment from the empty pRS313

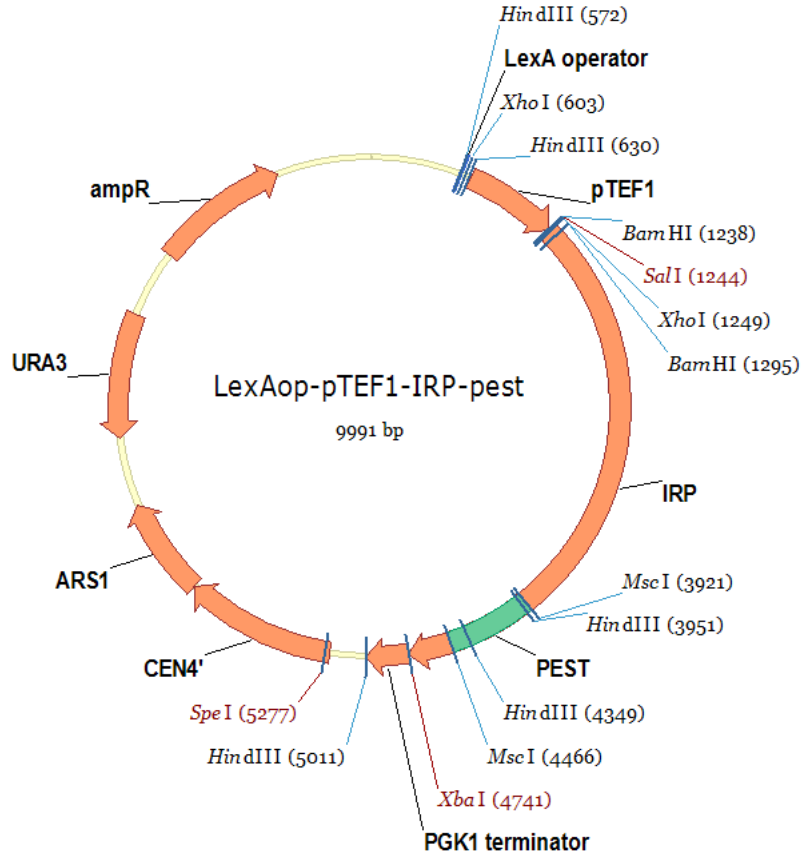

FIGURE 3.19: Map of the repressor plasmid containing the *TEF1* promoter and IRP, modified with the PEST degradation tag.

control plasmid. Plasmids producing a 1.2kb fragment following restriction enzyme digestion with BamHI and SalI were sent for sequencing to confirm the sequence and orientation of the P_{FUS1} promoter and LexA gene.

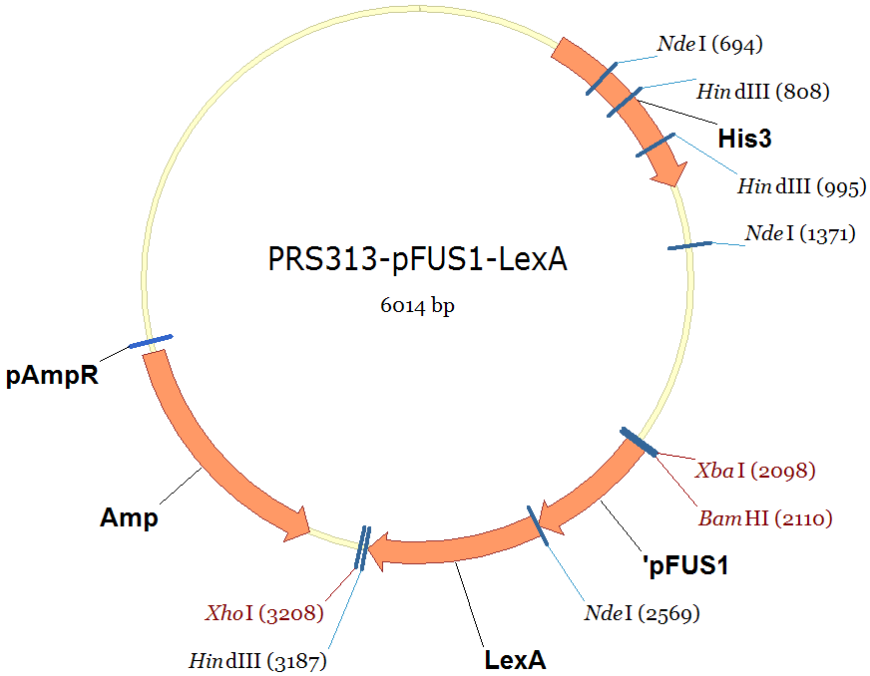

FIGURE 3.20: Map of the pRS313-pFUS1-LexA de-repressor plasmid.

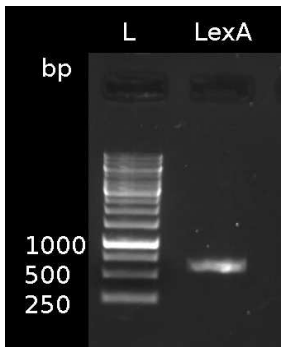

FIGURE 3.21: PCR product of the LexA gene from *E. coli* genomic DNA. L is the (Fermentas) 1kb ladder for size determination, and LexA is the PCR product.

3.7 Conclusion

Following confirmation of the sequencing data for each of the plasmids, the circuit was transformed into *S. cerevisiae sst2*Δ (Euroscarf strain Y06055). *S. cerevisiae sst2*Δ strain is hyper-sensitive to pheromone due to the mutation in the *SST2* gene that inhibits the negative feed-back loop in the yeast pheromone response pathway [49], and would enable the strongest possible response to pheromone induction for luciferase expression.

S. cerevisiae was transformed sequentially, first with the reporter plasmid, growing the cells in YPD with -*LEU* drop out supplement. These cells provided a luciferase control

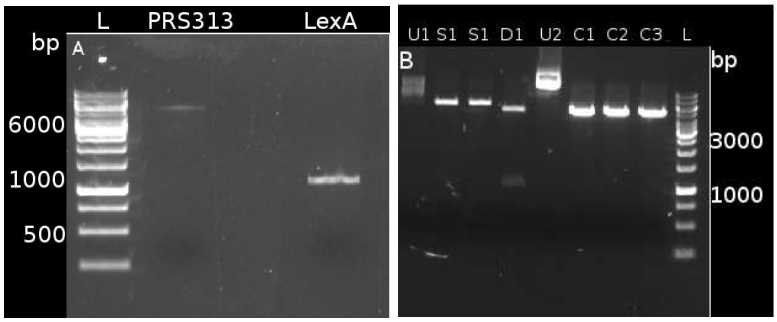

FIGURE 3.22: Ligation of the LexA gene with the pRS313 de-repressor plasmid. A. L is the (Fermentas) 1kb ladder, pRS313 is the pRS313 vector digested with the BamHI and SalI restriction enzymes. LexA is the P_{FUS1}-LexA fragment. B. pRS313 plasmid after ligation with the P_{FUS1}-LexA fragment digested with BamHI and SalI restriction enzymes. S1 is BamHI single digest. S2 is SalI single digest. D1 is BamHI and SalI double digest. C1 is control pRS313 plasmid cut with the SalI enzyme. C2 is control pRS313 plasmid cut with BamHI. C3 is control pRS313 plasmid cut with SalI and BamHI restriction enzymes. U2 undigested pRS313 plasmid DNA.

strain that could be used for comparison with the circuit luciferase expression. The yeast were next transformed with the de-repressor plasmid, and grown on YPD medium with *-LEU, -HIS* drop out supplement. The cells containing both the reporter and de-repressor plasmids were finally transformed with one of the four repressor plasmids, and grown on YP medium with *-LEU, -HIS, -URA* drop out supplement. The final round of transformations produced four different strains of cells, each containing one of the four circuits, named after the repressor plasmid they contain (table 3.1). In addition to the four circuit strains, and addition strain was transformed with only the reporter plasmid, to create a "luciferase control" strain, to measure maximum luciferase expression from the *sst2*Δ in the absence of LexA and IRP interactions.

Circuit Variation	IRP Respressor Variation
P_{DCD1}	P_{DCD1} weak constitutive expression of the IRP
P_{TEF1}	P_{TEF1} strong constitutive expression of the IRP
P_{DCD1}-PEST	P_{DCD1} weak constitutive expression of short half-life IRP_{PEST}
P_{TEF1}-PEST	P_{TEF1} strong constitutive expression of short half-life IRP_{PEST}
Luciferase control	Reporter plasmid only. No repressor or de-repressor plasmids

TABLE 3.1: Table of the four circuit variants (strains) that had been constructed for tuning the expression of the luciferase reporter gene. All cells contain the pheromone inducible luciferase gene plasmid, and the pheromone inducible LexA gene plasmid. These plasmids are then combined with the third circuit plasmid, which constitutively expresses the wild type IRP or short half-life IRP_{PEST} repressor, with either the low expression *DCD1* or the high expression *TEF1* promoter.

The set of four circuit designs enables the tuning of the circuit for short and long half-life as well as the overall intra-cellular levels of the IRP repressor. Model simulations indicate the abundance of IRP is a key control point in the circuit therefore the modifications prevent the repressor from overwhelming the LexA de-repressor and preventing luciferase expression (see chapter 5). The choice of circuit combination also provides the flexibility to investigate tuning the dynamic behaviour of the luciferase output through the interaction of IRP and LexA. Strong IRP expression can be coupled with a short half-life, and vice versa to investigate the effect on luciferase signal amplitude and period.

CHAPTER

FOUR

RESULTS - CIRCUIT CHARACTERIZATION

4.1 Introduction

Following construction of the gene circuit plasmids, the plasmids were transformed into *S. cerevisiae sst2*Δ forming four separate strains of circuit, P_{DCD1}, P_{TEF1}, P_{DCD1}-PEST, and P_{TEF1}-PEST representing the *DCD1* weak promoter, and the *TEF1* strong promoter expressing the IRP repressor, and the short half-life IRP_{PEST} variant. The cells were grown individually in broth culture and the expression of the circuit components measured following activation of the circuit by stimulation of the pheromone response pathway with 100nM α-factor pheromone. The expression of the components was measured using a range of techniques to obtain quantitative data at the various hierarchical expression levels within the cell. Luminescence was measured as the circuit "output" using the luminometer and the *in vivo* luminescence assay to obtain real-time kinetic data of luciferase expression. Luminescence measurements provided general circuit performance data that could be used to determine the behaviour of the circuit and compare with model predictions (detailed in chapter 5).

Protein levels were measured to determine the level of the LexA and IRP repressors, and infer kinetic parameters from the effect on the change in the luciferase output of

the circuit. The mRNA expression of the components was also measured using RT-qPCR to quantify the change in transcription levels. The IRP is repressed at the level of transcription by LexA, therefore reduction in IRP mRNA levels can be attributed to the activity of LexA, and the kinetics of the LexA repressor investigated through both the protein levels of the repressor, and also the reduction in IRP mRNA transcription.

4.2 Growth Rate Investigation

In addition to quantifying the relative levels of the components of the circuit, the host yeast strain was checked for the effect of maintaining the circuit on the growth rate of the cells, as changes in the output of the circuit could be due to the burden of maintaining three plasmids. The P_{TEF1} circuit was transformed into the pTC5 lab strain of *S. cerevisiae* to test alongside the pheromone hyper-sensitive *sst2*Δ mutant strain that was used for expressing the circuit. The data would then enable the investigation of effect on growth rate of both the circuit and the *sst2*Δ mutation. It was hypothesized that the P_{TEF1} circuit would exert the greatest burden on the cells due to the higher expression level of the IRP protein. If a difference was observed between the cells carrying the circuit plasmids and those without, the remaining plasmid circuits would be investigated further. The growth rate data was recorded (figure 4.1), and the growth rate was calculated from the slope of the exponential growth phase (table 4.1). The growth

strain	Slope	R^2	doubling time (hours)
pTC5	0.45	0.99	2.2
pTC5 + circuit	0.4	0.99	2.5
*sst2*Δ	0.31	0.99	3.2
*sst2*Δ + circuit	0.3	0.99	3.3

TABLE 4.1: Growth rate of yeast strains pTC5 and sst2Δ with and without the gene circuit to investigate the metabolic burden of maintaining the plasmids. The data was calculated using figure 4.1

rate data revealed that the plasmids did not produce a significant lag in the growth rate of the cells hosting the circuit (table 4.1). The doubling time of the pTC5 strain was approximately 2.2 hours, and increased to 2.5 hours following plasmid transformation. For the *sst2*Δ strain the growth rate was approximately 3.2 hours and increased to 3.3

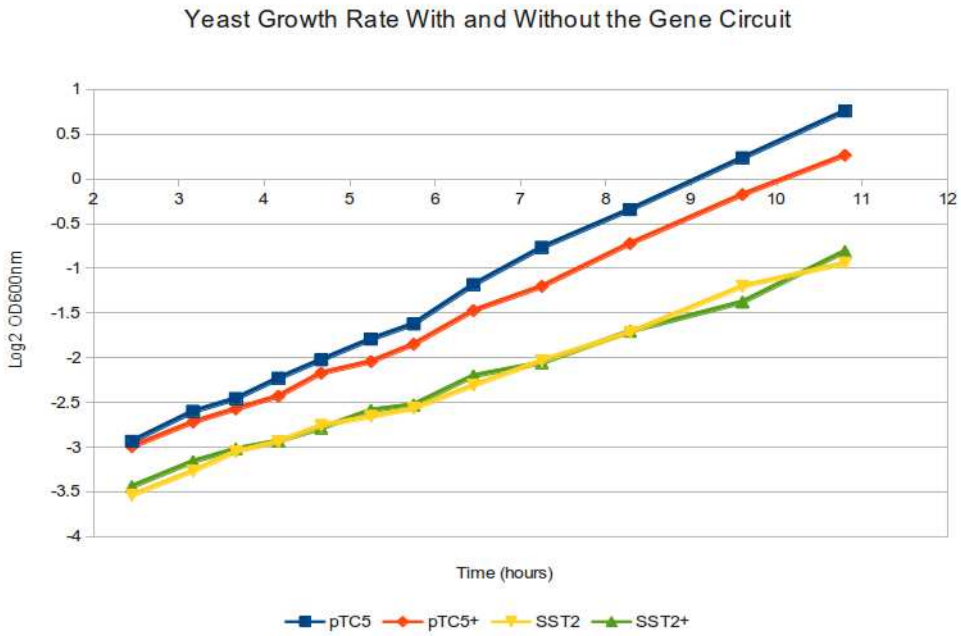

FIGURE 4.1: Log log plot of the growth curve for the *sst2*Δ strain +/- circuit (yellow and green), and the standard pTC5 laboratory strain +/- circuit (blue and orange).

hours following transformation. The *sst2*Δ strain had a slower growth rate than pTC5, but was still sufficient for studying the gene circuit, and enabled hyper-sensitivity to pheromone and potentially a higher-fold increase in luciferase expression [93]. The *sst2*Δ strains transformed with the gene circuit plasmids maintained 97% of the growth rate of the cells without the plasmids, therefore it was concluded that the gene circuit does not introduce a significant burden on the metabolism of the host cell.

4.3 Luminescence Measurement

Luminescence data were collected for each of the four circuit strains, in triplicate, from cultures that were induced with 100nM pheromone and cultures that were un-induced with pheromone (baseline luminescence). In addition, luminescence data was collected from the same yeast strain, containing only the pheromone inducible reporter plasmid as a “luciferase control” strain. This strain represents the maximum expression of luciferase that could be produced by the *sst2*Δ yeast cells, for comparison with the circuit expression levels.

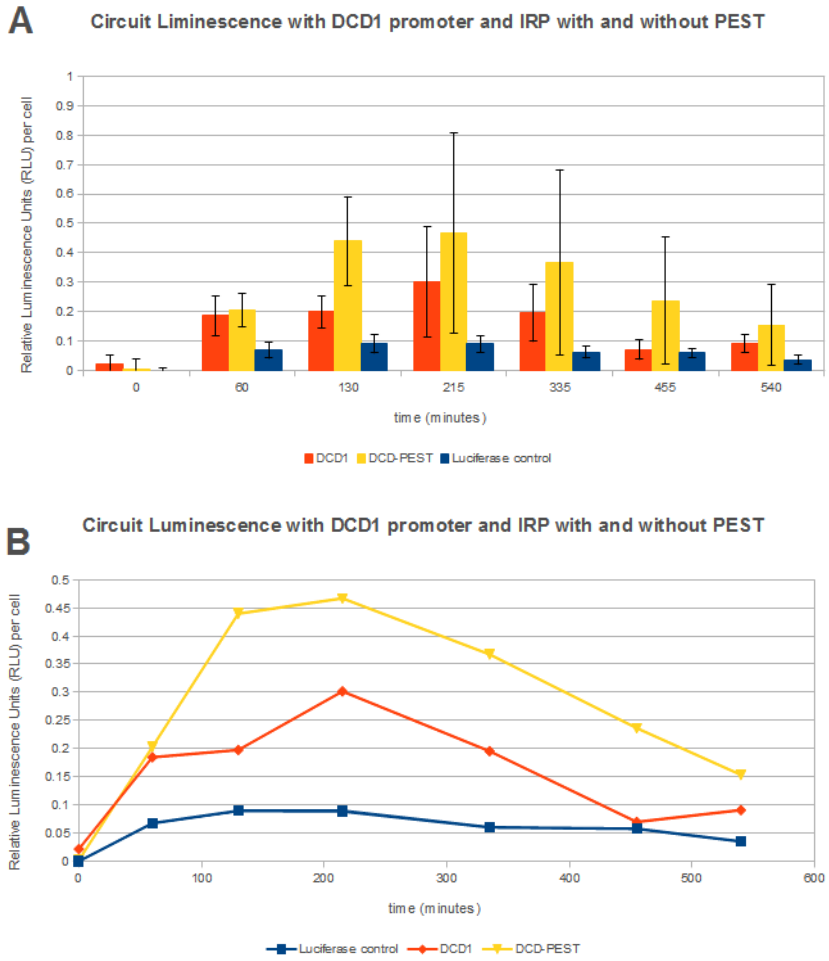

FIGURE 4.2: Luminescence expression for the P_{DCD1} and P_{DCD1}-PEST circuits (n=9). A. Time course measurements of luciferase expression for induced cells. The error bars represent the standard deviation from the mean calculated for each time point measurement. B. Scatter plot of the same data used in plot A, demonstrating the dynamic range of the luciferase expression of the circuits.

Due to the high sensitivity of the luminometer, three biological replicates were performed, and three replicate samples collected at each time point. Each sample was measured three times in the luminometer, providing a final n number of nine data points for each time point. In addition, at each time-point, samples were collected in triplicate for western blot and RT-qPCR analysis from each replicate culture. The data was normalised for cell growth using the cell count for each time point. The data for the luciferase expression of the P_{DCD1} and P_{DCD1}-PEST circuits demonstrated a higher level of luciferase expression from the P_{DCD1} circuit, and higher again from the P_{DCD1}-PEST circuit, compared with the luciferase control (figure 4.2). The baseline expression of un-induced cells was investigated to check the functionality of the repressors (figure 4.3).

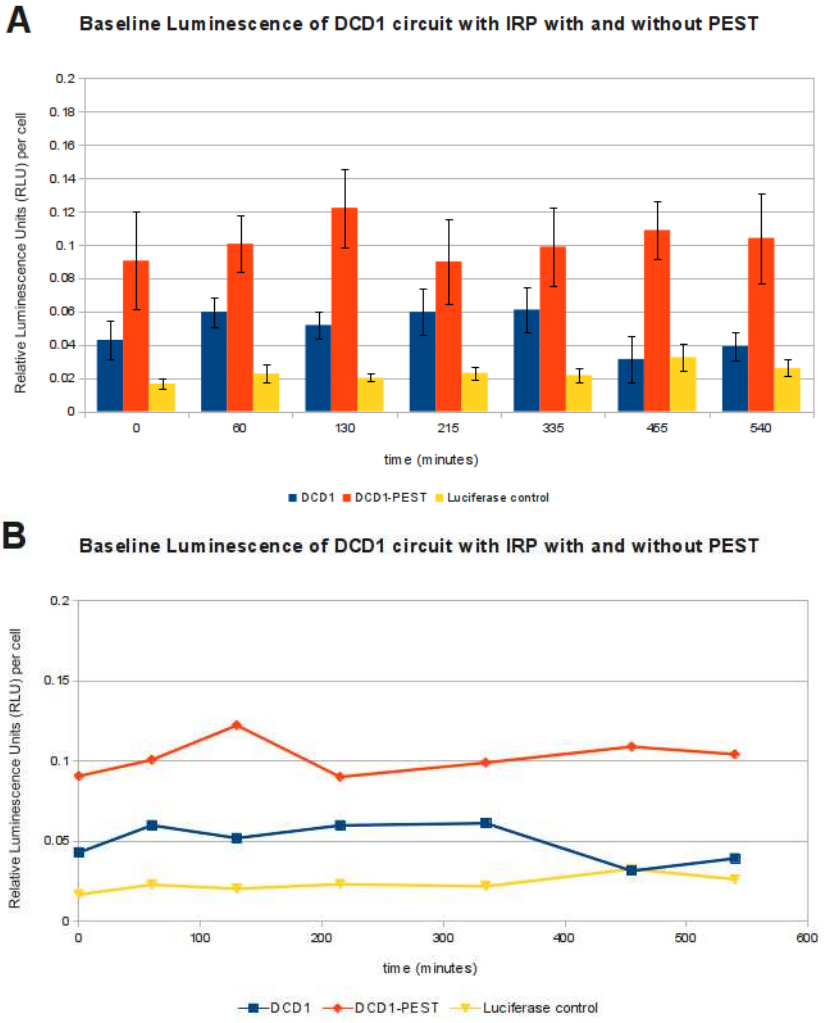

FIGURE 4.3: Baseline luciferase expression for the P_{DCD1} and P_{DCD1}-PEST circuits (n=9). A. Time course measurements of baseline expression for un-induced cells. The error bars represent the standard deviation from the mean calculated for each time point measurement. B. Scatter plot of the same data used in plot A, demonstrating the dynamic range of the luciferase expression of the circuits.

The data show that the P_{DCD1} and P_{DCD1}-PEST circuits do not repress the basal level of expression of luciferase below the basal level of the luciferase control. The expression of the P_{DCD1} and P_{DCD1}-PEST circuits is higher than the control, particularly for the P_{DCD1}-PEST with the short half-life IRP_{PEST}. The data were collected on different days however, introducing variation between the cell cultures. In addition, the baseline luminescence of the control provides an indication of the minimal level of luciferase expression of the reporter gene on the reporter plasmid, however the data from the control cannot be directly related to the circuit strains as in indication of the baseline luminescence from the circuit.

The relative fold change was calculated in order to ascertain a fold increase in luciferase expression for each of the circuits. The luminescence measurement was converted to relative luminescence units per ml using the cell count from the cellometer, and the RLU per cell for pheromone stimulated cells divided by the RLU per cell for the non-pheromone stimulated cells, at each time point. Using this data, it was possible to compare the fold increase in luciferase expression between pheromone stimulated and unstimulated cells. A higher fold change indicates a higher expression in relation to the baseline expression level.

As each strain has the same theoretical maximum expression level of luciferase from the pheromone inducible FUS1 promoter, a fold change difference between strains can be equated to a larger difference between the induced level and uninduced basal level of expression of luciferase. Analysis of the fold change data showed the P_{DCD1} circuit indicated a directional increase in luciferase expression compared with the luciferase control strain4.4). However, the higher level of variation in the data prevents an exact calculation of the fold change difference between the circuit and the control strains.

Maximum luciferase expression was achieved at approximately 200 minutes, compared with 120 minutes from the control, and the transition to the OFF-state requires the same level of time for both the P_{DCD1} circuits and the control, indicating the IRP repressor did not rapidly inhibit luciferase translation following pheromone-induction.

The P_{DCD1}-PEST circuit demonstrated a similar rate of increase in luminescence as the P_{DCD1} circuit, similar to that of the control. The P_{DCD1}-PEST circuit also required 200 minutes to reach maximum luciferase output, as observed from the P_{DCD1} circuit (figure 4.4). The P_{DCD1}-PEST circuit achieved a maximum fold change increase in luciferase expression of approximately 6 fold. This fold change increase was not as high as the P_{DCD1} circuit which was presumed to be due to the attenuated half-life of the IRP and the lower abundance of the repressor in the cell, enabling a higher expression from un-induced cells. Reduction of signal after 200 minute is also slower in the P_{DCD1}-PEST circuit compared with the P_{DCD1} circuit, also suggesting the shorter half-life IRP reduces the overall level of repression of luciferase translation.

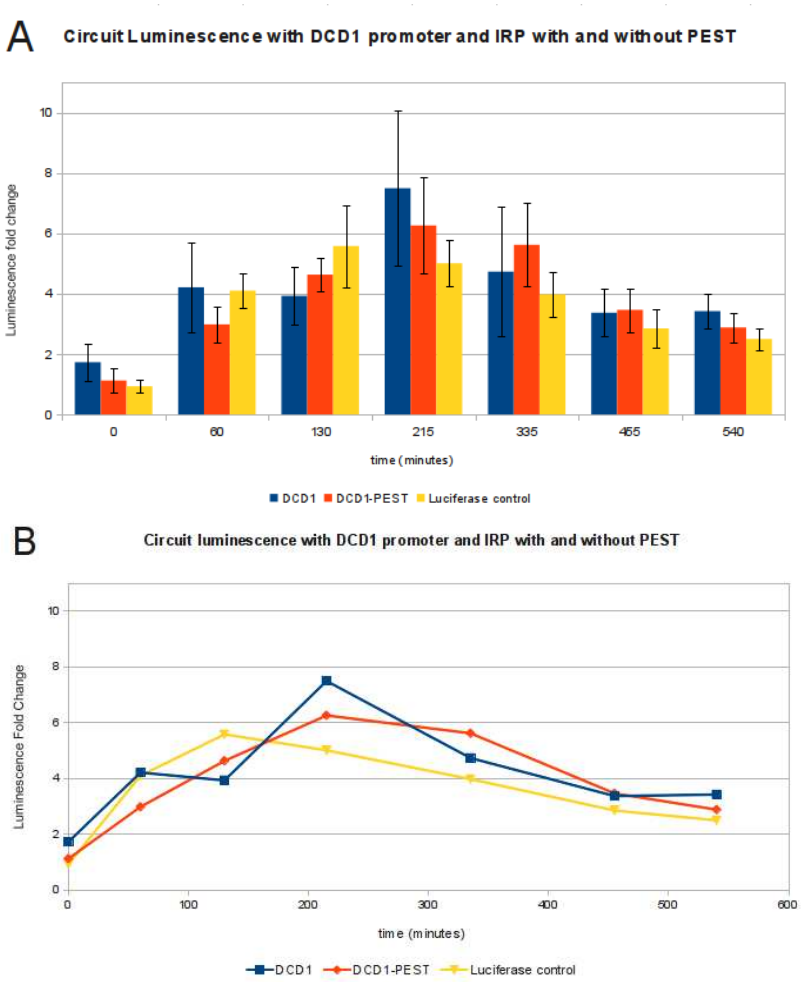

FIGURE 4.4: Fold change in luciferase expression by the P_{DCD1} and P_{DCD1}-PEST circuits compared with the luciferase control (n=9). A. Time course measurements of the relative fold change in luminescence between the induced and un-induced cells expressing the P_{DCD1} and P_{DCD1}-PEST gene circuits. The error bars represent the standard deviation from the mean calculated for each time point measurement. B. Scatter plot of the same data used in plot A, demonstrating the dynamic range of the luciferase expression of the circuits.

The experiment was repeated with the P_{TEF1} and P_{TEF1}-PEST circuits (figures 4.5 and 4.6). As expected, the luminescence expression was reduced from the P_{TEF1}, and P_{TEF1}-PEST circuits, which was attributed to the higher level of IRP expression from the *TEF1* promoter. The baseline data was also plotted and showed a lower level of baseline expression for the P_{TEF1}, when compared with the luciferase control and the P_{DCD1} circuits (figure 4.3). The P_{TEF1}-PEST circuit showed a higher level of expression than the P_{TEF1} circuit, similar to the level of the luciferase control, due to the shorter half-life of the IRP_{PEST} repressor (figure 4.6 B). The fold change in luciferase was plotted for the P_{TEF1} and P_{TEF1}-PEST circuits (figure 4.6). The data showed a reduction in the

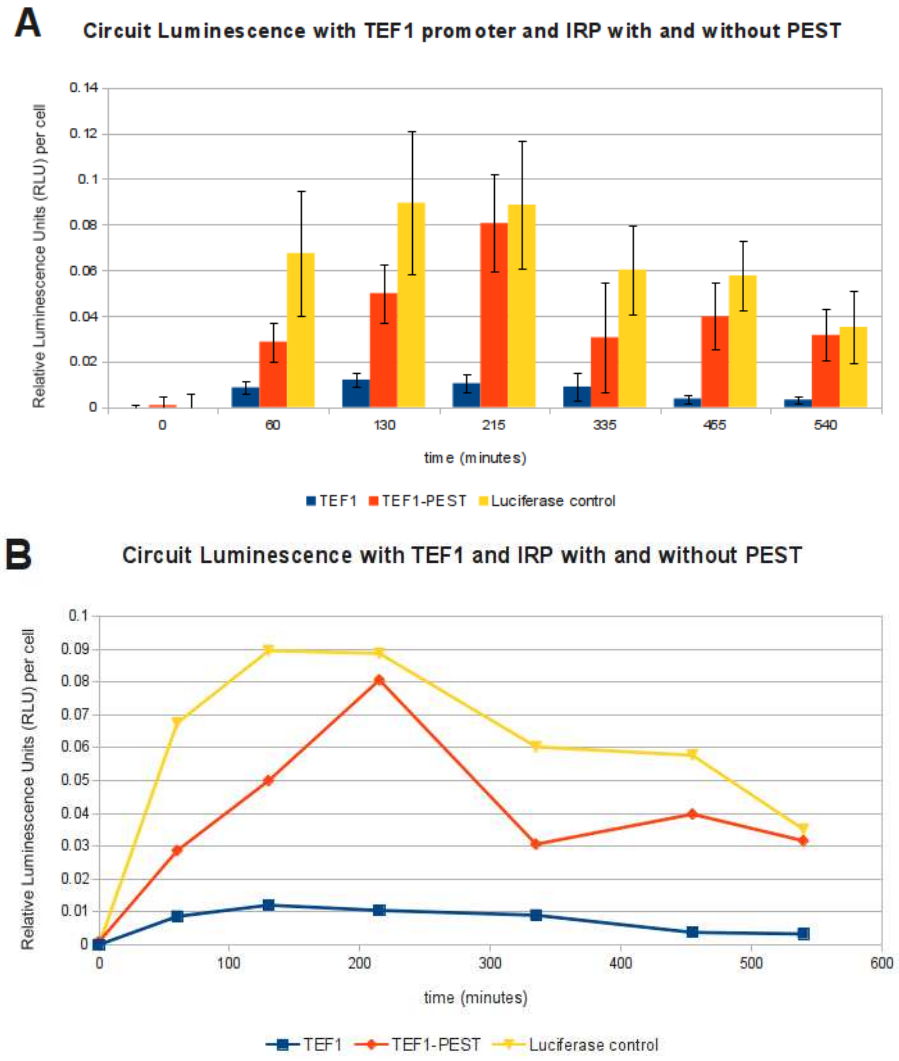

FIGURE 4.5: Luminescence expression for the P_{TEF1} and P_{TEF1}-PEST circuits (n=9). A. Time course measurements of luciferase expression for induced cells. The error bars represent the standard deviation from the mean calculated for each time point measurement. B. Scatter plot of the same data used in plot A, demonstrating the dynamic range of the luciferase expression of the circuits.

output of luciferase compared with the P_{DCD1} and P_{DCD1}-PEST circuits (figure 4.7 and 4.4). The P_{TEF1} promoter expressed a higher level of IRP in the cell, and inhibits the translation of luciferase mRNA expressed during pheromone-induction of the circuit (and was confirmed by western blot, figure 4.17). The low level of luciferase output from the P_{TEF1} circuit also indicated the high level of IRP in the cell was sufficiently high to negate repression of IRP transcription by pheromone-induced expression of LexA during the period of the pheromone response pathway. The P_{TEF1} circuit reaches maximum output of luciferase within 120 minutes, comparable with the control however the fold change induction is less than 4 fold, compared with approximately 5 fold

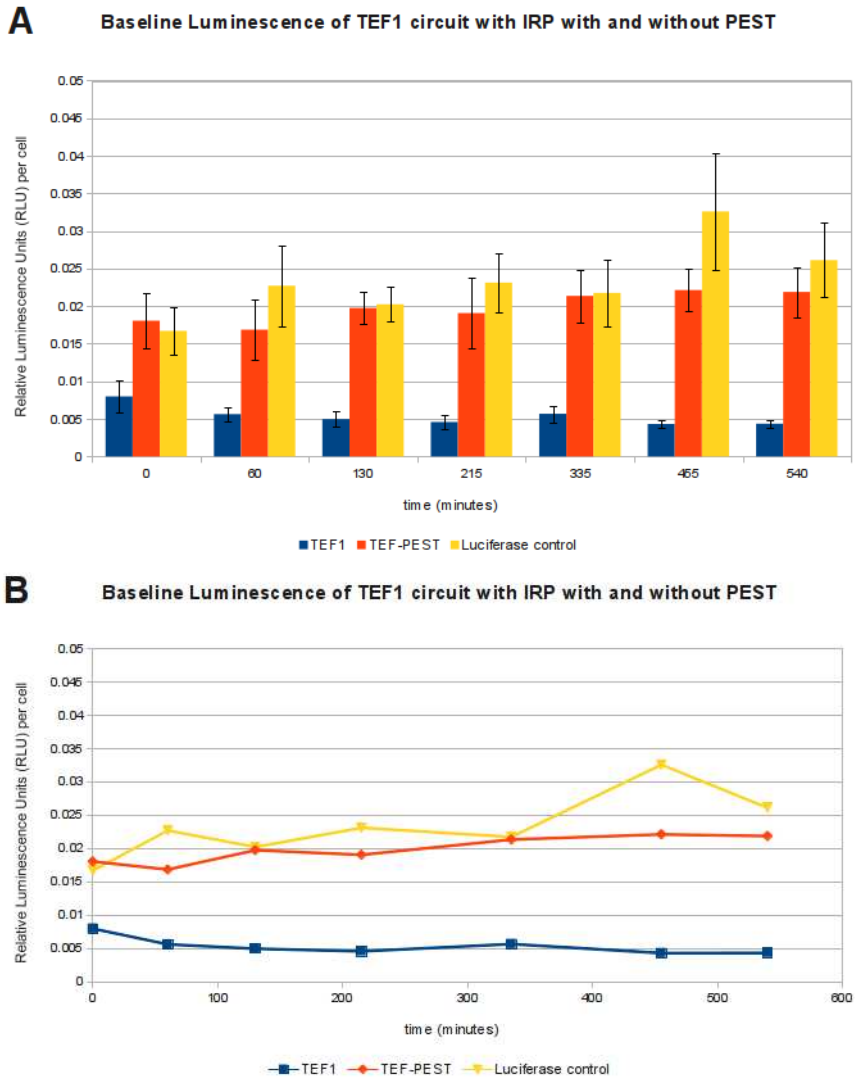

FIGURE 4.6: Baseline luciferase expression for the P_{TEF1} and P_{TEF1}-PEST circuits (n=9). A. Time course measurements of baseline expression for un-induced cells. The error bars represent the standard deviation from the mean calculated for each time point measurement. B. Scatter plot of the same data used in plot A, demonstrating the dynamic range of the luciferase expression of the circuits.

from the control. The P_{TEF1}-PEST circuit data showed a higher level of luciferase expression following pheromone-induction compared with the P_{TEF1} circuit (figure 4.7) with a 1.5 times higher-fold change increase in luciferase. The P_{TEF1}-PEST circuit luciferase expression reached approximately 5-fold induction, similar to the fold change increase of the control. However the P_{TEF1}-PEST circuit requires approximately 200 minutes to achieve maximum output (as with the P_{DCD1} circuits), compared with 120 minutes with the control. After 200 minutes the luminescence fold change reduced to approximately 3 fold within 100 minutes, making a faster return to the OFF-state for

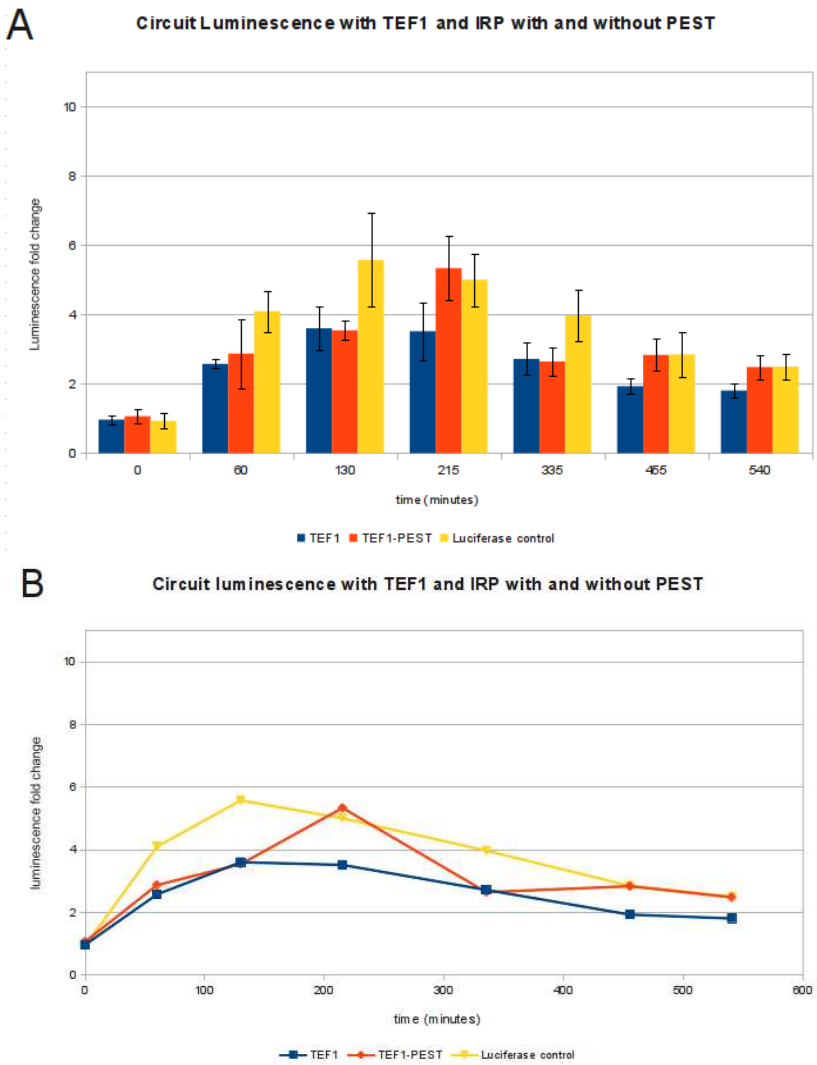

FIGURE 4.7: Fold change in luciferase expression by the P_{TEF1} and P_{TEF1}-PEST circuits compared with the luciferase control (n=9). A. Time course measurements of the relative fold change in luminescence between the induced and un-induced cells expressing the P_{TEF1} and P_{TEF1}-PEST gene circuits. The error bars represent the standard deviation from the mean calculated for each time point measurement. B. Scatter plot of the same data used in plot A, demonstrating the dynamic range of the luciferase expression of the circuits.

the P_{TEF1}-PEST circuit when compared with the P_{TEF1} circuit, as well as the P_{DCD1} circuits. The observation of a faster reduction in the luminescence fold change from the P_{TEF1}-PEST circuit was likely to be a combination of effects of the reduction in the activity of the pheromone response pathway and the decreased repression, relative to the P_{TEF1} circuit.

The maximum fold-induction in luminescence was calculated for each circuit (figure 4.8). The maximum fold-induction was observed from the P_{DCD1} and P_{DCD1} -PEST

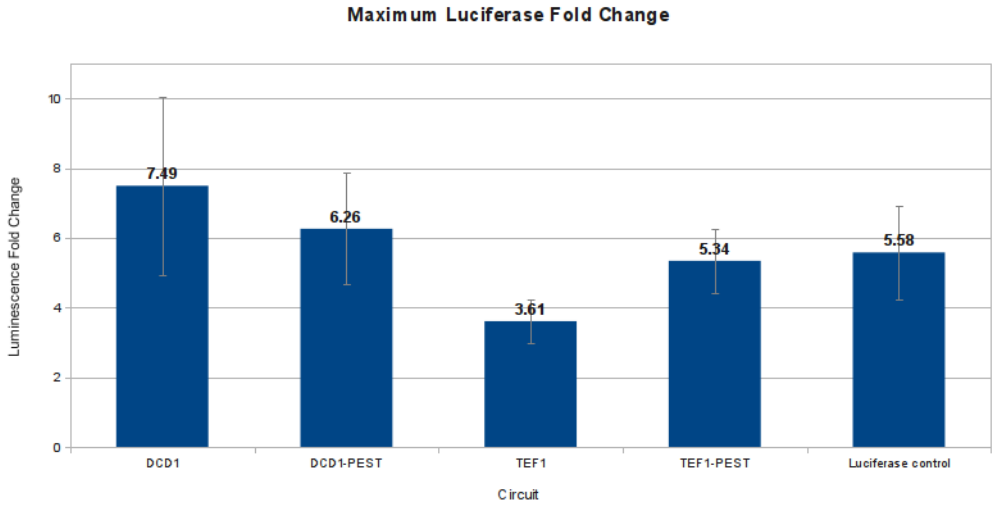

FIGURE 4.8: Maximum luminescence fold change for each circuit (n=9). The control circuit is the pheromone-induced luciferase reporter gene alone, with no repression. The fold change is the ratio of luminescence between the induced and un-induced cells, and the graph shows the maximum ratio achieved by each of the circuits during time course experiment.

circuits, compared with the P_{DCD1} circuits. The P_{TEF1} circuit is repressed by the strong promoter, and addition of the short half-life IRP in the P_{TEF1}-PEST restores the fold-change to a similar level as the control strain. The fold-change luminescence measurements indicated the expression level of the luciferase reporter gene is governed primarily by the rate of IRP production by the repressor plasmid, as predicted by the model (chapter 5, figures 5.13 and 5.14). The pheromone-induced LexA repressor does not appear to have a strong influence on the level of luciferase output as induction times are slower in all of the circuits compared with the control strain (figures 4.4 and 4.7). This could be due to the time required for transcriptional inhibition to reduce the protein abundance in the cell, especially with the long half-life IRP repressor. A high level of IRP mRNA template in the cell coupled with a long-lived repressor protein provides a large reservoir of repressor molecules and template. The IRP's function as a translational repressor also enables it to rapidly inhibit the expression of additional luciferase protein from the mRNA template, enabling it to repress the output of the circuit within a shorter time-scale than LexA can repress IRP expression. As the model predicted from sensitivity and metabolic control analysis (chapter 5), the largest change in the dynamics of the luciferase output were observed from directly perturbing the constitutive promoter strength and half-life of the IRP.

In addition to assessing the circuit behaviour, a set of control circuits were created containing only the reporter and repressor plasmids (figure 4.9, for comparison with the full circuit schematic, see figure 1.17, chapter 1.5.6, page 47). The *S. cerevisiae sst2*Δ cells were not transformed with the de-repressor plasmid containing the pheromone-inducible LexA repressor and therefore constitutively repress the luciferase reporter gene. Differences in the expression of the luciferase reporter gene between the control and full circuits could be attributed to the activity of the LexA de-repressor. The

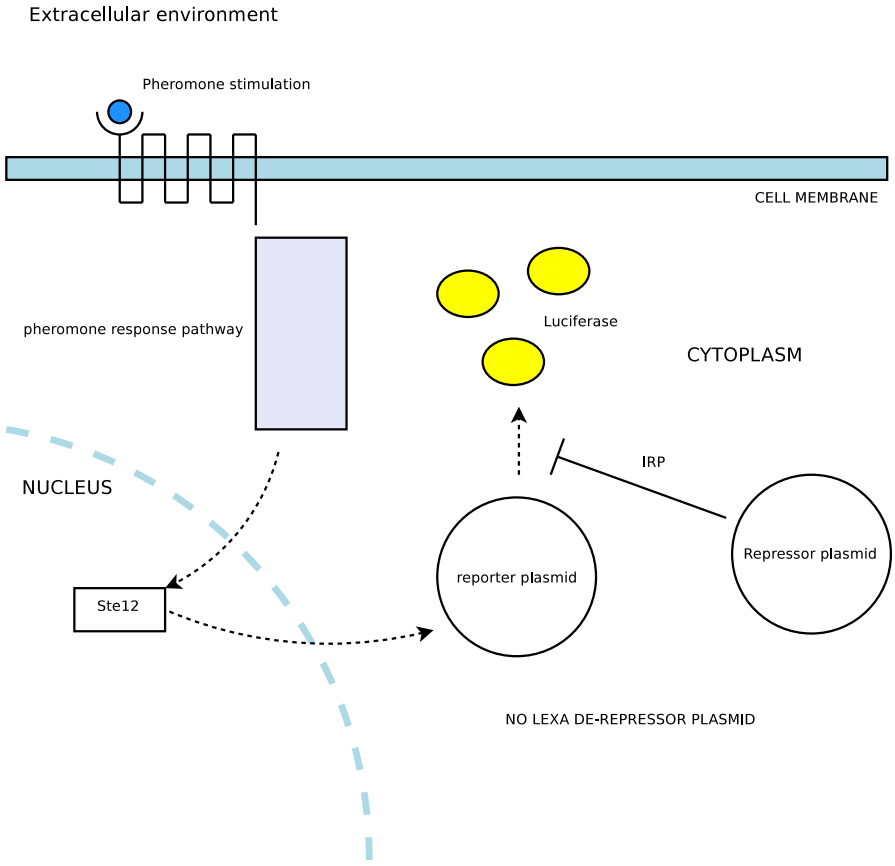

FIGURE 4.9: Schematic diagram of the control experiment with cells transformed with the repressor and reporter plasmids only (no LexA de-repressor plasmid).

data from the P_{DCD1} and P_{DCD1}-PEST control circuits showed a continuous level of repression of luciferase output by the circuits (figure 4.10). Taking 240 minutes as time-point for maximum luciferase expression (see figure 4.10), P_{DCD1} repressed luciferase expression by approximately 60% and P_{DCD1}-PEST by approximately 70%. Across the entire time course experiment, the P_{DCD1} control circuit inhibits expression of luciferase by approximately 80% and the P_{DCD1}-PEST control circuit by 50%. The short half-life IRP_{PEST} repressor had its efficacy reduced by approximately 30% with the addition

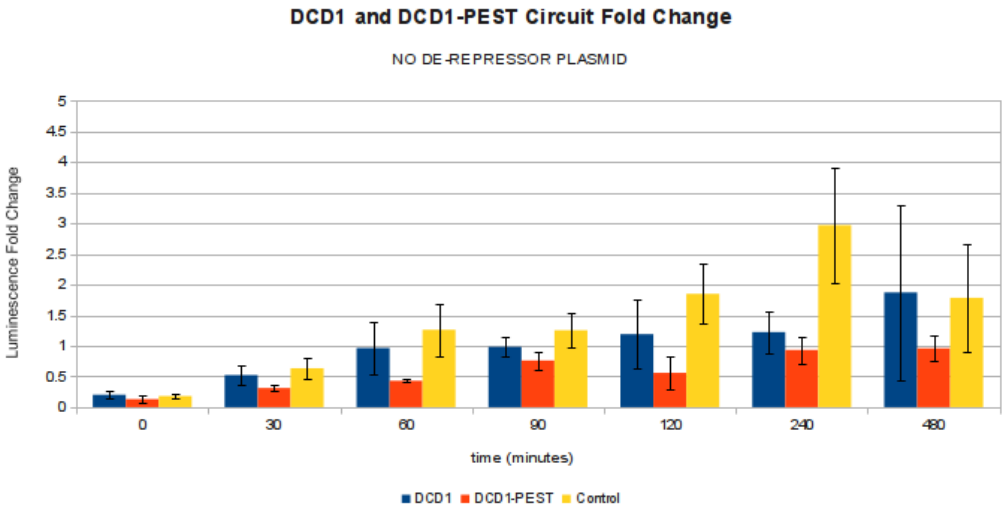

FIGURE 4.10: Fold change in luciferase output from the P_{DCD1} and P_{DCD1}-PEST control circuits minus the de-repressor plasmid (n=3).

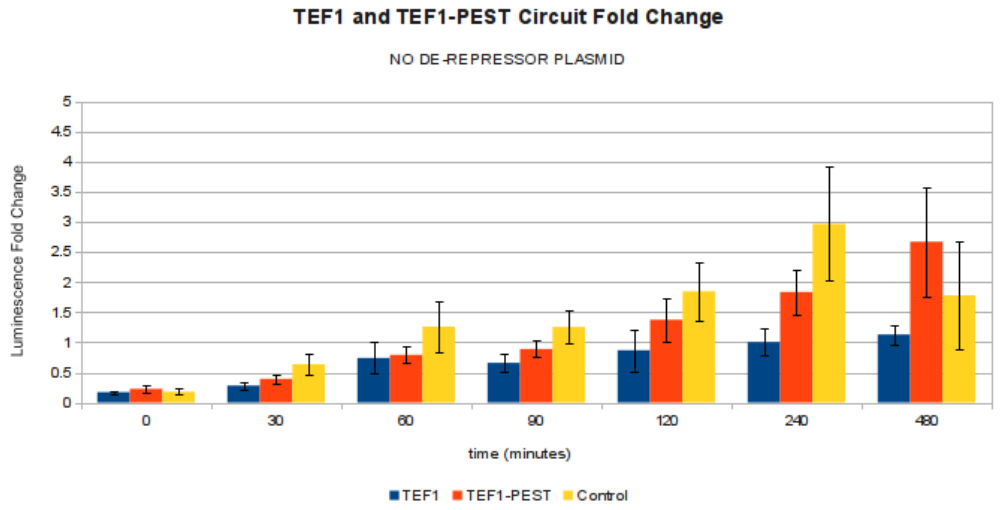

FIGURE 4.11: Fold change in luciferase output from the P_{TEF1} and P_{TEF1}-PEST circuits minus the de-repressor plasmid.

of the PEST degradation tag, compared with the wild-type IRP. The experiment was repeated with the P_{TEF1} and P_{TEF1}-PEST control circuits (figure 4.11). The data from the P_{TEF1} and P_{TEF1}-PEST circuits also demonstrated constitutive repression of luciferase expression, as with the P_{DCD1} and P_{DCD1}-PEST control circuits. At maximum expression of luciferase, taken at 240 minutes in the experiment luciferase expression was inhibited by approximately 66% by the P_{TEF1} circuit, and approximately 40% by the P_{TEF1}-PEST circuit. As observed in the P_{DCD1}-PEST circuit, the PEST degradation tagged IRP repressor activity was reduced by approximately 30% compared with the

wild-type IRP.

For both the P_{DCD1} and P_{TEF1} control circuits, there is no fold change increase in luciferase expression over the control, as observed from the full P_{DCD1} and P_{DCD1}-PEST circuits (figure 4.4). The data provided evidence that the LexA repressor is functioning within the circuit to repress the IRP repression on luciferase translation, and the circuits are functioning as designed. As with the circuit luminescence assay however, there was a large amount of variability in the data which prevents forming significant conclusions on the expression of the circuit components and the characteristics of the circuit.

4.3.1 Luciferase Signal to Noise Ratio

The project goal was to investigate the reduction in basal expression of luciferase by the interaction of the circuit components. Basal expression was recorded as the luminescence from non-pheromone induced cells, and compared with luminescence measured from the pheromone induced cells. The signal to noise ratio (SNR) provides a method of investigating the difference between the signal from the circuit (luminescence following pheromone induction) and the background luminescence (basal expression from non-induced cells).

A low SNR would indicate a smaller difference between luciferase expression from pheromone induced and non-induced cells, and a high SNR indicates a larger difference between basal and pheromone induced luciferase expression, which can be attributed to the activity of the circuit, repressing basal luciferase expression.

The SNR was calculated using the ratio of the mean luminescence measurement at each time point, and the standard deviation from the mean at each time point (section 2.29.4). The luminescence SNR was then compared between circuit strains, and the control strain (pheromone induced luciferase expression with no additional repressors), for each of the time course experiments.

The SNR data for the P_{DCD1} circuits shows an increase in the SNR during the pheromone response (figure 4.12), which was expected as the circuit has a higher output

of luciferase (figure 4.4). The SNR decreases rapidly following the two hour period of the pheromone response compared with the control (figure 4.12 A). This indicates the circuit becomes more noisy in the OFF-state than the control with only the pheromone-inducible reporter gene without the repressor interactions.

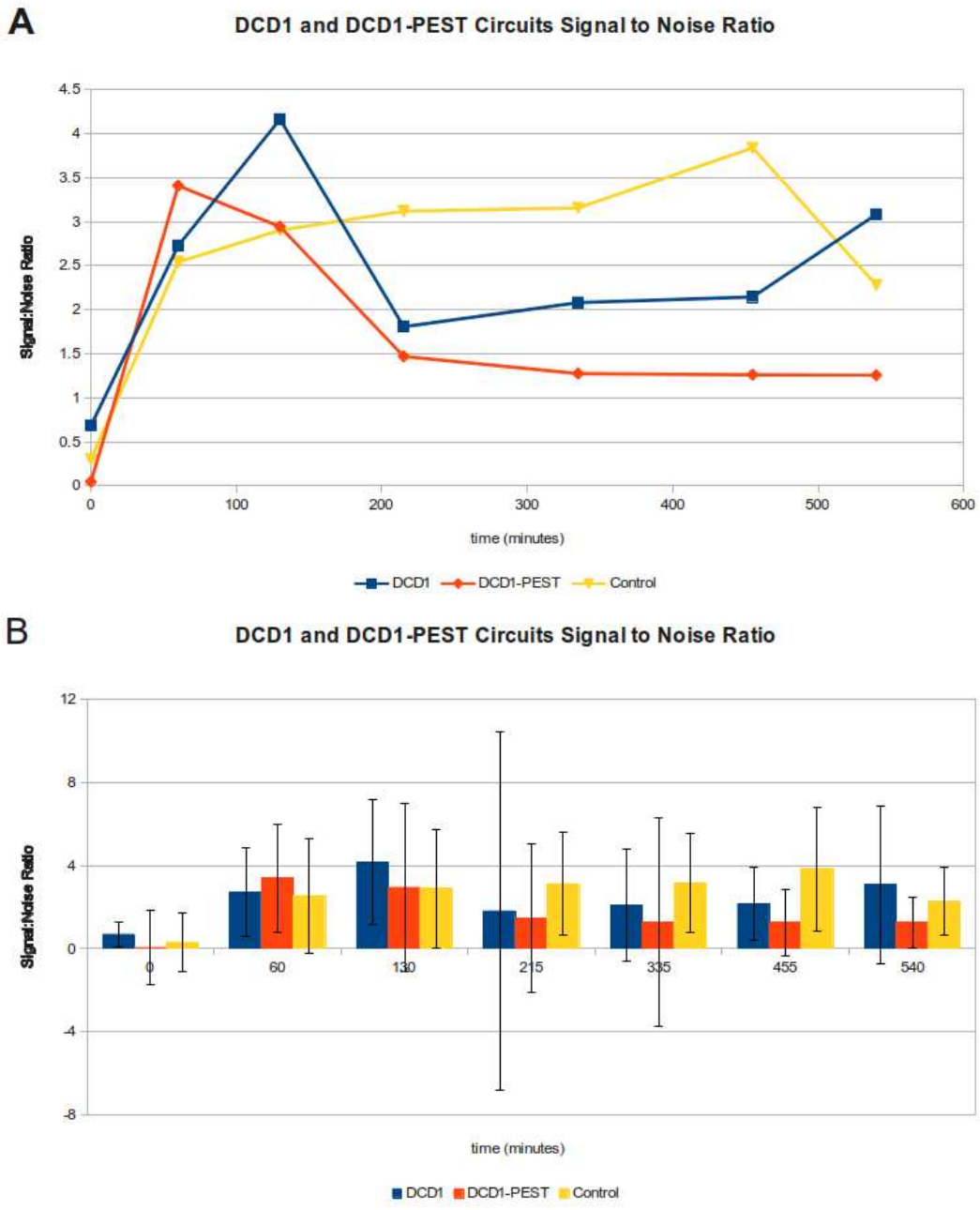

FIGURE 4.12: Luciferase expression signal to noise ratio (SNR) for the P_{DCD1} and P_{DCD1}-PEST circuits, and the control (pheromone induced luciferase expression only) using the luminescence time course data. A and B are the same data represented by two charts due to the size of the error. A. The change in the SNR ratio over time for each circuit throughout the experiments. B. The noise propagation throughout the time course experiment for each circuit.

The SNR for the P_{TEF1} circuits was also calculated (figure 4.13). The data also show the same increase in SNR as the circuit is induced (figure 4.13 A). The P_{TEF1} circuits retain a higher SNR compared with the control strain during the first two hours of the pheromone response, despite the P_{TEF1}-PEST circuit producing the same level of luciferase output as the control (figure 4.7).

The P_{TEF1}-PEST circuit maintains a longer period of high SNR than the P_{DCD1}, P_{DCD1}-PEST and P_{TEF1} circuits over approximately 200 minutes (figure 4.12), and this was attributed to the P_{TEF1}-PEST circuit taking longer to reach maximum luciferase output. Both of the P_{TEF1} and P_{TEF1}-PEST circuits also reduce the SNR rapidly as the luciferase levels reduce, and the circuits returned to the OFF-state (figure 4.13 A). The lower SNR is maintained for the remainder of the time course, as with the P_{DCD1} circuits, compared with the control strain.

The data indicated that both the P_{TEF1} and P_{DCD1} circuits had a lower level of noise in the ON-state compared with the control, and a higher level of noise in the OFF-state, and in transitioning from the ON to the OFF-state (figures 4.12 and 4.13).

The modelling data indicated the addition of the PEST tag to the IRP repressor may increase noise in the expression of luciferase (chapter 5, figure 5.18). A lower SNR was observed from the P_{DCD1}-PEST circuit, compared with the P_{DCD1} circuit (figure 4.12), however it was not observed from the P_{TEF1}-PEST circuit compared with the P_{TEF1} circuit (figure 4.13).

In addition, the SNR was plotted as a function of the luminescence measurement for each of the circuits, to confirm the SNR was increasing with the increase in luminescence (figure 4.14).

The signal to noise ratio of luciferase expression in the control circuits (without the LexA component) was calculated (figures 4.15 and 4.16) for comparison with the full circuit (figures 4.12 and 4.13). The SNR data from the control circuits showed a similar, or lower SNR compared with the control strain (pheromone inducible luciferase expression without the repressors), during pheromone induction (the first 120 minutes

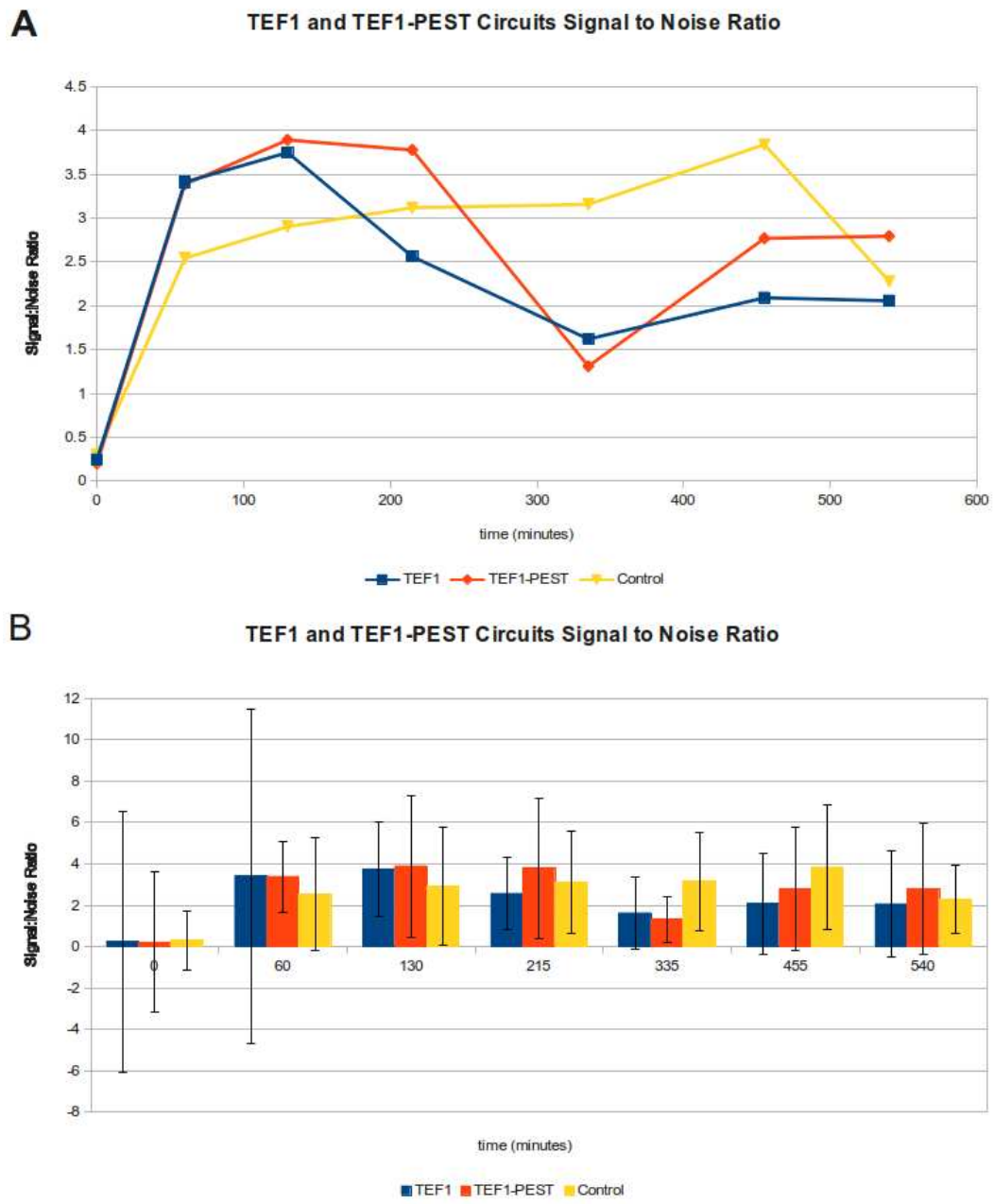

FIGURE 4.13: Luciferase expression signal to noise ratio (SNR) calculated for the P_{TEF1} and P_{TEF1}-PEST circuits , and the control (pheromone induced luciferase expression only), using the luminometer time course data. A and B are the same data represented by two charts due to the size of the error. A. The change in the SNR ratio over time, for each circuit throughout the time course experiment. B. The noise propagation throughout the time course experiment for each circuit.

of the time course experiment), and this was attributed to the lower expression level of luciferase from the control circuits.

The circuits without the LexA de-repressor demonstrated a higher SNR following pheromone activation, compared with the pheromone induced control strain, indicating

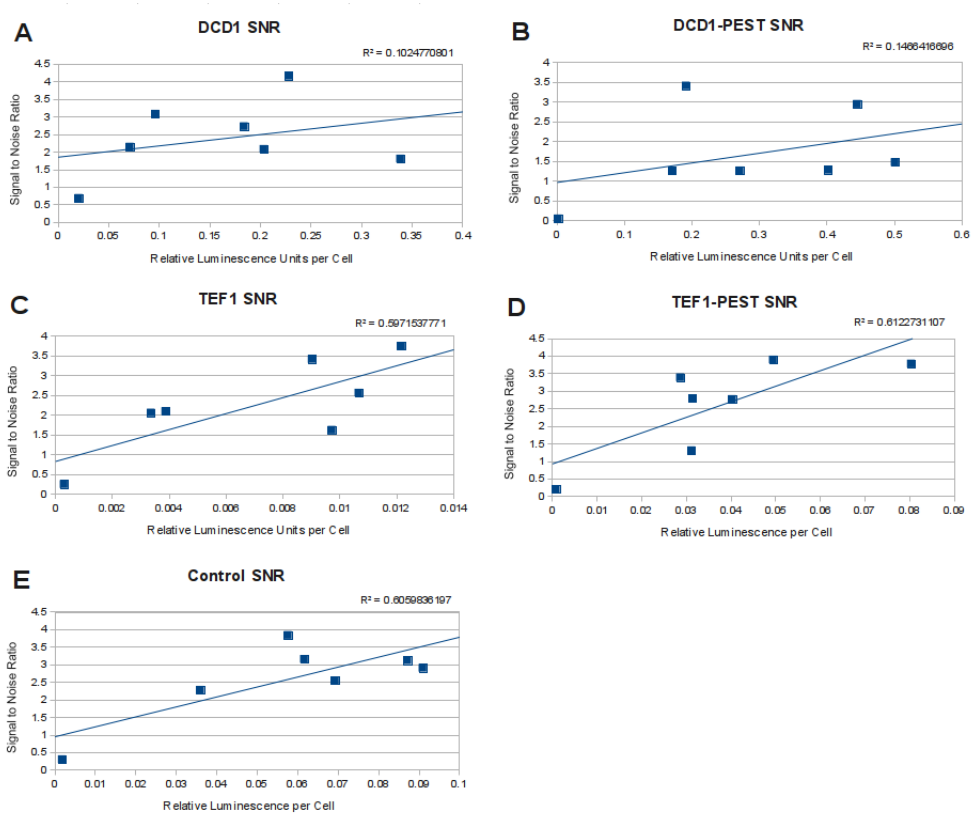

FIGURE 4.14: Signal to Noise Ratio plotted as a function of luminescence for each of the four gene circuits. A. P_{DCD1} circuit with the *DCD1* promoter and wild-type IRP. B. P_{DCD1}-PEST circuit with the *DCD1* promoter and short half-life IRP, C. P_{TEF1} circuit with the *TEF1* promoter and wild-type IRP, P_{TEF1}-PEST circuit with the strong *TEF1* promoter and short half-life IRP. E. Control strain with the pheromone-induced luciferase reporter gene with no repression.

luciferase expression from the circuit without LexA was less noisy than the control strain without the repressors, as well as the full circuit with all of the components.

Noise propagation in the SNR throughout the experiments however was high (figures 4.12 B, 4.13 B, 4.15 B, and 4.16 B), corresponding with the high level of variation in the luminescence data which was used to make the calculations. The size of the error propagation prevents making significant conclusions from the data, and the changes in SNR can only be used as anecdotal evidence of the behaviour of the circuit. More data is required to further investigate the significance of the changes in SNR observed in this work, as well as the contributions of the circuit components to noise (see chapter 6).

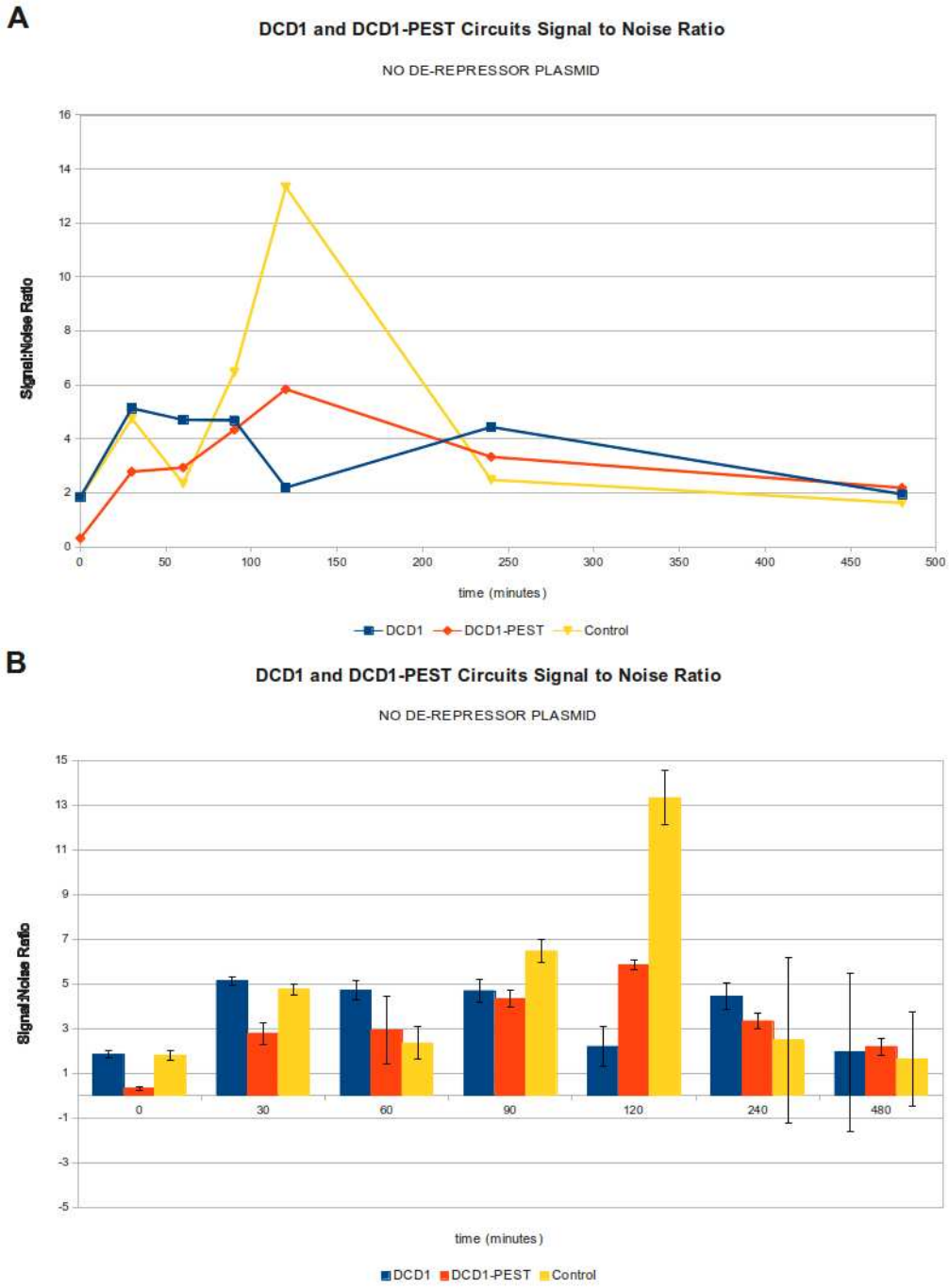

FIGURE 4.15: A and B are the same data represented by two charts due to the size of the error. A. Signal to Noise Ratio calculated for the P_{DCD1} and P_{DCD1}-PEST circuits without LexA, and the control (pheromone induced luciferase expression only), calculated using the luminometer time course data. B. The noise propagation throughout the time course experiment for each circuit.

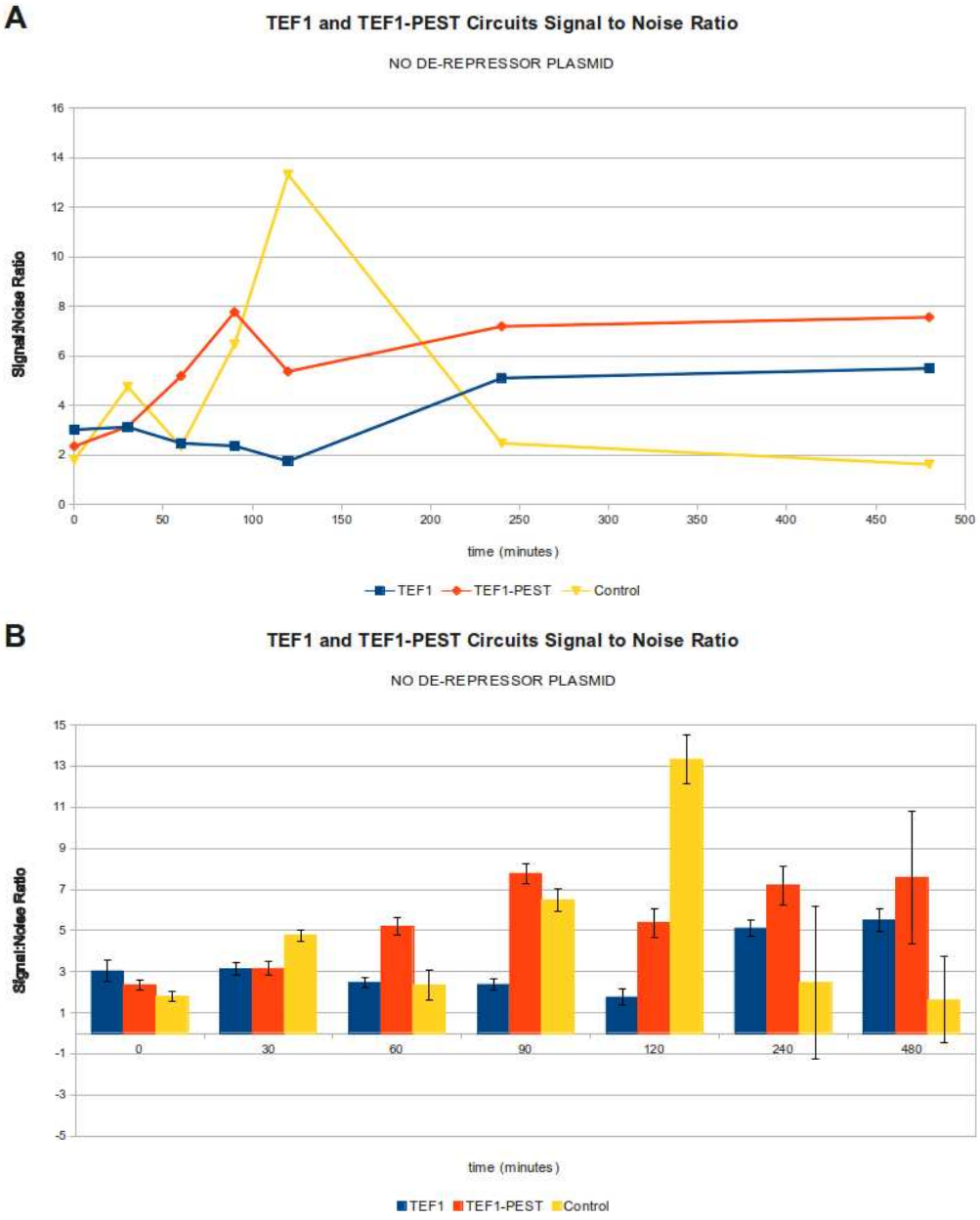

FIGURE 4.16: A and B are the same data represented by two charts due to the size of the error. A. Signal to Noise Ratio calculated for the P_{TEF1} and P_{TEF1}-PEST circuits without LexA, and the control (pheromone induced luciferase expression only), calculated using the luminometer time course data. B. The noise propagation throughout the time course experiment for each circuit.

4.4 Protein Quantification

From the luciferase data it was concluded that the initial time zero, plus the 1 hour, 2 hours, 4 hours, and 5 hours time course samples would provide representative samples for investigating protein expression by western blot. Antibodies for luciferase, LexA,

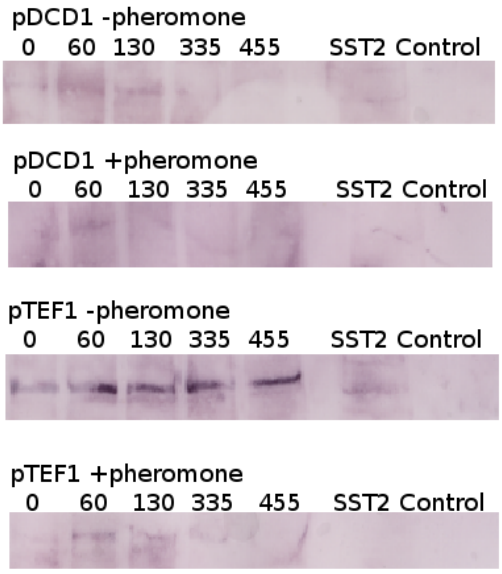

FIGURE 4.17: Representative IRP antibody western blot of the P_{DCD1} and P_{TEF1} gene circuits with (+) and without (-) pheromone-induction. Lanes are labelled with the sampling time (minutes). sst2 Control is the *S. cerevisiae sst2*Δ control strain that was not transformed with the circuit plasmids.

and IRP were purchased from a range of suppliers. An equivalent of 4ml of culture at an OD_{600nm}=1.0 was collected at each time point however, luciferase and IRP proved to be very difficult to detect by western blot (figure 4.17, and 4.20). Luciferase quantification by western blot was not essential as the luminometer provided accurate measurement of this component however, the presence of LexA and IRP could only be inferred from the luminescence data and not quantified. Western blot detection was attempted using a number of secondary antibody conjugates, and corresponding detection strategies were tested in order to enhance LexA and IRP detection. Fluorescein Isothiocyanate (FITC), Horseradish peroxidase (HRP), and infra-red conjugated secondary antibodies were used, and finally alkaline phosphatase bound secondary antibodies with the Promega BCIP/NBT colour development substrate kit. Of these methods, only the alkaline phosphatase conjugate secondary antibodies detected the IRP, but only from the P_{TEF1} and P_{TEF1}-PEST circuits with the higher expression level of IRP (figures 4.17 and 4.20). The data from the IRP western blot (figure 4.17) revealed a signal for the *TEF1* promoter IRP expression, and a weak signal for the *DCD1* promoter which disappeared following pheromone-induction, making it difficult to quantify the fold change reduction in IRP expression from P_{DCD1} and P_{DCD1}-PEST samples. The P_{TEF1} promoter IRP

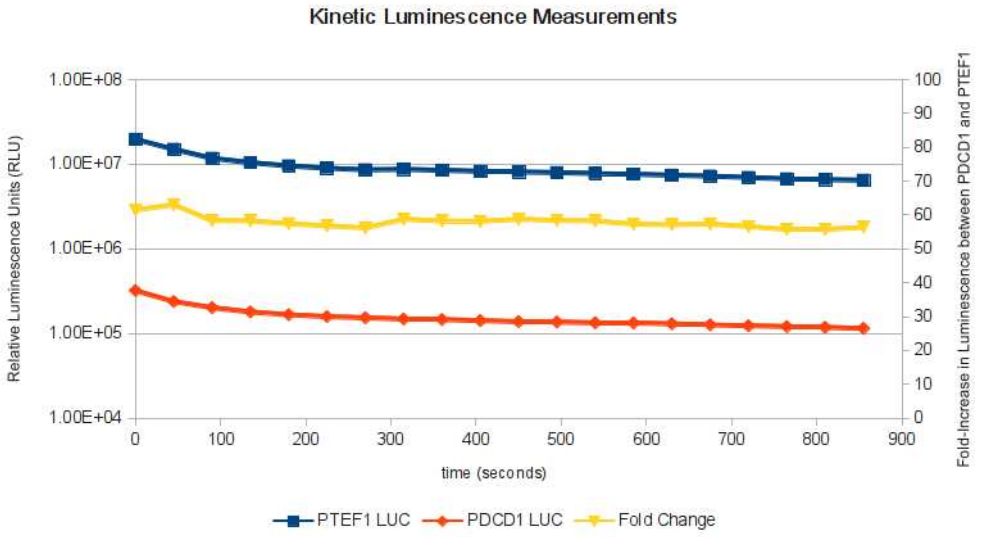

FIGURE 4.18: Luminescence assay of the constitutive expression of luciferase from the *DCD1* (orange) and *TEF1* (blue) promoters. The fold-increase in expression of luciferase from P_{TEF1} compared with P_{DCD1} is plotted on the secondary axis (yellow). Figure recreated with permission from Malys N. (2011, McCarthy Lab unpublished data).

expression signal however was stronger and enabled partial quantification of IRP expression. Kinetic measurements of luminescence provided indication that the *DCD1* promoter expression level was approximately 60 times lower than the *TEF1* promoter (Malys N. and Pietroni P. (2011). McCarthy lab, data unpublished) (figure 4.18), therefore although there was a low level of IRP expression from the *DCD1* promoter, expression could also be estimated from the P_{TEF1} expression levels. For estimating parameter values for the model, the P_{TEF1} circuit could be used to quantify expression which can then be estimated for the P_{DCD1} circuit. Data from the P_{TEF1} circuit IRP expression level was collected and the relative levels of protein expression calculated to investigate the repression of IRP by LexA expressed by pheromone-induced cells (figures 4.19 and 4.20). The results indicated the IRP component of the P_{TEF1} circuit was inhibited during pheromone-induction, compared with non-pheromone-induced cells (figure 4.19). From the data it could be inferred the IRP was being repressed by the pheromone-induced expression of the LexA repressor. There was an increase in IRP expression observed between the initial measurement and the remaining time course measurements, which could not be directly explained and was likely due to variation in the sampling and western blot assay, however measurements across the time period of

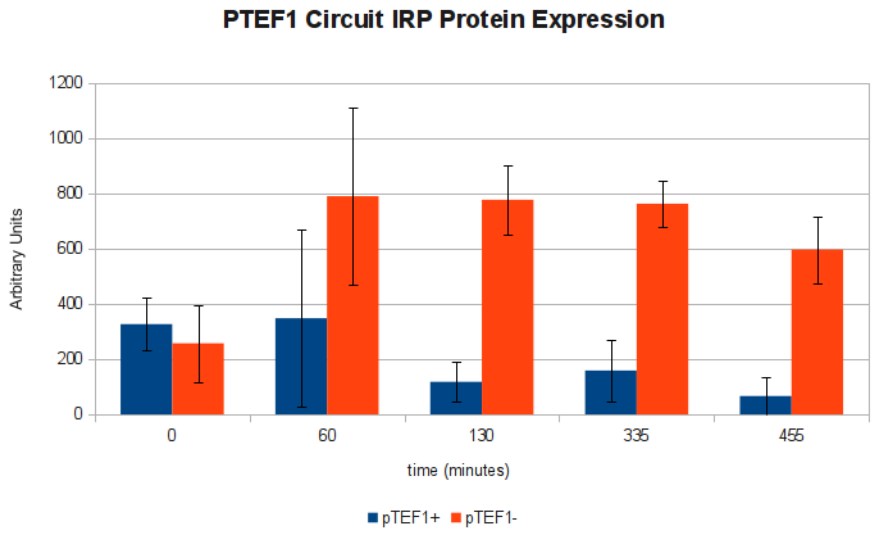

FIGURE 4.19: Analysis of IRP expression from the P_{TEF1} circuit obtained by western blot (n=3). The data show time course expression levels of IRP from the P_{TEF1} circuit. The y axis shows arbitrary units calculated from image intensity.

the experiment were consistent. The average difference between IRP expression for the induced and un-induced cells was calculated across the time course experiment for the P_{TEF1} and was approximately 203 arbitrary units, compared with approximately 636 from the un-induced cells, indicating IRP expression was reduced by approximately 70% during pheromone-induction. If the P_{DCD1} promoter produced 60 times lower expression level of IRP than the P_{TEF1} promoter, and LexA also inhibited the IRP by 70% in the P_{DCD1} and P_{DCD1}-PEST circuits, then this provided an explanation as to why it was not possible to quantify IRP expression from these circuits.

Western blot data was collected for the P_{TEF1}-PEST and expression compared with the P_{TEF1} circuit with the wild-type IRP (figures 4.19 and 4.21). The data showed a reduction in the expression of IRP_{PEST} in the pheromone induced cells, compared with un-induced cells during the period of pheromone-induction (figure 4.21). The average reduction in IRP_{PEST} expression in the induced cells was calculated across the time course experiment as approximately 40% of the un-induced cells, indicating IRP_{PEST} was inhibited by approximately 60% during pheromone-induction. Also, the overall level of IRP_{PEST} was reduced by approximately 10-fold, compared with the wild-type IRP measurements (figures 4.19 and 4.20), and this lower abundance was attributed to the increased degradation rate of the protein due to the PEST sequence. The antibodies

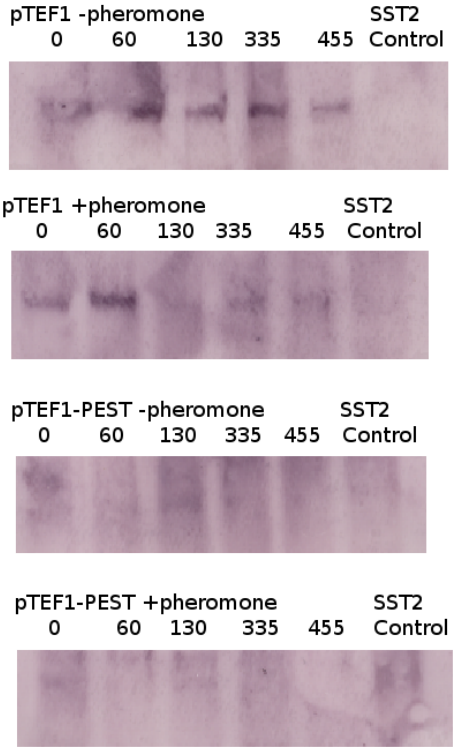

FIGURE 4.20: Representative western blot of the IRP antibodies for the the P_{TEF1} and P_{TEF1}-PEST circuits, with (+) and without (-) pheromone-induction. Lanes are labelled with the sampling time (minutes). The *sst2* Control is the *S. cerevisiae sst2*Δ control strain that was not transformed with the circuit plasmids.

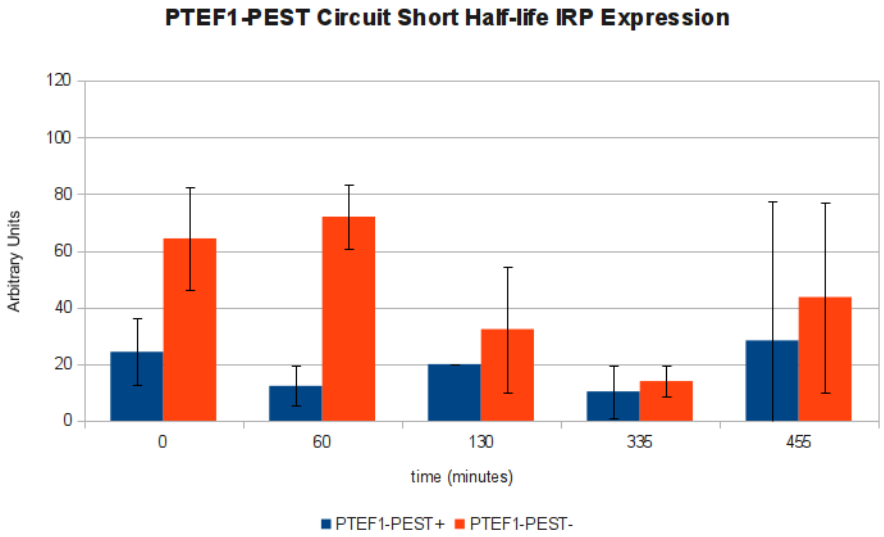

FIGURE 4.21: Analysis of the short half-life IRP_{PEST} expression levels for the P_{TEF1}-PEST circuit (n=3). The graph shows the time course expression levels of pheromone-induced cells (P_{TEF1}-PEST +) and non-induced cells (P_{TEF1}-PEST -). The y axis shows arbitrary units calculated from image intensity.

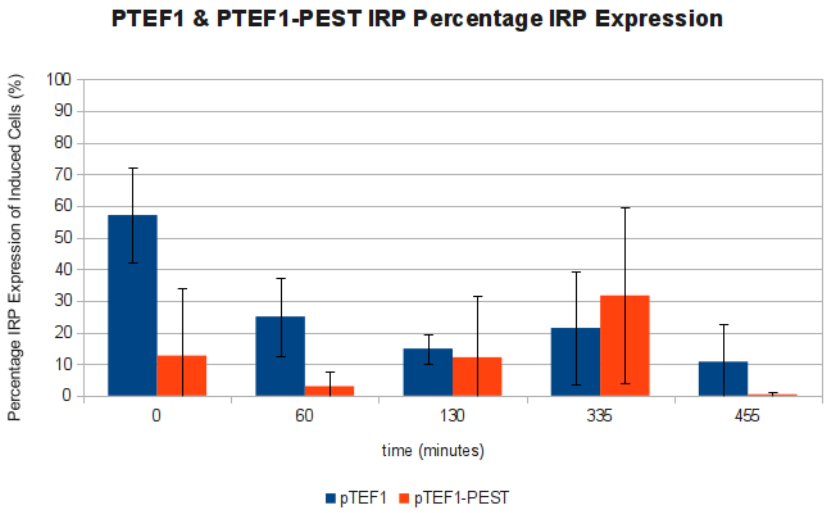

FIGURE 4.22: Percentage IRP expression for the P_{TEF1} and P_{TEF1}-PEST circuits analysed by western blot (n=3). The graph shows the percentage expression level of IRP in pheromone-induced cells as a function of non-induced cells during the time course experiment.

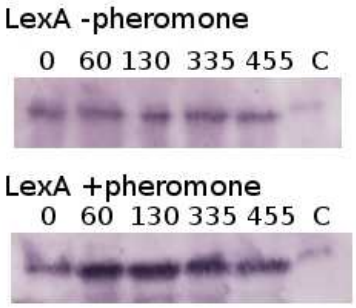

FIGURE 4.23: Representative western blot of LexA protein expression from the P_{DCD1} circuit, visualized using the alkaline phosphatase method. Lanes are labelled with the sampling time (minutes). C is the *S. cerevisiae sst2*Δ control strain that was not transformed with the circuit plasmids.

against LexA proved to be more sensitive than those against IRP and LexA could be detected by western blot (figure 4.23). As with the luciferase luminometer data, there was a high level of variation from the LexA western blot image analysis (figure 4.24). Also, there was a high baseline signal from un-induced cells, that reduced fold-change calculations as were performed with the luminometer data. The baseline data therefore was subtracted from the data from the induced cells and presented in figure 4.25 A and B. The data indicated a directional increase in LexA expression in the pheromone-induced cells, compared with the un-induced cells (figure 4.25 A and B). The directional change was approximately the same for the P_{DCD1} and P_{TEF1} circuits. Using the initial

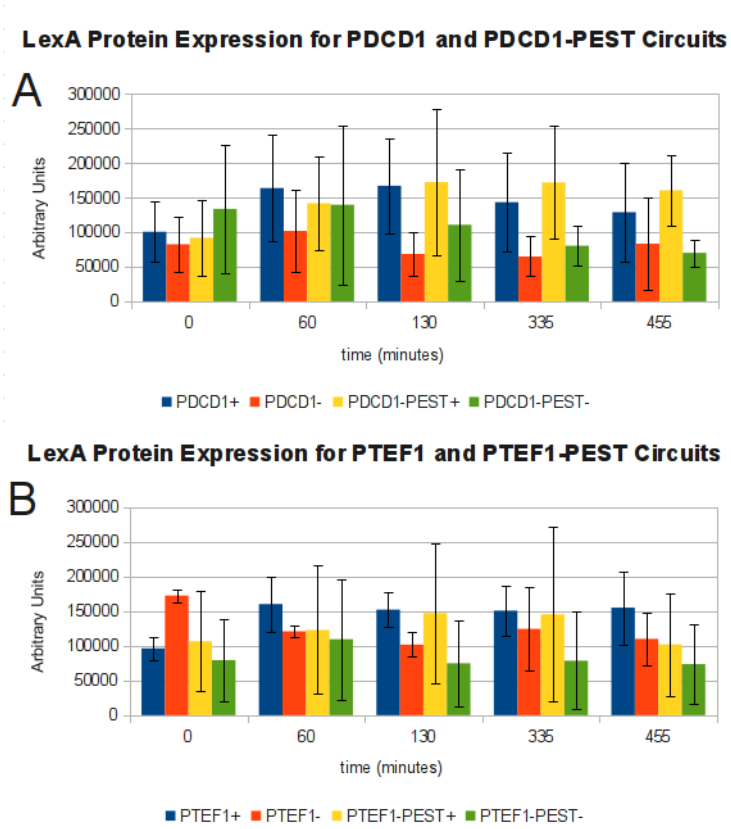

FIGURE 4.24: Graph of LexA protein expression from the gene circuits obtained by western blot (n=3). A) shows the P_{DCD1} and P_{DCD1}-PEST circuit time course samples with (+) and without (-) pheromone induction. B) shows the P_{TEF1} and P_{TEF1}-PEST circuit time course samples with (+) and without (-) pheromone induction.

measurement as a baseline and comparing the expression levels through the time course experiment, the P_{DCD1} circuit demonstrated a fold-increase in LexA expression of approximately five-fold (approximately 2,000 units at the initial measurement to 10,000 units at 130 minutes (figure 4.25 A), and a three fold-increase in the P_{TEF1} circuits (2,000 units to approximately 7,000 units at 130 minutes, figure 4.25 B).

A positive control (recombinant protein) was not available for the IRP, luciferase, and LexA proteins, so a calibration curve for quantitative western blotting could not be performed. Also, the variability in the band intensity from the alkaline phosphatase reaction (figure 4.23) meant accurate quantification of a calibration curve would be difficult, and not quantitative in terms of molecules per cell. The alkaline phosphatase detection method was found to amplify weak signals however the assay is an enzyme reaction and variation in the time allocated for the reaction, and variation in the substrate

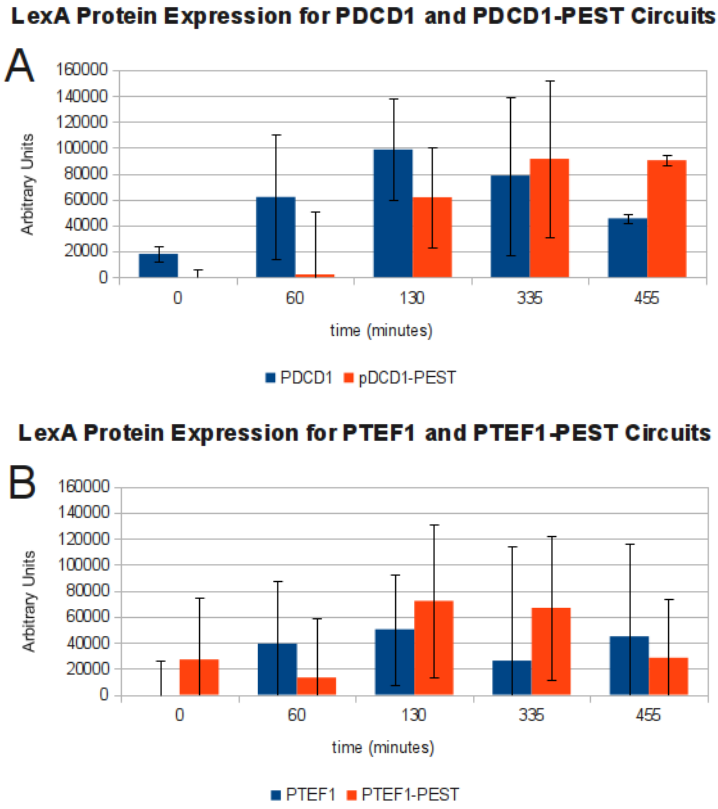

FIGURE 4.25: Graph of the base-line corrected change in LexA protein expression, comparing pheromone induced with non-induced cells, obtained by western blot (n=3). A) P_{DCD1} and P_{DCD1}-PEST LexA protein expression. B) P_{TEF1} and P_{TEF1}-PEST LexA protein expression. The y axis shows arbitrary units calculated from image intensity.

concentration across the membrane can effect the intensity of the band, adding noise to the data.

The western blot data did however enable the measurement of relative-fold changes between the pheromone-induced and non-induced cells, providing information on the relative increases of the components which can be related to the luminescence data (figures 4.4 and 4.7). Although not quantitative, western blotting provided confirmation of expression of the circuit components from the plasmids.

4.5 mRNA Quantification

mRNA expression of the circuit components were quantified using reverse transcription-quantitative PCR (RT-qPCR). As with the protein quantification, the initial time zero,

plus the 1 hour, 2 hours, 4 hours, and 5 hours time course samples would provide representative samples for investigating mRNA expression.

4.5.1 qPCR Housekeeping Gene Selection

A range of housekeeping genes were investigated to find a stable gene with which to normalize the expression of the circuit components. The study by Teste M. *et al* was used to identify a library of housekeeping genes that could be used with *S. cerevisiae* (table 4.2) [257]. The housekeeping genes used in the study were cross referenced with the SGD database tool "Expression Connection" [258] using the micro-array gene expression databases for the yeast pheromone response pathway (figure 4.26) [85]. The micro-array data suggested UBC6 (YER100W), ALG9 (YNL219C), TDH3

ORF	Gene	Maximum fold increase	Maximum fold decrease
YNL219C	**ALG9**	1.0	-1.3
YDR519W	FPR2	1.3	-1.1
YGL040C	**HEM2**	1.1	-1.1
YBR011C	IPP1	1.6	-1.4
YIL075C	RPN2	1.8	-1.1
YDR167W	TAF10	1.0	-1.8
YGR192C	**TDH3**	1.4	-1.2
YER100W	**UBC6**	1.1	-1.2

.

TABLE 4.2: Table of housekeeping gene expression fluctuation during the yeast pheromone response. Expression data was calculated using the SGD Expression Connection software, using micro-array data from Roberts *et al.* [85]. Fold change was calculated as the ratio between gene expression levels in pheromone-induced cells and un-induced cells. Genes marked in bold were identified as potential housekeeping genes for the circuit RT-qPCR experiments.

(YGR192C) and HEM2 (YGL040C) would be suitable housekeeping genes for use with RT-qPCR during the mating response as these genes demonstrates the smallest fluctuation in fold change during the pheromone response (table 4.2). From figure 4.26, UBC6 was consistent throughout the two hour period of the mating response, and was cited as a reproducible house keeping in the Teste study, and therefore was used as the housekeeping gene for assaying the gene circuit mRNA expression levels [257].

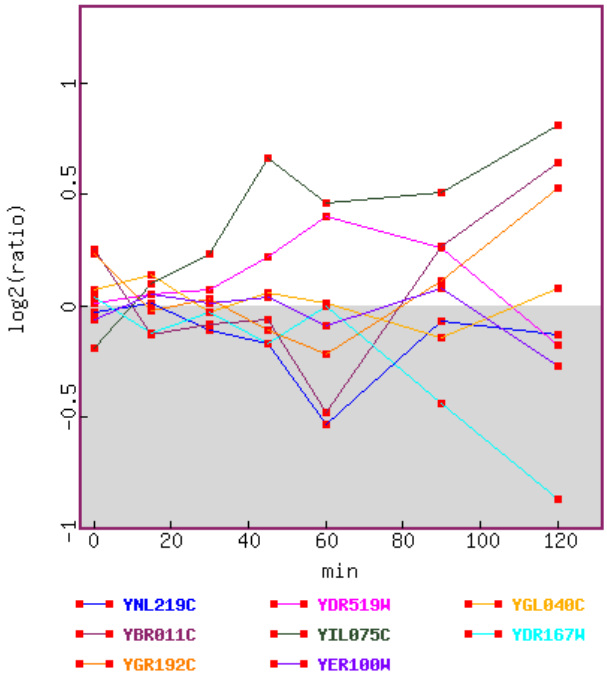

FIGURE 4.26: Graph of the gene expression changes for the housekeeping genes during the yeast pheromone response. The data is plotted with Log2 ratio of induced:un-induced gene expression levels against the two hour duration of the yeast pheromone response pathway.

4.5.2 Primer Validation

RT-qPCR primers were designed for luciferase, LexA, and IRP genes, and validated using plasmid DNA for the circuit primers, and genomic DNA for the housekeeping genes (table 4.3). Miniprep plasmid and genomic DNA samples containing approximately 50ng of DNA were diluted 10^{-1}, 10^{-2}, 10^{-5}, and 10^{-7} and a RT-qPCR experiment performed in duplicate for each primer pair. A standard RT-qPCR cycle was run, as detailed in material and method chapter 2 and the data observed for a signal from each of the primers (figures 4.27). Following the RT-qPCR experiment all of the gene circuit primers, as well as the UBC6, HEM2, and TDH3 housekeeping primers produced a signal at the 10^{-7} dilution using the SYBR green detection method. The ALG9 primers produced a signal at 10^{-5} dilution. The primers therefore demonstrate a detection range of approximately 50ng to 5fg and an amplification efficiency of 1.8 (table 4.3).

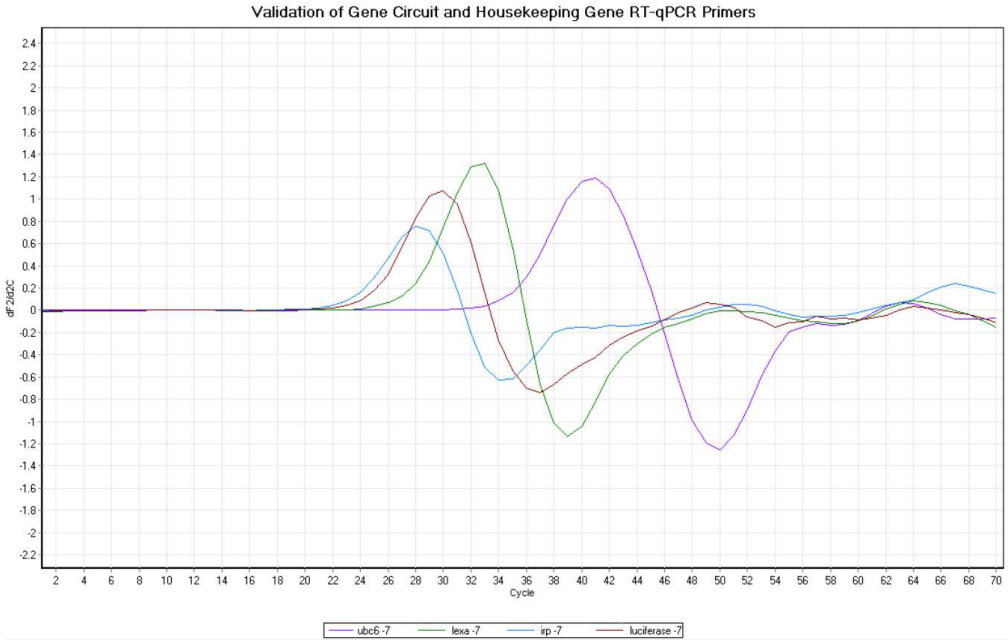

FIGURE 4.27: Representative results of the RT-qPCR primer validation experiment. The graph shows the RT-qPCR SYBR green signal for the luciferase, LexA, IRP and UBC6 housekeeping gene RT-qPCR primers performed on plasmid and genomic DNA.

Name	Take Off	Amplification
UBC6 1	40.6	1.83
UBC6 2	36	1.8
LexA 1	28.8	1.83
LexA 2	28.8	1.8
IRP 1	24.5	1.82
IRP 2	24.6	1.88
luciferase 1	25.5	1.87
luciferase 2	26.3	1.79

TABLE 4.3: Table of RT-qPCR validation data for the circuit and housekeeping gene primers. The table shows the take off cycle and PCR amplification efficiency for each of the RT-qPCR Primers validated on genomic and plasmid DNA.

4.5.3 Sample Preparation

For each of the gene circuit time-points, the total mRNA was extracted as in the material and methods section (chapter 2) and RT-qPCR used to quantify the relative amount of mRNA in each. Each RT-qPCR reaction was repeated in triplicate, and the amplification efficiency checked for a minimum of >1.6. For each of the RT-qPCR experiments, the initial time point (t0) was used as the calibrator for the instrument software in order to perform comparative quantification. The data was then expressed as the relative increase

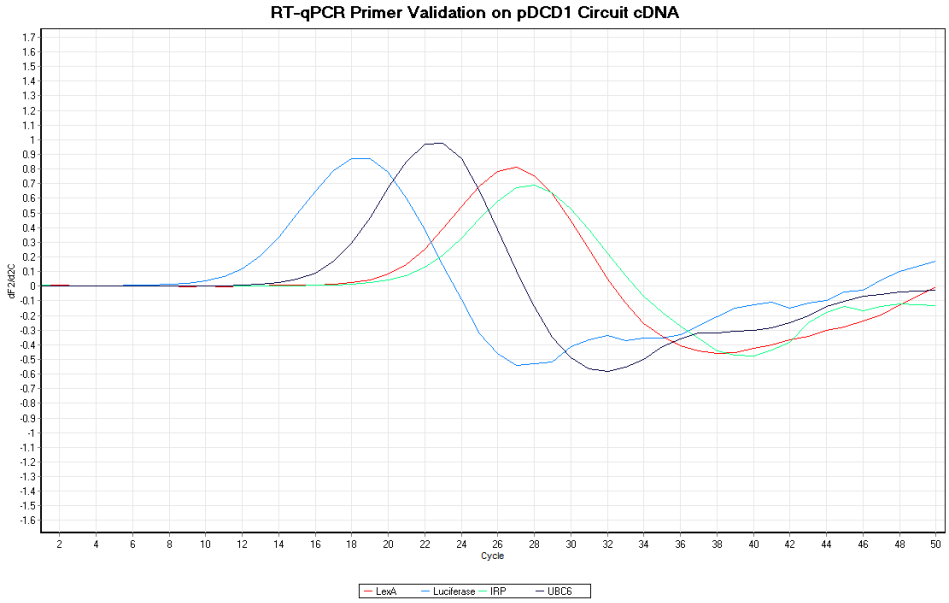

FIGURE 4.28: Representative results of the RT-qPCR primers with P_{DCD1} cDNA. The graph shows the RT-qPCR SYBR green signal for the luciferase, LexA, IRP and UBC6 housekeeping gene RT-qPCR primers performed on plasmid derived cDNA

Name	Take Off	Amplification
LexA 1	21.3	1.68
LexA 2	21.7	1.71
LexA 3	22.0	1.71
luciferase 1	13.3	1.66
luciferase 2	13.0	1.66
luciferase 3	13.5	1.69
IRP 1	23.0	1.64
IRP 2	22.8	1.39
IRP 3	22.6	1.66
UBC6 1	17.5	1.68
UBC6 2	17.9	1.71
UBC6 3	18.1	1.74

TABLE 4.4: Representative RT-qPCR data using P_{DCD1} circuit plasmid cDNA.

from the initial time point measurement. Each of the data points was then normalised against the signal from the UBC6 housekeeping gene, and the fold change calculated for the induced cells compared with the un-induced cells. Finally, the mRNA expression data was logged to base 2 to scale the data. A representative result from an RT-qPCR experiment on P_{DCD1} plasmid derived cDNA is shown in figure 4.28.

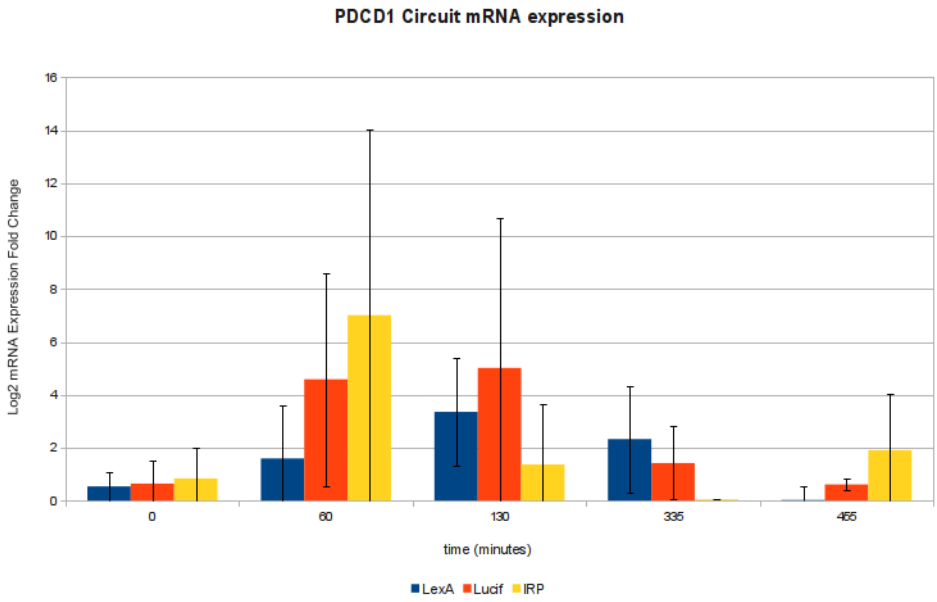

FIGURE 4.29: Graph of the mRNA transcription data for the P_{DCD1} circuit following pheromone induction, obtained by RT-qPCR.

4.5.4 pDCD1 Circuit qPCR Analysis

The data from the RT-qPCR experiments demonstrated a high level of variation. The results from the P_{DCD1} circuit however show a directional increase in luciferase and LexA mRNA expression during the first two hours of pheromone response (figure 4.29). The luminometer data for P_{DCD1} showed maximum luminescence between 200 and 300 minutes (figure 4.4). This delay between mRNA expression and the detection of luminescence indicated the IRP was repressing translation of luciferase mRNA.

4.5.5 pTEF1 Circuit qPCR Analysis

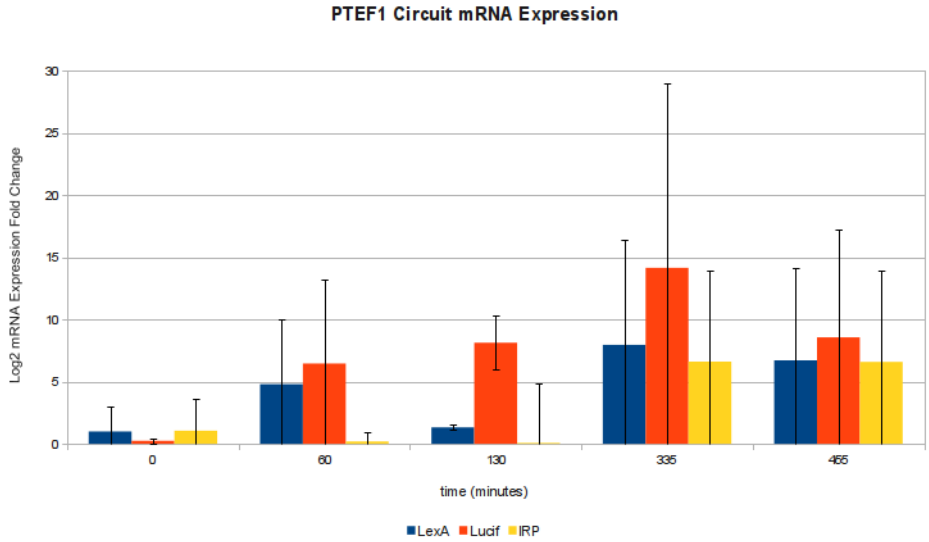

FIGURE 4.30: Graph of the mRNA transcription data for the P_{TEF1} circuit following pheromone induction, obtained by RT-qPCR.

The data from the P_{TEF1} circuit showed luciferase mRNA expression increased in the first two hours of the pheromone response (figure 4.30). The luminescence data from the P_{TEF1} circuit also showed maximum luminescence between 200 and 300 minutes, and luciferase expression in the P_{TEF1} circuit was lower than the P_{DCD1} circuit (figures 4.7, and 4.8). This data further confirmed the IRP was functioning as a repressor of luciferase mRNA translation, as transcription levels for the P_{DCD1} and P_{TEF1} circuits were comparable but luminescence expression profiles were different (figures 4.4, and 4.7) .

The data for the IRP mRNA expression indicated a directional decrease in expression during the pheromone response from the P_{TEF1} data, indicating LexA was repressing transcription of IRP mRNA. The variation in the data for IRP mRNA transcription is extremely high however and may have been due to the low level of expression of IRP from the *DCD1* promoter creating noise in mRNA expression. However this is speculative and further replicate data was required to obtain more reliable data.

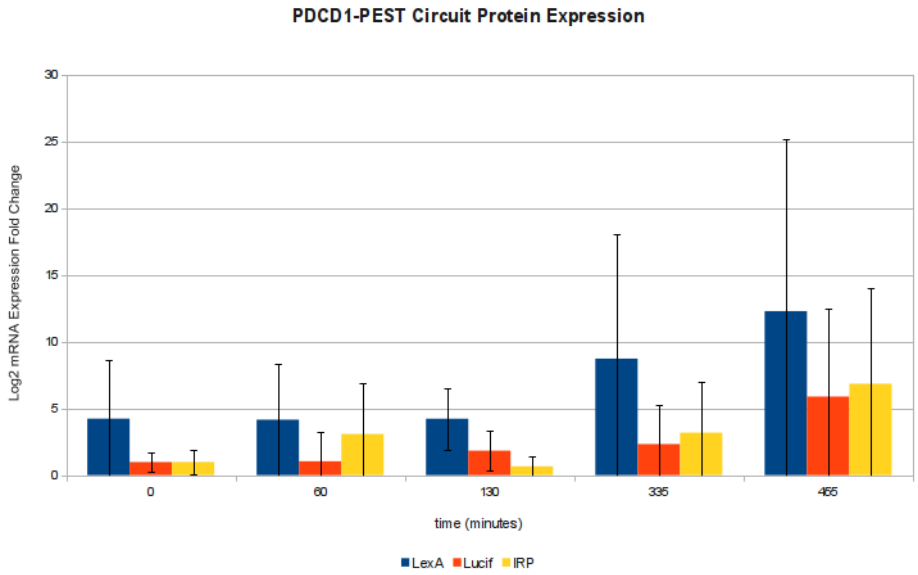

FIGURE 4.31: Graph of the mRNA transcription data for the P_{DCD1}-PEST circuit following pheromone induction, obtained by RT-qPCR.

4.5.6 pDCD1-PEST Circuit qPCR Analysis

The P_{DCD1}-PEST circuit demonstrates the same trend in mRNA transcription as the P_{DCD1} circuit (figures 4.29 and 4.31). The P_{DCD1}-PEST circuit mRNA transcription shows a decrease in the expression of IRP during the first two hours of the pheromone response. It could be inferred that this is due to pheromone-induced expression of LexA, inhibiting IRP_{PEST} transcription. Luciferase and LexA also remain up-regulated during the pheromone response, however the data are highly variable, limiting a conclusive analysis of the relative changes in transcription.

4.5.7 pTEF1 Circuit qPCR Analysis

The P_{TEF1}-PEST circuit data demonstrated an increase in luciferase mRNA transcription in the first hour of the pheromone response (figure 4.32). There was a directional decrease in IRP expression compared to the P_{DCD1}-PEST data (figure 4.31) with no increase in transcription towards the end of the experiment as observed in the P_{DCD1} circuits, however as with the previous experiments there is a high degree of variability in the data.

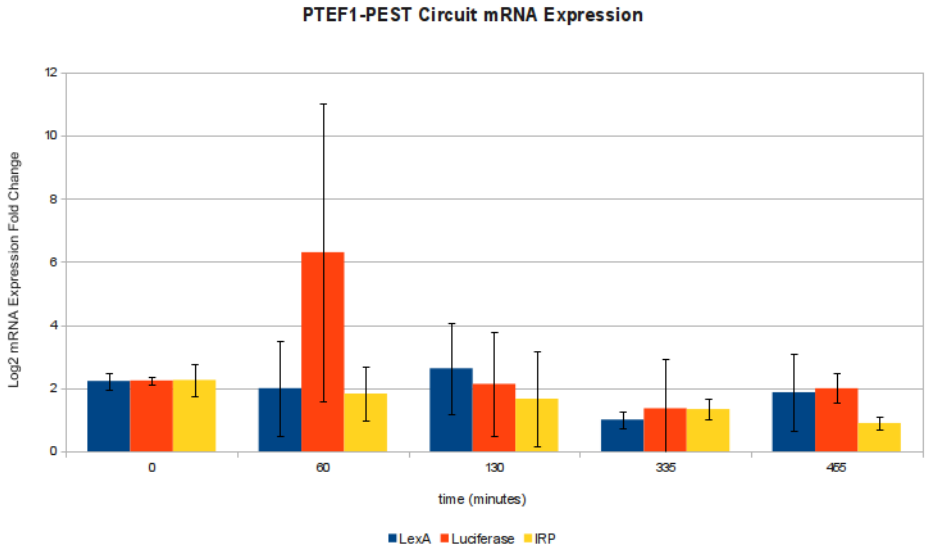

FIGURE 4.32: Graph of the mRNA expression data for the P_{TEF1}-PEST circuit following pheromone induction, obtained by RT-qPCR.

4.5.8 qPCR Analysis Summary

The data from the RT-qPCR experiments were not sufficiently accurate for modelling, as the data has a high degree of variability. The data do however provide evidence that the components are present in the circuits and interacting as there is a signal from the IRP, luciferase, and IRP transcription, and a directional change in the transcription levels of each (figures 4.29, 4.30, 4.31, and 4.32).

A larger number of replicates was required to obtain more reliable RT-qPCR data for model fitting (chapter 5), and to find statistically-significant differences in the mRNA transcription levels during the pheromone response. Also, a larger number of housekeeping genes should be run with each circuit to enable normalisation across a range of housekeeping genes as this may also reduce the level of variability in the data.

In this work, three biological replicates were performed however for MIQE guidelines it would be preferable to study ten housekeeping genes and perform ten biological replicates of each experiment [259, 260]. However, for logistical reasons this was not possible for this project. The RT-qPCR data provides confirmation of the interactions of the components in the circuit, particularly for the P_{DCD1} and P_{TEF1} circuits where it can be seen that IRP mRNA levels are being down-regulated in the P_{DCD1} circuit (figure

4.31), but not in the P_{TEF1} circuit (figure 4.32), providing insight into the kinetics of the LexA repressor. Inhibition of IRP is observed in the P_{DCD1} circuit (figure 4.29), but a higher expression level of LexA is required for inhibiting IRP transcription from the P_{TEF1} promoter.

4.6 Conclusion

The luciferase expression from the variations of the circuit indicated there was a quantifiable difference in the dynamic behaviour of the circuits as a result of incorporating the high and low expression levels of the IRP repressor and perturbing the IRP half-life through the addition of the PEST degradation tag. Western blot data indicated that fusing the PEST-rich C-terminal domain of the Mateus and Avery yEGFP3 reporter gene to the IRP conferred a reduction in half-life of the IRP as seen in the yEGFP (figure 4.21) [216].

Increasing the expression level of IRP reduces the overall signal output (the amplitude of the signal), as expected, and changing the half-life can increase the signal and also tune the deactivation of the circuit, reducing the period of the signal (figure 4.7). The IRP and LexA repressors enable a higher-fold change in the output of the circuit between the OFF and ON-states. The P_{DCD1} circuit can be used to achieve the highest fold change output of the cells (figure 4.4) and the reduction in half-life can be used to extend the period compared with the control. The circuit therefore provides an increase in luciferase compared with a pheromone-induced reporter gene alone, and a tunable output in terms of luciferase expression.

CHAPTER

FIVE

MODELLING

5.1 Introduction

Modelling of biological systems is an integral part of the design and development process in synthetic biology [8, 10]. Mathematical modelling provides hypothesis generation for designing gene circuits [178, 180], and computer simulation enables prediction of complex behaviours that would be difficult to anticipate or investigate experimentally [261, 262]. The process of computer modelling followed by laboratory experimentation creates an iterative cycle of design and development that is a characteristic of synthetic biology [14, 263].

The gene circuit model created for this project was required to predict control points and the dynamic range of the output of the circuit, providing an indication of which components and interactions would be most significant, prior to laboratory implementation. Later, as the circuit was characterized in the lab, gene and protein expression data from the components could be collected and used to parameterise the model, allowing it to evolve into an accurate predictive tool for understanding the experimental observations. This approach provides direction for the project at inception,

and then early prediction of behaviour as the circuit is tested in the laboratory. A model that is parameterised entirely by experimental data is a valuable predictive tool that can be used for further refinement of the circuit behaviour [14, 169, 173].

The gene circuit was designed to be activated by the Ste12 transcription factor from the yeast pheromone response pathway. The mating response pathway formed the input to the circuit and involved a complex cascade of reactions that ultimately resulted in the activation of the mating response genes [55]. It was decided at the beginning of the project to include the mating response pathway in the model as an "input module" for the gene circuit reactions for accommodating the dynamic behaviour of the MAP kinase cascade [128]. Currently, the most detailed model of the pathway that has been published is by Bente Kofahl and Edda Klipp [128]. The model attempts to provide a complete simulation of the pathway, including G-protein activation, the MAP kinase cascade, and activation of the Ste12 transcription factor (figures 1.8 and 1.9). Although the model contains all of the components of the yeast pheromone response pathway, the individual reactions are not required when simulating the gene circuit, and adds additional computational time to the simulations. Therefore in this project, the model was simplified into a single rate equation to translate the flux through the pathway into a generic input function for the circuit. To achieve this, the model was reduced to only the components directly interacting with the cascade. The input was limited to the formation of the complete Ste5 complex - Ste5, Ste7, Ste11,Gβγ, Fus3, designated "complexD" in the model (figure 1.9) [128]. The output was taken as the phosphorylated form of Ste12. The Bar1 and Far1 reactions controlling cell elongation, chemotaxis, and pheromone degradation were removed, and the components of the MAP kinase cascade fixed at their initial values. Changing the initial concentration of complexD and observing the change in flux through the final reaction (v34: the phosphorylation of Ste12 to Ste12pp) (table 5.1) enabled simulation of only the cascade components directly involved in the phosphorylation of Ste12.

As the model was based on simple mass-action kinetics the resulting output was a linear increase in output in response to increasing input (see figure 5.1 A). The flux through the reactions other than v34 were found to be minimal compared with reaction v34, therefore it was assumed that the flux through the mating response

pathway model is primarily towards the phosphorylation of Ste12 (see figure 5.1 B). The slope of the graph produced from scanning through reaction v34, was taken as the rate of phosphorylation of Ste12. The model was then simplified by removing all of the components between complexD and the double phosphorylated form of Ste12, replacing them with a single reaction, multiplied by the rate constant for reaction v34 from the original model (see equation 5.2). The final "minimal model" took the input from the activated trans-membrane G-protein and activated the Ste12 transcription factor through a single reaction, representing the flux of the reactions through the MAP kinase cascade (equation 5.1).

$$\frac{d[Fus3pp]}{dt} = -(k1_{v33} \cdot [Fus3pp]] - (k1_{v34} \cdot [Ste12] \cdot [Fus3pp] + (k1_{v35} \cdot [Ste12active]) + (k1_{v32} \cdot [complexD])) \tag{5.1}$$

Reaction	Reaction Details	Rate (reactions per minute)
v32	complexD $\rightleftharpoons$ Fus3pp	3.42x10^6
v33	Fus3pp $\rightleftharpoons$ Fus3	50
v34	Ste12 + Fus3pp $\rightleftharpoons$ Ste12 active	18
v35	Ste12 active $\rightleftharpoons$ Ste12 + Fus3pp	10

TABLE 5.1: The simplified MAPK model based on the yeast pheromone response pathway model by Kofahl and Klipp [128]. The rate equations and parameters for all reactions were taken from the model by Kofahl and Klipp, and v32 was modified with the rate calculated from equation 5.2. "=" denotes a reversible reaction.

The rates of the reactions v32, v33, v34, v35 were taken from the Kofahl and Klipp model, and are listed in table 5.1. Reaction v32 was modified to incorporate the rate of flux through the MAP kinase cascade (figure 5.2) calculated from equation 5.2:

$$r = k \cdot [complexD] \tag{5.2}$$

Equation 5.2 was derived from the observation of the change in flux through reaction v34 as a function of the initial concentration of complexD. The rate "r" is equal to the slope of the line from the flux through v34 multiplied by the concentration of complexD. The slope of the line in figure 5.1A was calculated as r = 3.42x10^6. The complexes that make up the MAPK cascade were then deleted and a single reaction inserted for the

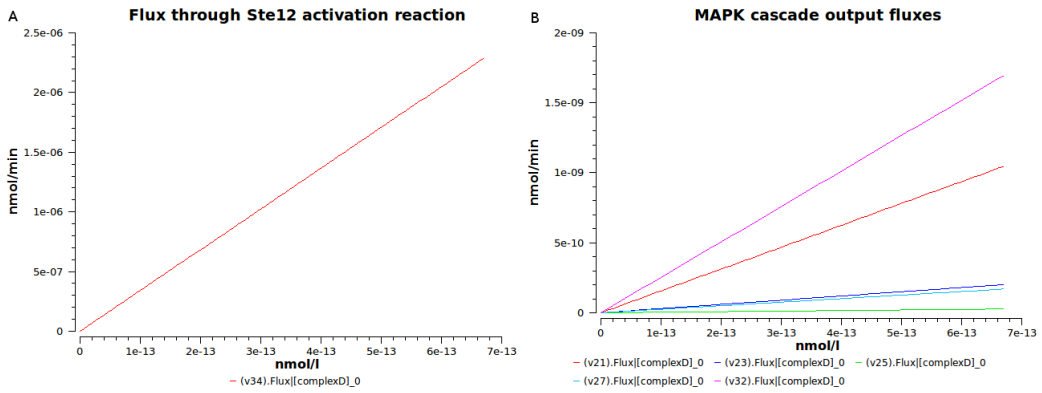

FIGURE 5.1: MAP kinase cascade outputs plotted with varying initial concentration of complexD. The graph shows the initial concentration of complexD (x axis) as a function of the concentration of Ste12 active (y axis). The concentration of Ste12active is shown in nmol/l however this is an arbitrary measure of concentration and is not an accurate prediction of the cellular concentration of phosphorylated Ste12. A: Flux through reaction v34, representing the change in the concentration of phosphorylated Ste12. B: flux through the reactions producing the individual components of the MAP kinase cascade, excluding phosphorylated Ste12.

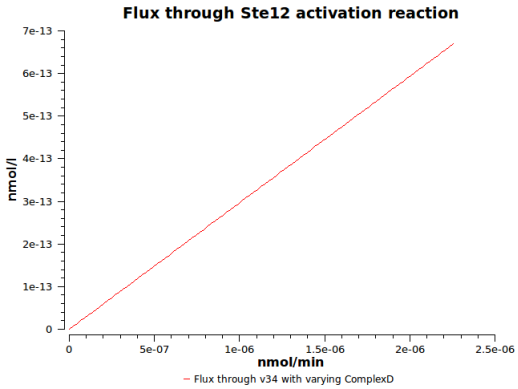

FIGURE 5.2: Flux through reaction v34 in minimal model incorporating a single rate equation to describe the behaviour of the MAP kinase cascade.

transformation of complexD to Ste12pp, with the reaction constant *r*. This model was then run and the output behaviour checked for the new single step model. The new simplified model was used to parameter scan increasing concentrations of complexD and observe the rate of change in the concentration of phosphorylated Ste12 through reaction v34 (see figure 5.2). The slope of the graph was taken as 3.37×10^{6}, which is comparable with the original slope of the MAP kinase model (figure 5.1). From this result, the new simplified model was taken as representative of the original Kofahl and Klipp model for the rate of the activation of Ste12.

5.2 Modelling Eukaryotic Signal Cascades

After further investigation of the Kofahl and Klipp model (chapter 1.4), it was found that the model was not sensitive to the level of pheromone and that the steady-state concentrations of phosphorylated forms of Fus3 and Ste12 were always zero, with fixed levels of pheromone input. Some of the model parameter values were obtained from the literature, but many had been fitted to validate the behaviour against wet lab data of various mutants of the pheromone response pathway [128]. To this end, the model predicts the results it was designed to predict, however it has little further utility, and cannot thus be used to investigate other dynamics of the pheromone response, such as ultra-sensitivity and the dynamic range of the signal response, as reported by O'Shaughnessey or Yi [46, 126].

5.2.1 A Revised Mating Pathway Model

A mechanistic model of the mating pathway was constructed from current descriptions of the pathway found in the literature (table 5.2) [55, 80]. There are no published studies that accurately quantify the components of the yeast mating pathway. A comprehensive search of the literature was conducted to find the most accurate quantification of the components of the MAPK pathway in yeast (table 5.3). The behaviour of the model using these parameters was compared with published experimental observation [126], and models by Kholodenko, Huang and Ferrell, and Wang [120, 131, 264]. The model incorporated the MAP kinase cascade reactions of the yeast pheromone response pathway and models the successive phosphorylation of the MAPK components, Ste11 (MAPKKK), Ste7 (MAPKK), and Fus3 (MAPK) (figure 5.3). Ste5 was not included in the model reactions as it was believed it is primarily involved in maintaining signal specificity rather than facilitating signal transduction [71, 268]. The kinases interact as they would with the scaffold (sequential phosphorylation), but the scaffold itself is not modelled directly. Each kinase must be doubly phosphorylated before it can phosphorylate the next kinase in the cascade, and non-specific phosphorylases can de-phosphorylate the kinases at each reaction in the cascade. There was no experimental

Biochemical Event	Model Reaction
Ste20 binds to Ste11 #1	ste20 + mapkkk $\rightleftharpoons$ ste20-mapkkk
Ste20 phosphorylation of Ste11 #1	ste20-mapkkk $\rightarrow$ mapkkk-p + ste20
Ste20 binds to Ste11 #2	ste20 + mapkkk-p $\rightarrow$ ste20-mapkkk-p
Ste20 phosphorylation of Ste11 #2	ste20-mapkkk-p $\rightarrow$ mapkkk-pp + ste20
Ste11 binds to Ste7 #1	mapkkk-pp + mapkk $\rightleftharpoons$ mapkkk-pp-mapkk
Ste11 phosphorylates Ste7 #1	mapkkk-pp-mapkk $\rightarrow$ mapkkk-pp + mapkk-p
Ste11 binds to Ste7 #2	mapkkk-pp + mapkk-p $\rightleftharpoons$ mapkkk-pp-mapkk-p
Ste11 phosphorylates Ste7 #2	mapkkk-pp-mapkk-p $\rightarrow$ mapkkk-pp + mapkk-pp
Ste7 binds to Fus3 #1	mapkk-pp + mapk $\rightleftharpoons$ mapkk-pp-mapk
Ste7 phosphorylates Fus3 #1	mapkk-pp-mapk $\rightarrow$ mapkk-pp + mapk-p
Ste7 binds to Fus3 #2	mapkk-pp + mapk-p $\rightleftharpoons$ mapkk-pp-mapk-p
Ste7 phosphorylates Fus3 #2	mapkk-pp-mapk-p $\rightarrow$ mapkk-pp + mapk-pp
Fus3 up-regulates expression of msg5	mapk-pp $\rightarrow$ msg5 + mapk-pp
msg5 binds to Fus3 #1	msg5 + mapk-pp $\rightarrow$ msg5-mapk-pp
msg5 de-phosphorylates Fus3 #1	msg5-mapk-pp $\rightarrow$ ptc1 + mapk-p
msg5 binds to Fus3 #2	msg5 + mapk-p $\rightleftharpoons$ msg5-mapk-p
msg5 de-phosphorylates Fus3 #2	msg5-mapk-p $\rightarrow$ ptc1 + mapk
msg5 is degraded	msg5 $\rightarrow$
nsp binds to Ste7 #1	mapkk-pp + nsp $\rightleftharpoons$ mapkk-pp-nsp
nsp de-phosphorylates Ste7 #1	mapkk-pp-nsp $\rightarrow$ mapkk-p + nsp
nsp binds to Ste7 #2	mapkk-p + nsp $\rightleftharpoons$ mapkk-p-nsp
nsp de-phosphorylates Ste7 #2	mapkk-p-nsp $\rightarrow$ mapkk + nsp
nsp binds to Ste11 #1	mapkkk-pp + nsp $\rightleftharpoons$ mapkkk-pp-nsp
nsp de-phosphorylates Ste11 #1	mapkkk-pp-nsp $\rightarrow$ mapkkk-p + nsp
nsp binds to Ste11 #2	mapkkk-p + nsp $\rightleftharpoons$ mapkkk-p-nsp
nsp de-phosphorylate Ste11 #2	mapkkk-p-nsp $\rightarrow$ mapkkk + nsp
Phosphorylated Fus3 activates Ste12	mapk-pp + ste12 $\rightarrow$ mapk-pp + ste12-active
Ste12 de-activates	ste12-active $\rightarrow$ ste12

TABLE 5.2: Revised MAPK model reactions. The reactions are constructed for input into the Copasi software package.

Model Species	Particle Number	Source
Ste20	260	Ghaemmaghami S. [265]
Ste7	736	Slaughter B. [266]
Ste11	672	Ghaemmaghami S. [265]
Fus3	848	Slaughter B. [266]
non-specific phosphatase	1000	Ghaemmaghami S. [265]
msg5	1000	Maeder C. [267]
Ste12	1390	Ghaemmaghami S. [265]

TABLE 5.3: Parameter values for the revised MAPK model. Parameter values were calculated in number of molecules per cell, and input into the model as particle numbers in Copasi.

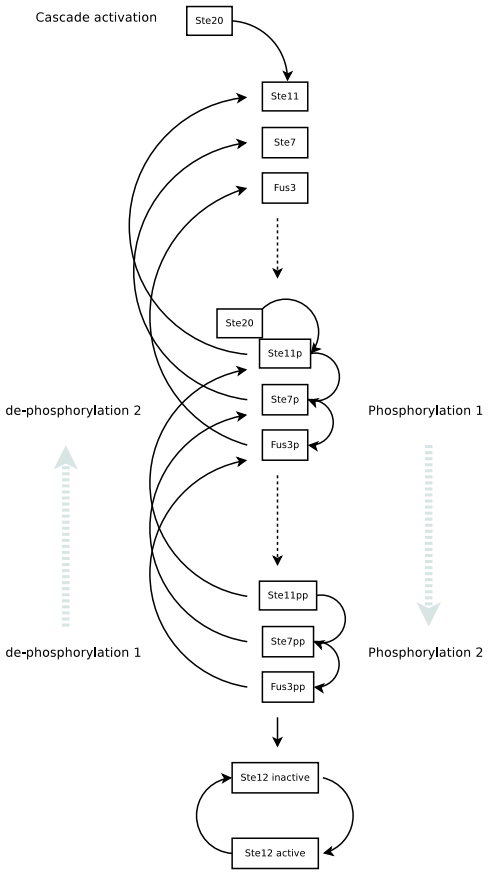

FIGURE 5.3: Diagrammatic representation of the revised MAPK model.

data to demonstrate the mechanism of the procession of phosphorylation states during the cascade, specifically whether the three kinases must bind as a complex and then sequentially phosphorylate each other, or whether each binds separately, and randomly. The Kofahl and Klipp model also follows this process of association and dissociation of the kinases, with the Ste5 scaffold, and the complex can dissociate completely at each phosphorylation step (figure 1.9). The revised model incorporated reversible binding between each kinase, and double phosphorylation events for each, which are believed to create the ultra-sensitive switch-like response, bistability, and hysteresis behaviours observed from the MAPK cascade [77]. The model also included each component binding to form a complex before being released in a phosphorylated or de-phosphorylated form (equation 5.3).

$$E_{i-1^*} + E_i \rightleftharpoons E_i E_{i-1^*} \rightarrow E_{i-1^*} + E_i^* \tag{5.3}$$

These events introduced the nested feedback loops (dephosphorylation of each kinase separately, and simultaneously with the cascade of phosphorylation reactions (figure 5.3)) modelled by Goldbeter and Markevich, and are also believed to be the feature of MAPK cascades that facilitates ultra-sensitivity and bistability [76, 77, 123].

The model therefore combines the complex formation approach of Kofahl and Klipp with generic MAPK cascade models such as Huang and Ferrell, and Kholodenko. [122, 124, 128, 131]. The dynamics of the G-protein cycle were not included in the model as it was believed the G-protein cycle would be at steady-state to initiate the phosphorylation of the MAPK cascade and would not have further downstream effects [56, 126].

5.2.1.1 Simulation Results

Time course simulations showed the MAP kinases are sequentially phosphorylated, followed by accumulation of active Ste12 within two hours, as observed experimentally in the yeast mating pathway (figures 5.4) [55, 195, 196]. A parameter scan of the input to the cascade (taken as Ste20 in the model) revealed the hyper-sensitivity characteristic of MAPK cascades. By varying the level of signal entering the cascade, the model demonstrated a threshold concentration of input producing a switch-like response in the output (figure 5.5 A and B). Figure 5.5 shows that introducing reversible complex formation between kinases and phosphatases produces the ultra-sensitive response, as shown in published models of signal cascades [77, 124, 269]. This behaviour could not be reproduced with the Kofahl and Klipp model. Moore published an observation that filamentous growth (the chemotropic response to pheromone) was not detected at pheromone concentrations less than 10nM, indicating that this is the activation threshold of the mating response [195]. A parameter scan of the input signal was run on the model and it was found that the new model demonstrates an ultra-sensitive switch-like response at a concentration of approximately 15nM activating signal (figure 5.5 B). The model is therefore demonstrating the switch-like behaviour replicated in previous published models of MAP kinase cascades [132, 270], and the switching point is at a biologically-relevant concentration of the activation signal [195].

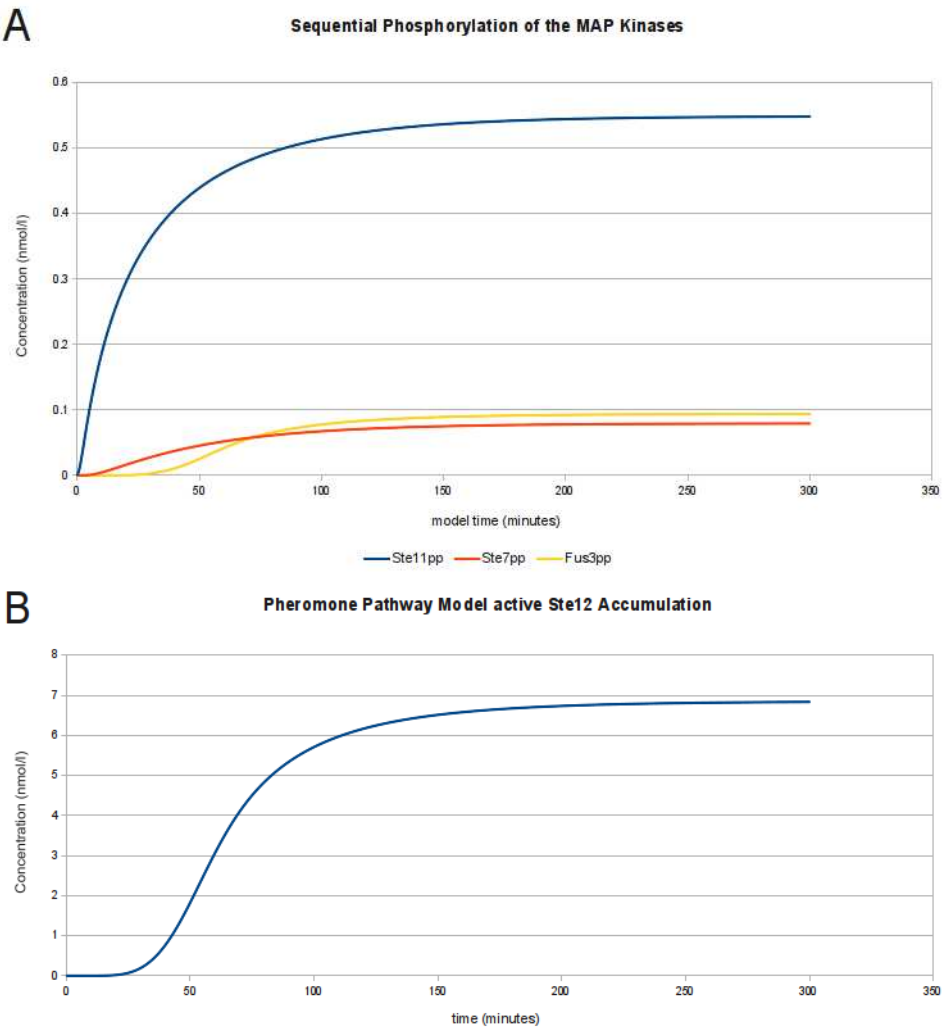

FIGURE 5.4: Time course simulation of the revised MAPK model. A. Accumulation of the double phosphorylated forms of the MAP kinases. B. Accumulation of active Ste12.

The revised MAPK model demonstrated the same characteristics as published simplified models of MAP kinase cascades [122, 124]. The model also reproduces experimental observations of the pheromone response of yeast in terms of the time scale, order of events, and activation threshold [227]. This revised MAPK model provides a more relevant chassis with which to study the effect of the dynamic behaviour of MAPK cascades on gene circuitry that are coupled to the yeast pheromone response pathway. The model requires additional reactions to incorporate the additional components of the pheromone response pathway, such as the G-protein cycle and the α-factor Ste2 receptor binding. The model also requires additional parameterisation to tune the ultra-sensitivity behaviour, as currently the rate of activation is very steep

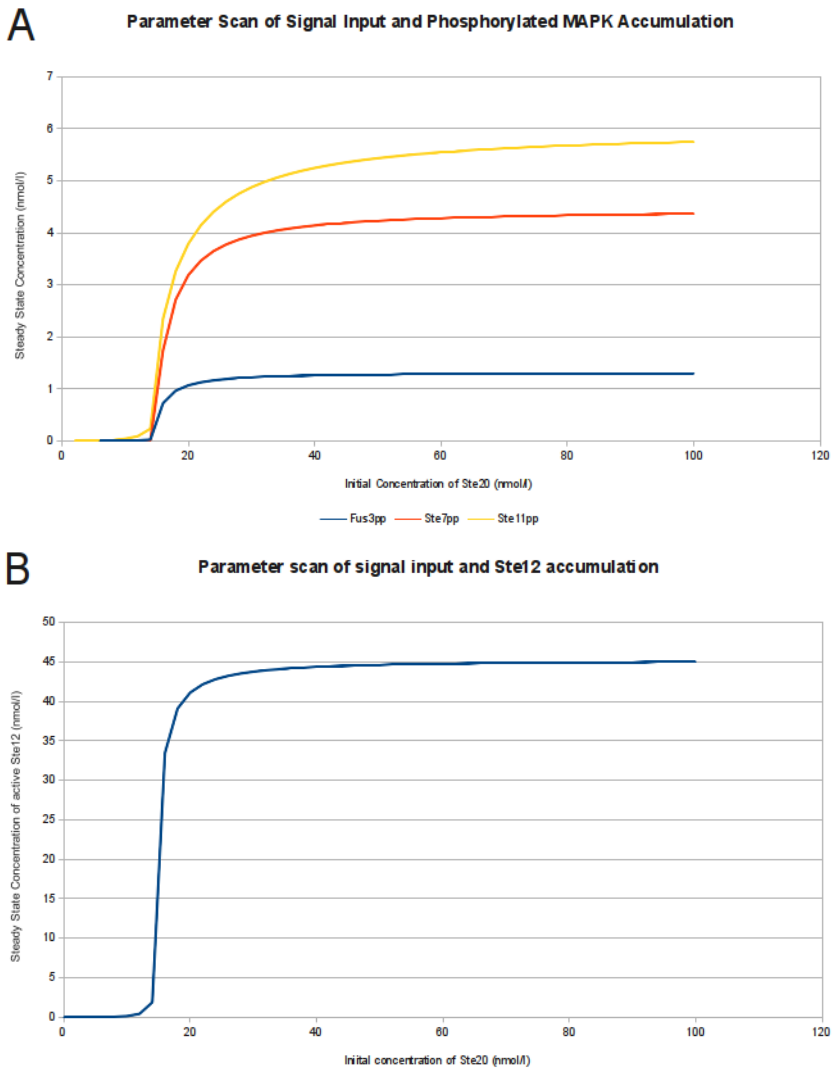

FIGURE 5.5: [Accumulation of the phosphorylated forms of the MAP kinases and active Ste12 in the revised MAPK model in response to varying levels of input. A. The accumulation of the double phosphorylated forms of the MAP kinases. B. The accumulation of the Fus3pp activated Ste12. The model demonstrates an ultra-sensitive response with a threshold activation concentration of approximately 15nM.

compared with published models of MAPK cascades (figure 5.5) ([133]). A rate law could be investigated in terms of using Michaelis-Menten type kinetics with a Hill coefficient instead of simple mass-action kinetics [122]. Also, future research in the McCarthy lab includes QconCAT mass spectrometry analysis of the yeast pheromone response pathway, therefore this model can be parameterised using quantitative data, and incorporated into this research program.

5.3 Modelling the Gene Circuit

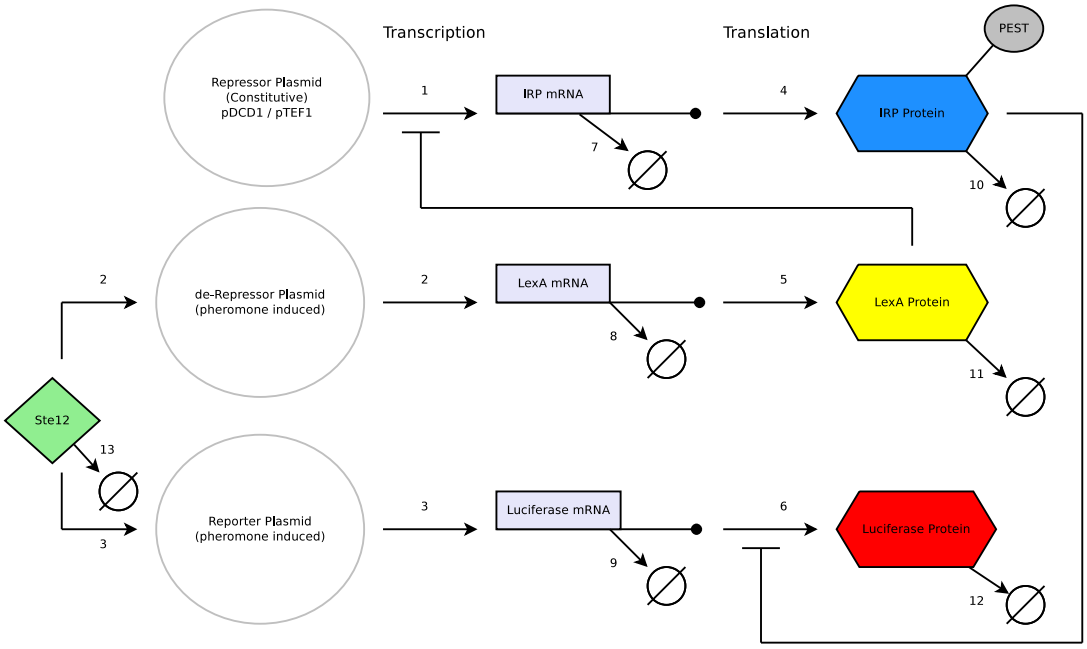

FIGURE 5.6: Schematic overview of the gene circuit model. Arrows represent reactions, and shapes represent species in the model. The plasmids were not present as species in the model, however rate equations represented the transcription process from the plasmid DNA to the mRNA species, which is then present as a species in the model. Reaction numbers are detailed in table 5.4.

A model of the gene circuit was devised that would simulate the interactions of the components, enabling investigation of the system both prior to, and during laboratory experimentation (figure 5.6). A mechanistic model was built using mass-action kinetics and Michaelis-Menten formalism, enabling representation of the activation of the pheromone-inducible genes and inhibition of transcription and translation by LexA and IRP respectively (table 5.4). Standard Michaelis-Menten equations were employed (equation 5.4) where v is the rate of the reaction, V_{max} is the maximum reaction rate, [s] is the substrate concentration, and K_M is the concentration of substrate that results in 50% of the maximum rate V_{max} [271]. Michaelis-Menten kinetics were used to model the processes of transcription and translation as these processes can saturate, and the rate equations provide the ability to tune the strength of the promoters through the V_{max} and K_M parameters [157, 190]. The circuit input signal, and degradation rates of protein and mRNA species were represented by mass-action kinetics as these processes are unlikely

No.	Biochemical Event	Model Reaction
1	IRP transcription (LexA inhibited)	$\rightarrow$ IRP-mRNA; LexA
2	LexA transcription (Ste12 activated)	$\rightarrow$ LexA-mRNA; Ste12
3	luciferase transcription (Ste12 activated)	$\rightarrow$ Lucif-mRNA; Ste12
4	IRP translation	IRP-mRNA $\rightarrow$ IRP + IRP-mRNA
5	LexA translation	LexA-mRNA $\rightarrow$ LexA + LexA-mRNA
6	luciferase translation (IRP inhibited)	Lucif-mRNA $\rightarrow$ Lucif + Lucif-mRNA; IRP
7	IRP mRNA degradation	IRP-mRNA $\rightarrow \oslash$
8	LexA mRNA degradation	LexA-mRNA $\rightarrow \oslash$
9	luciferase mRNA degradation	Lucif-mRNA $\rightarrow \oslash$
10	IRP degradation	IRP $\rightarrow \oslash$
11	LexA degradation	LexA $\rightarrow \oslash$
12	luciferase degradation	Lucif $\rightarrow \oslash$
13	Ste12 degradation	Ste12 $\rightarrow \oslash$

TABLE 5.4: Gene circuit model reactions. semi-colon denotes a modifier in the reaction that either activates or inhibits the reaction from occurring. The reactions are constructed for input into the Copasi software package. The reaction numbers relate to the reactions outlined in figure 5.6

to exhibit saturation phenomena under the model conditions.

$$V = \frac{V_{max} \cdot [s]}{K_M + [s]} \tag{5.4}$$

Standard Michaelis-Menten equations cannot account for basal "leak" in gene expression (genes expressed at a low level in the absence of specific activation) [157, 190, 272]. Therefore, gene expression for transcription was modelled using a modified form of the Michaelis-Menten equation, based on published work by Ajo-Franklin, and included the parameter S_0, representing low level of constitutive reaction flow, which in the circuit model would represent basal expression [190].

The rate equation in 5.5 was used, where v is the rate of transcription, S_0 is the rate of basal transcription, V_{max} is the maximum rate of transcription, *[A]* is the concentration of the activating signal, and K_M is the concentration of the activating signal that results in 50% of the maximum rate of transcription [157, 190]. The activating signal represents a generic term for transcription factor, polymerase, ribosomes and the DNA/RNA replication machinery of the cell.

$$v = S_0 + \frac{V_{max} \cdot [A]}{[A] + K_M} \tag{5.5}$$

Inhibition of transcription and translation was modelled using a further modified form of the Michaelis-Menten equation, incorporating a competitive inhibitor (equation 5.6), which is the same form as equation 5.4 with the addition of the concentration of the repressor, *[i]* and K_i which is the concentration of inhibitor that results in 50% inhibition of the reaction. For translation the parameter for basal expression S_0 was removed.

$$v = \frac{V_{max} \cdot [A]}{[A] + K_M \cdot (1 + \frac{[i]}{K_i})} \tag{5.6}$$

The Michaelis-Menten formalism enables a more biologically-relevant behaviour with inhibition, activation, and saturation of the reactions however, it is designed for enzyme-catalysed reactions and is derived from equation 5.7.

$$E + S \rightleftharpoons ES \rightarrow E + P \tag{5.7}$$

The Michaelis-Menten scheme in equation 5.7 assumes a low concentration of enzymes, a concentration of substrate that is not rate limiting, and that the enzyme-substrate complex forms much faster than the formation of product, and the reaction occurs in a well-mixed homogeneous suspension [273]. The processes of transcription and translation involve hundreds of components in huge complexes [274]. The dynamic behaviours of each of these interactions is over-simplified in the generic Michaelis-Menten rate equation. However, there are insufficient data available to provide accurate mathematical representation to incorporate these processes, and the Michaelis-Menten formalism provides a convenient framework to begin modelling biological interactions. It would also be impractical to build a model with rate equations incorporating all of the molecular interactions during DNA replication and protein synthesis.

The Michaelis-Menten formalism contains V_{max} and K_M parameters to describe the rate of an enzyme-catalysed reaction and the catalytic turnover rate [271]. These functions can be used in the modelling of synthetic gene circuits to simulate different strength promoters [275, 276]. V_{max} can be used to represent maximal promoter turnover time, and saturation concentration of RNA polymerase (i.e. the maximal rate at which RNA polymerases can be recruited, and transcription initiated). K_M can be used to represent the concentration of RNA polymerase molecules and/or relevant transcription

factors that produces 50% of maximal promoter output. The substrate can represent the components of the transcription initiation pathway; in the case of this project, the substrate was taken as the transcription factor, functioning as the activating signal to recruit an RNA polymerase to the promoter and initiate transcription. This interpretation necessarily assumes that the concentration of transcription factor is the limiting factor in transcription initiation. Therefore, the fraction *[A]* / K_M + *[A]* represents the proportion of promoters occupied by transcription-competent RNA polymerase for any given transcription-factor concentration, where *[A]* replaces the substrate in the classical Michaelis-Menten rate equation [276]. When *[A]* tends to infinity, the ratio approaches 1 and represents saturation of the reaction. When *[A]* tends to zero the equation can be expressed as $v = (V_{max}/K_M)\cdot [A]$, where the factor V_{max}/K_M represents the promoter strength [276]. Therefore, the three dependent parameters: V_{max}, K_M, and V_{max}/K_M are sufficient to compare promoter strength. The same Michaelis-Menten formalism can be used with translation, replacing the polymerase enzyme with the ribosome.

The circuit was initially modelled in isolation (without the interaction with the MAP kinase cascade) to enable observation of interactions between the components in response to externally-stipulated levels of stimulus. To this end, circuit activation was modelled using a generic "activation-species", which caused catalytic stimulation of the expression of luciferase and LexA mRNA species (equations 5.8, and 5.9).

As described above, the parameter "S_0" was added to the transcription equations to account for a continuous low (basal) level of transcription from the IRP, LexA and luciferase genes. There is no published data for the specific levels of basal expression of the genes expressed on the circuit plasmids, however research indicates the up-regulation is between 10 and 100-fold [277]. For the purposes of the model therefore, basal expression was set to 1% of V_{max}.

$$Luciferase\ transcription = S_0 + \frac{V_{max} \cdot [activator]}{K_M + [activator]} \tag{5.8}$$

$$LexA\ transcription = S_0 + \frac{V_{max} \cdot [activator]}{K_M + [activator]} \tag{5.9}$$

$$IRP\ transcription = S_0 + \frac{V_{max} \cdot [IRP\ activator]}{[IRP\ activator] + K_M \cdot (1 + \frac{[LexA]}{K_i})} \tag{5.10}$$

The parameter "activator" represents the output of the yeast pheromone response pathway and incorporates Ste12 transcription factor activation and RNA polymerase binding to the promoter. For IRP transcription, the formula incorporates competitive inhibition by the LexA repressor (equation 5.10) "IRP activator" represents the cellular signal that regulates constitutive expression of the *DCD1* or *TEF1* promoters, and provides a continuous activation of transcription via a model species with fixed concentration, therefore simulating constitutive gene expression. The denominator includes the classical Michaelis-Menten competitive inhibitor expression *[i]/[K_i]*, enabling increasing levels of LexA to increase inhibition on the transcription of the IRP, at a rate that can be tuned with the parameter K_i. Using this format enables LexA to compete with the cellular transcription machinery to suppress the production of IRP mRNA.

For translation, the same formulae are used. For luciferase translation, as with IRP transcription, a competitive inhibitor function is added to the Michaelis-Menten equation. As the concentration of IRP increases, it increases the inhibition of translation of luciferase mRNA to luciferase protein (equation 5.11).

$$Luciferase\ translation = \frac{V_{max} \cdot [Luciferase\ mRNA]}{[Luciferase\ mRNA] + K_M \cdot (1 + \frac{[IRP]}{K_i})} \tag{5.11}$$

LexA and IRP translation are represented with standard Michaelis-Menten equations with their respective mRNA species as substrate (equations 5.12, and 5.13).

$$LexA\ translation = \frac{V_{max} \cdot [LexA\ mRNA]}{K_M + [LexA\ mRNA]} \tag{5.12}$$

$$IRP\ translation = \frac{V_{max} \cdot [IRP\ mRNA]}{K_M + [IRP\ mRNA]} \tag{5.13}$$

The model was initially constructed with default values of 1 for all parameters, making all reaction rates equal throughout the model and enabling it to be interrogated in terms of just the interactions between the components. While this approach did not provide a quantitative model of the system (components cannot be quantified, nor the outputs

species	steady-state particle number
Luciferase	5
LexA	7
IRP	14

TABLE 5.5: Steady-state particle numbers of the circuit components in the absence of pheromone (the "OFF-state").

taken as accurate representations of what might be expected experimentally), the model did nevertheless simulate the *qualitative* interactions between the components in the model and therefore could be used to provide useful indications of how the interactions of the components might behave as a system *in vitro*. The steady-state values for the components were expressed as particle numbers, because concentration is irrelevant in the model at this time. The model was constructed to enable the investigation of the relative changes in the levels of the components, and the changes to the steady state of the system.

To begin, no input signal was added for up-regulation of the pheromone response and a default number of 1000 particles of the IRP signal was added to saturate the reaction for the constitutive expression of the repressor plasmid. This was defined as the circuit "OFF'-state". The steady of state of the system was determined and set as the initial conditions of the model to simulate the system under normal growth conditions, prior to circuit activation (table 5.5). The data indicated luciferase expression was lower than LexA as the IRP was repressing the translation of luciferase mRNA. The model predicted that IRP inhibits luciferase basal expression by approximately 30%. Addition of input signal for the pheromone-induced promoters resulted in the expression of the reporter gene, and the circuit transitioning to the "ON-state". The input signal was fixed at a constant, arbitrary value of 1000 to saturate the system (as with the IRP signal) and the steady-state level of the pheromone induced components LexA and luciferase was determined (table 5.6). Analysis of the steady-state particle numbers demonstrated the IRP inhibited luciferase expression by approximately 20% (taking LexA as 100% output). LexA also inhibited IRP, producing approximately 85% (15% inhibition) of the output observed in the OFF-state (table 5.5). From the observations of the model, it was implied that the components were interacting as expected, in terms of IRP inhibiting luciferase, and LexA inhibiting the IRP. The model next required parameterisation with

species	steady-state particle number
Luciferase	8
LexA	10
IRP	11

TABLE 5.6: steady-state particle numbers of the circuit components in the "ON-state".

biological-relevant values for each of the reactions. This would provide a more accurate model of the system, which could be compared with experimental observations.

5.4 Model Parameterisation

In order to begin to build an *in vivo* representative model of the circuit, the model was parameterised using published literature values for the rate of transcription and translation in *S. cerevisiae*, as well as rates of protein and mRNA degradation (table 5.7). The model was run to steady-state with no signal, the model state set to the initial state, signal added, and sensitivity analysis performed on the model to investigate control points that could provide the ability to tune the circuit *in vitro*. The results are

Parameter	value	source
Transcription	800bp/min	Zenklusen, D. [278]
Translation	9.3 codons/sec	Bonven B. [279]
average mRNA half-life	23 minutes	Wang Y. [280]
average protein half-life	43 minutes	Belle A. [281]

TABLE 5.7: Table of generic parameter values from the published literature, used for the first round of parameterisation of the gene circuit model.

summarized in table 5.8. Sensitivity analysis indicated the steady-state of the system was sensitive to the IRP transcription rate, as this directly influences the level of IRP in the system (table 5.8). This raised the concern that IRP protein concentration above a certain threshold would effectively sequester even induced luciferase mRNA, and thus prevent the circuit switching to the "ON"-state. Sensitivity analysis provided the hypothesis that implementing varying rates of transcription of the IRP gene would enable tuning the circuit, in terms of the level of IRP protein abundance, and therefore the level of luciferase expression. A strong and weak promoter for IRP would avoid the risk of the IRP repressing the luciferase reporter gene to such an extent that it was not

Parameter	Sensitivity Value
IRP basal transcription	$1.6x10^{10}$
IRP transcription V_{max}	3313
IRP transcription K_M	4281
IRP transcription K_i	227
luciferase basal transcription	2214
luciferase transcription V_{max}	935
luciferase transcription K_M	1422
LexA basal transcription	2832.86
LexA transcription V_{max}	1887
LexA transcription K_M	62
IRP translation K_M	1573
IRP translation V_{max}	1305
luciferase translation K_M	1937
luciferase translation V_{max}	104
luciferase translation K_i	585
LexA translation K_M	154
LexA translation V_{max}	1202
IRP mRNA degradation	2593
LexA mRNA degradation	4262
luciferase mRNA degradation	48
IRP degradation	106
LexA degradation	1171
luciferase degradation	1210
signal degradation	2667

TABLE 5.8: Sensitivity analysis of the gene circuit model performed on all parameter values, as a function of all non-constant concentrations of species. Larger values indicate that perturbing the parameter will have a larger effect on the steady-state of the components of the system.

detectable, requiring the re-designing of the circuit later in the project. Therefore, the strong *TEF1* and the weaker *DCD1* yeast promoters were included in the experimental phase (chapter 4) as promoters for the repressor plasmid to investigate the effect of high and low abundance of IRP on the expression of luciferase.

The interaction of the repressors (the K_i values) did not score highly in the metabolic control analysis, indicating that they were not strong control elements of the system. This indicated the model is not yet accurately simulating the interaction of the IRP and LexA repressors. However, the luciferase translation K_i sensitivity was 585, and the IRP transcription K_i is 227. This indicated that the system was approximately twice as sensitive to translational inhibition compared with transcriptional inhibition. It was

hypothesized that this was representative of the difference in the time scales for the reactions, as inhibition of translation results in a change in protein abundance more rapidly than inhibiting transcription, which would not have an effect on downstream protein abundance until all of the pre-existing mRNA template had been degraded. This observation, together with the high sensitivity to IRP transcription indicated the framework of the model was approximating the system.

In addition to sensitivity analysis, metabolic control analysis was also used to investigate the model (table 5.9). The metabolic control coefficient could be represented as a stacked bar chart enabling graphical interpretation of the data. The figures 5.7 (page 148), 5.13 (page 156), and 5.14 (page 157), attempt to represent the concentration control coefficients of the models. A positive control coefficient value indicates that a reaction has a positive effect on the steady state concentration of a reactant, and a negative control coefficient indicates a negative effect on the steady state concentration of a reactant in the model. Larger values indicate a larger effect. For example, in figure 5.13 the largest positive control coefficient for luciferase is the IRP degradation reaction. Increasing the rate of IRP degradation is predicted by the model to effect the largest increase in the steady-state concentration of luciferase, compared with the other reactions of the model. The concentration control coefficient for metabolic control analysis are synonymous with sensitivity analysis [149] and confirmed that luciferase expression could be increased by increasing the IRP degradation rate and decreasing the IRP transcription rate. The model therefore provides valuable insight into the controlling interactions within the network even in the absence of experimental data, and provides useful indicators to refine or re-design the circuit before *in vitro* construction.

The circuit was designed prior to the project to maximize the ratio of the expression of the reporter gene between the OFF and the ON-state, through lowering of the basal expression level. The original strategy for circuit design included IRP regulation as a means to maximize the induced/uninduced ratio of reporter gene expression. The model indicated that, while the IRP repressor does indeed inhibit the basal level of output in the OFF-state, the level of output in the ON-state is similarly inhibited, and indeed to a similar extent. This therefore reduces the overall system output and produces an

	luciferase mRNA	LexA mRNA	IRP mRNA	LexA	luciferase	IRP
luciferase transcription	1.140	0	0	0	0.880	0
LexA transcription	-0.012	0.952	-0.039	-0.039	0.002	-0.039
IRP transcription	0.302	0	1	0	-0.038	1
LexA translation	-0.007	-0.417	-0.024	0.583	0.001	-0.024
IRP translation	0.151	0	-0.5	0	-0.019	0.5
luciferase translation	-1.114	0	0	0	0.140	0
IRP degradation	-0.302	0	0	0	0.038	-1
luciferase degradation	0	0	0	0	-1	0
LexA degradation	0.013	0	0.041	-1	-0.002	0.041
luciferase basal transcription	0.114	0	0	0	0.088	0
LexA basal transcription	-0.001	0.048	-0.002	0.048	7.5e-05	-0.002
LexA mRNA degradation	0.007	-0.583	0.024	-0.583	-0.001	0.024
luciferase mRNA degradation	-0.140	0	0	0	-0.108	0
IRP mRNA degradation	-0.151	0	-0.5	0	0.019	-0.5

TABLE 5.9: Metabolic control analysis of the gene circuit model, used to construct figure 5.7.

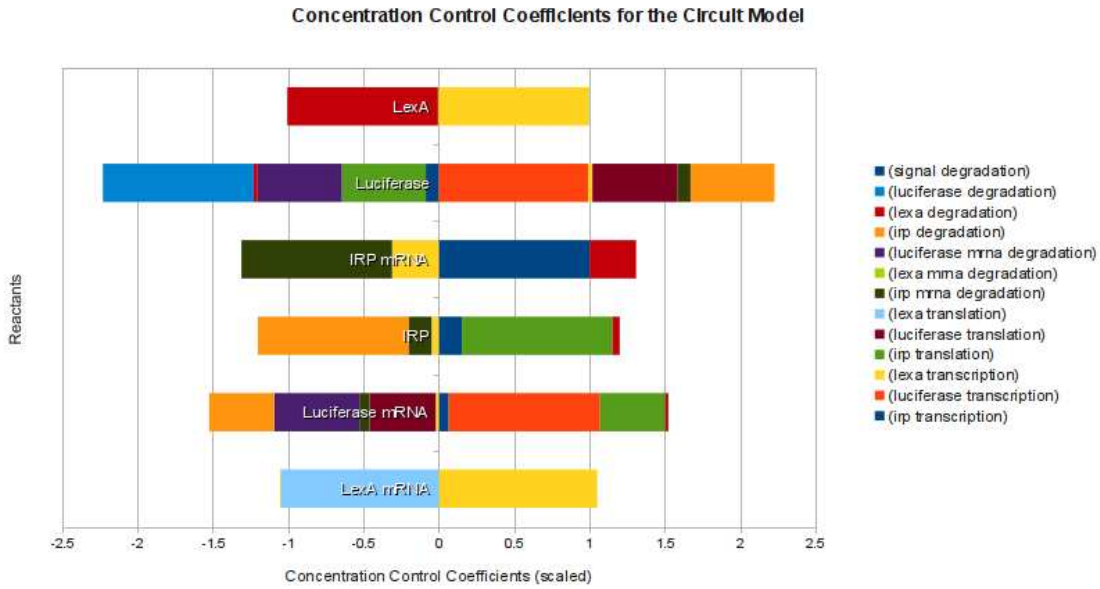

FIGURE 5.7: The concentration control coefficients for the gene circuit model. The graph shows the reactants in the model on the x axis, and the y axis shows the concentration control coefficients for each of the reactions in terms of the control exerted by the concentration of the reactants. The concentration control coefficient illustrates the effect of perturbing the rate of a reactions on the steady state concentration of a reactant in the network. A negative control coefficient has a negative effect on the reactant steady state concentration, and a positive control coefficient a positive effect.

essentially unchanged ratio of induced/uninduced expression. The model also predicts that the inhibitor must be rapidly removed from the system, to enable induced output to reach levels typically observed in a simple "direct" pheromone-induced reporter system. To address these concerns raised by the model, it was decided to investigate the effects of altering the half-life of the IRP protein and the strength of the constitutive promoter in the *in vivo* system. To this end, "strong" and "weak" constitutive promoters (P_{TEF1}, and P_{DCD1}, respectively) were implemented in the repressor plasmid, and a modified IRP with a shorter half-life (IRP_{PEST}) was constructed (chapter 3).

5.4.1 Further Parameterisation and the Final Model

Further parameterisation was performed to improve on the generic parameter values, using a literature search of published kinetic parameters. Kinetic data for all of the model reactions was not available prior to the project however, text mining can provide

a range of existing parameter values from the published literature and is used routinely in modelling biological systems [282, 283] (table 5.10). A generic rate of transcription in *S. cerevisiae* was taken as 800 base pairs per minute [278], and for translation; 9 codons per second [279]. Based on the length of the DNA and mRNA transcript of the component, the parameters for the gene circuit model were calculated as in table 5.10. The circuit model was further modified to reflect the laboratory implementation of the

Parameter	value	source
Luciferase mRNA transcription	2 minutes	Zenklusen *et al* [278]
LexA mRNA transcription	46 seconds	Zenklusen *et al* [278]
IRP mRNA transcription	3.3 minutes	Zenklusen *et al* [278]
Luciferase protein translation	2 minutes	Bonven B. [279]
LexA protein translation	20 seconds	Bonven B. [279]
IRP protein translation	1.6 minutes	Bonven B. [279]
LexA protein half-life	1 hour	Sassanfar M. [284]
IRP protein half-life	16 hours	Clarke S. [252]
Luciferase protein half-life	3 hours	Leclerc G. *et al* [285]
Ste12 numbers per cell	1390	Ghaemmaghami S. [265]
Ste12 protein half-life	25 minutes	Esch R. [286]
Fold change for gene up-regulation	10-100x	Buchler N. [277]

TABLE 5.10: Gene circuit parameter values sourced from published literature. Specific parameter values were calculated for each of the circuit components, improving on the generic parameter values.

strong (*TEF1*) and weak (*DCD1*) promoters, prompted by the initial metabolic control analysis results. The IRP transcription reaction was retained as a single rate law for both promoters, and an additional parameter was added to incorporate the change (equation 5.14) where p represents "promoter strength".

Copasi enables the use of "global quantities" that can be assigned values or custom mathematical expressions. This feature enabled the promoter strength to be modified globally, thus acting as an additional parameter value available for parameter scanning, optimization, or sensitivity analysis. Previous work on luciferase expression by Naglis Malys in the McCarthy lab (unpublished) gave estimates for *DCD1* promoter activity that were approximately 60 times lower than the *TEF1* promoter, therefore the rate equation maximum rate (V_{max}) and basal expression level (s) were reduced 60-fold to simulate the transcription rate of the weaker promoter. For P_{TEF1}, p was set to 1, and for P_{DCD1} p was set to 60. The global quantity p in equation 5.14 therefore represents the

term 1/promoter strength. Higher values of p represent weaker promoters, and lower values of p, stronger promoters.

$$IRP\ transcription = \frac{s}{p} + \frac{\frac{V_{max}}{p} \cdot [IRP\ activator]}{[IRP\ activator] + K_M \cdot (1 + \frac{[LexA]}{K_i})} \tag{5.14}$$

5.5 Stochastic Simulation of the Gene Circuit

Deterministic modelling of the gene circuit with coupled ordinary differential equations provided an average measurement of the behaviour of the circuit over time. However, coupled systems of chemical equations can also be represented as a stochastic process where the variables are numbers of molecules, and interactions are modelled as discrete events [287]. Deterministic modelling using ordinary differential equations to model biological systems assumes there are millions of molecules in a well mixed, homogeneous suspensions, that have an equal probability of interacting at any particular time.

While this is often sufficient for simulating simple *in vitro* biological behaviour, it is less appropriate for modelling complex, compartmentalised intracellular systems, particularly those involving interactions between small numbers of molecules [288, 289]. Stochastic modelling allows for random fluctuations in the interactions of small numbers of molecules, thus providing a more accurate simulation of the cellular environment [136]. The disadvantage in the use of stochastic simulations is that such models require a greater amount of computing time, as the interaction of every molecule at every time step must be calculated [288, 290]. Larger networks of interactions, such as MAP kinase cascades or whole genome models require large clusters and parallel computing to simulate short time courses [291].

The Gibson and Bruck stochastic algorithm essentially consists of an initialization step, a Monte Carlo step, and a update step [292]. The initialization step involves the recording of the number of molecules in the system, the reaction constants, and initialization (i.e. seeding) of the random number generators. The Monte Carlo step then generates two random numbers: these are used to determine the next reaction to

occur and the time interval, respectively. As the probability of a given reaction occurring is proportional to the number of substrate molecules, one can use the initial conditions to both determine the relative probabilities of *all* possible reactions, and to determine the probability of *any* given reaction occurring within a given time period. By combining these probabilities with random numbers, one can thus (via the first random number) determine *which* reaction occurred and (via the second random number) how much time elapsed before that reaction occurred. The update step simply updates the molecule numbers of the system to reflect the results of the reaction determined in the Monte Carlo step, and increases the time elapsed by the time step similarly determined. The process then repeats until the number of reactants is zero, or the maximum time for the simulation has been exceeded [287, 293]. The Gibson and Bruck stochastic algorithm is implemented in Copasi and therefore could be applied to the circuit model without additional software or programming [134].

Initial simulations of the system with the P_{DCD1} promoter exhibited transcriptional and translational bursting phenomena, due to the lower promoter strength and therefore IRP abundance. [245, 294]. This stochastic behaviour cannot be captured by deterministic solvers, as the behaviour is averaged over time [288]. Closer examination of IRP mRNA output by the *DCD1* promoter in the stochastic model revealed expression levels of the order of only 1 to 2 molecules per cell (figure 5.8), which is consistent with published literature on the promoter strength of P_{DCD1} [212, 242]. Such low levels of mRNA production would be expected to result in bursting behaviour in transcription and translation, as demonstrated in published models [294, 295]. Moreover, this behaviour would only be detectable in stochastic simulations, where molecules are treated as discrete entities rather than concentrations. This demonstrates the power of the stochastic approach, and this model therefore presented us with the opportunity to study noise in the circuit output using different strength promoters to control the expression level of the IRP repressor [288]. Stochastic trajectories represent a single set of probabilities for the change in the state of the system over time: as a crude approximation, they model a single possible response of a single cell. Therefore simulations were repeated 1000 times using the Copasi software, in order to generate a representative population. Optimally, many hundreds of thousands of trajectories

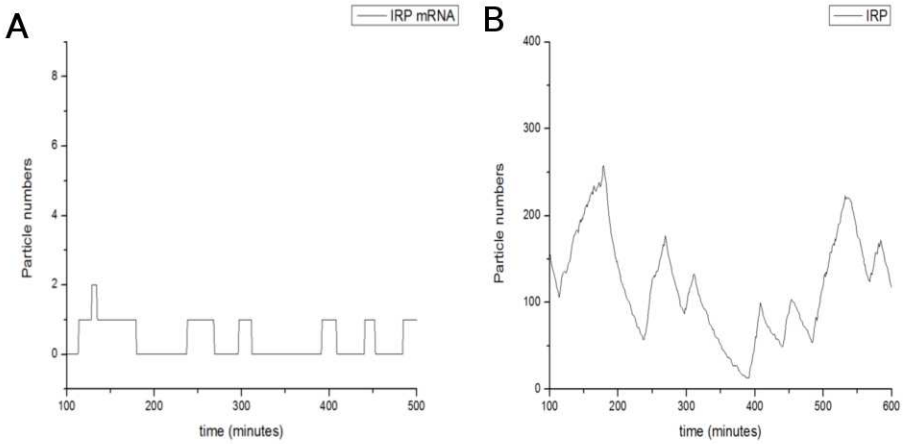

FIGURE 5.8: A single stochastic trajectory of the circuit model with the weaker *DCD1* promoter. A. The data demonstrate low levels of mRNA transcription, with one or two molecules being produced. B. Corresponding IRP protein translation from IRP mRNA. The protein numbers burst simultaneously with the production of mRNA.

should be run to obtain as close to the true mean of the population as possible. Copasi, however, outputs the raw data from each stochastic trajectory sequentially: manual processing of the sizeable data sets thus generated rapidly becomes a non-viable approach. To address this issue, a Python script was created that can calculate a global mean and standard deviation for the data produced by repeated simulations (see appendix B), enabling a greater number of repetitions, and automatic data handling and processing. Taking the mean of a large sample set essentially approximates the stochastic simulation to the deterministic simulation. The standard deviation, however incorporates the non-linear behaviour of the reactions, enabling observation of noise [288].

The model was run to steady-state to establish resting concentrations of all the components of the model. These values were then used to set the initial conditions for the simulation. 1000 replicate simulations were performed over a 600 minute period to correlate with the experimental data, and the results collated to generate an overall representation of the time course response of the model (the expression of the circuit components for each experiment therefore is plotted as a graph where n=1,000).

Data on the change in the number of protein species over time was collected from a model using parameters set to simulate the weaker *DCD1* promoter (figure 5.9), and a second model constructed, simulating the circuit with the stronger *TEF1* promoter

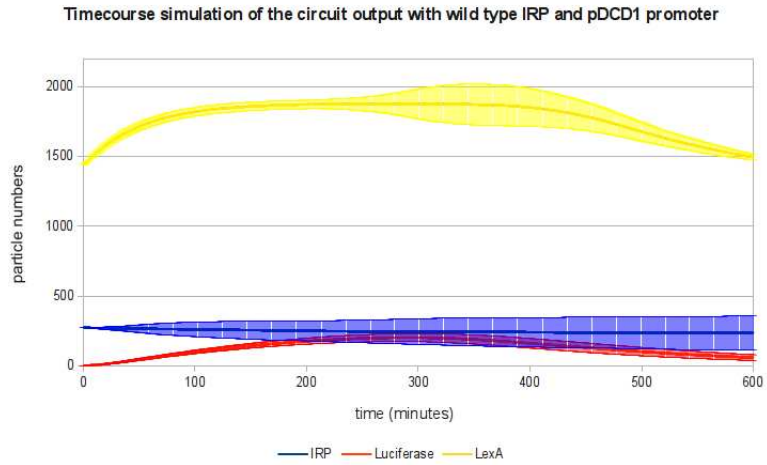

FIGURE 5.9: Stochastic simulation of the P_{DCD1} model showing the change over time of the protein components of the model.

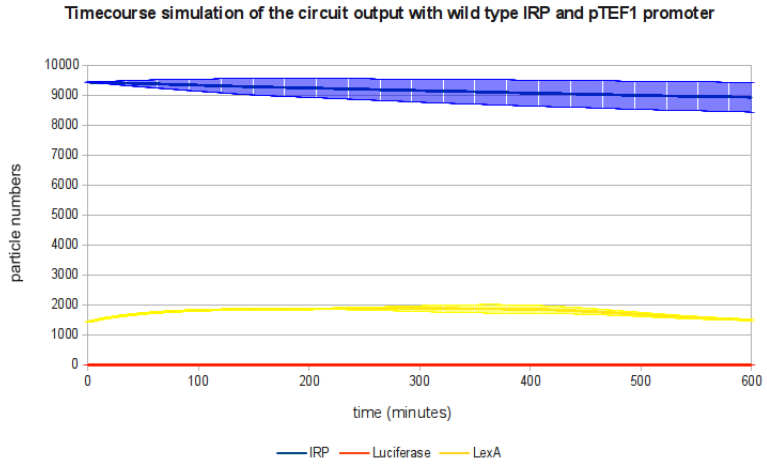

FIGURE 5.10: Stochastic simulation of the P_{TEF1} model showing the change over time of the protein components of the model.

(figure 5.10). From the time course simulations of the P_{DCD1} model, the IRP protein level (figure 5.9, blue) decreases slightly over time as the LexA protein (figure 5.9 yellow) increases during the time course and represses the expression of IRP. From figure 5.10 it can be seen that the luciferase expression level for the P_{TEF1} model does not increase significantly, compared with the P_{DCD1} model. The data suggested that the LexA repressor would not be capable of repressing the IRP in the P_{TEF1} model, and this prediction was confirmed by the experimental data (chapter 4).

The luciferase protein expression levels from the models are much lower than IRP and LexA and cannot be interpreted from figures 5.9 and 5.10, therefore they were plotted separately in figure 5.11. The luciferase expression level from the P_{DCD1} model

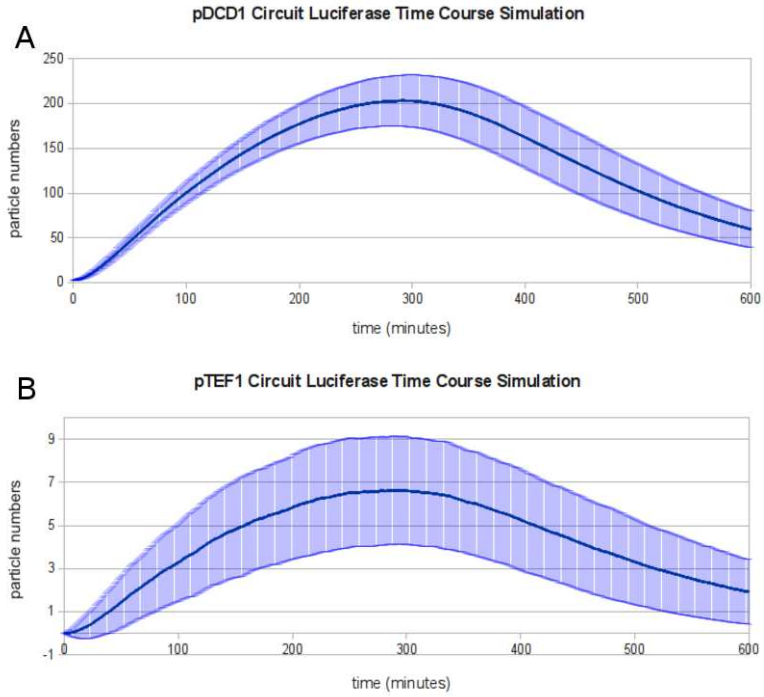

FIGURE 5.11: Stochastic simulation of the luciferase time course output of the P_{DCD1} and P_{TEF1} models.

was significantly higher than the P_{TEF1} model, by a factor of almost 25-fold. This observation was as expected as the stronger P_{TEF1} promoter would generate higher levels of IRP repressor. As might be expected, lower particle numbers appear strongly correlated with an increase in overall noise for that particular component of the model. Signal to noise ratios (SNR) were calculated for each circuit (figure 5.12) , confirming that this ratio indeed decreases as particle numbers decrease. The LexA SNR remained relatively constant through the simulation whereas the luciferase and IRP SNR decrease, indicating the circuit output becomes more noisy over time. It was noted from this data that experimental observations of luciferase expression could be subject to a high degree of noise due to low overall luciferase particle number, particularly in the P_{TEF1} strain with its high IRP expression, where luciferase levels my be low or even undetectable in this circuit.

Both of the parameterised models show the highest metabolic control coefficient for luciferase expression is from the degradation rate of the IRP (figure 5.13, table 5.11, and figure 5.14, table 5.12). No published studies have directly determined the half-life of the IRP protein, however there is data to indicate the value exceeds 16hrs [228, 252],

	luciferase mRNA	LexA mRNA	IRP	IRP mRNA	LexA	luciferase
IRP transcription	0.017	0	0.914	1	0	-0.897
luciferase transcription	1	0	0	0	0	1
LexA transcription	-0.003	1	-0.174	-0.191	0.235	0.171
IRP translation	0.019	0	1	0	0	-0.981
luciferase translation	-0.019	0	0	0	0	0.981
LexA translation	-0.014	0	-0.741	-0.810	1	0.726
IRP mRNA degradation	-0.017	0	-0.914	-1	0	0.897
LexA mRNA degradation	0.003	-1	0.174	0.191	-0.235	-0.171
luciferase mRNA degradation	-0.981	0	0	0	0	-0.981
IRP degradation	-0.019	0	-1	0	0	0.981
LexA degradation	0.014	0	0.741	0.810	-1	-0.726
luciferase degradation	0	0	0	0	0	-1
signal degradation	0	0	0	0	0	0

TABLE 5.11: Metabolic control analysis of the DCD1 based gene circuit model, used to construct figure 5.13.

	luciferase mRNA	LexA mRNA	IRP	IRP mRNA	LexA	luciferase
IRP transcription	0.017	0	0.914	1	0	-0.897
luciferase transcription	1	0	0	0	0	1
LexA transcription	-0.003	1	-0.174	-0.191	0.235	0.171
IRP translation	0.019	0	1	0	0	-0.981
luciferase translation	-0.019	0	0	0	0	0.981
LexA translation	-0.014	0	-0.741	-0.810	1	0.726
IRP mRNA degradation	-0.017	0	-0.914	-1	0	0.897
LexA mRNA degradation	0.003	-1	0.174	0.191	-0.235	-0.171
luciferase mRNA degradation	-0.981	0	0	0	0	-0.981
IRP degradation	-0.019	0	-1	0	0	0.981
LexA degradation	0.014	0	0.741	0.810	-1	-0.726
luciferase degradation	0	0	0	0	0	-1
signal degradation	0	0	0	0	0	0

TABLE 5.12: Metabolic control analysis of the TEF1 based gene circuit model, used to construct figure 5.14.

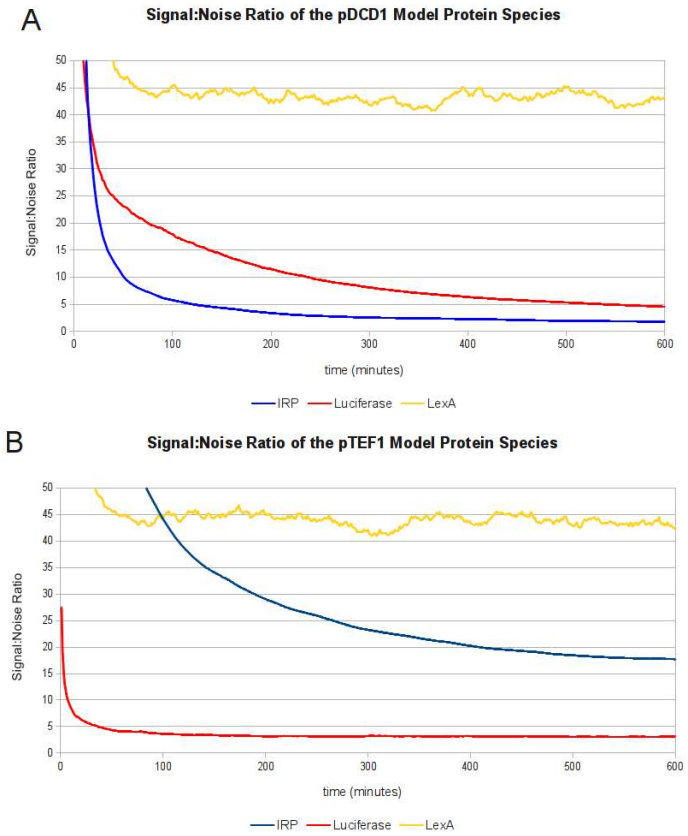

FIGURE 5.12: Signal to noise ratio for the P_{DCD1} and P_{TEF1} models. A) Is the signal to noise ratio for the P_{DCD1} model, and B) is the signal to noise ratio for the P_{TEF1} model. Graphs show the signal to noise ratio of the luciferase, LexA, and IRP protein components of the circuit during the time course simulation.

a time scale significantly longer than the total circuit response duration. As such, LexA-mediated reduction in IRP expression would be unlikely to contribute significantly to the *in vivo* circuit.

It was hypothesized from the model data that modifying the half-life of the IRP could increase expression level of luciferase from the circuit. The data prompted the investigation of the development of a Cln2 PEST degradation tagged IRP in the repressor plasmid (chapter 3.4.3), and the model adjusted to incorporate the predicted shorter half-life of 30 minutes [216].

At this point in the project there were four distinct versions of the circuit: P_{DCD1}, P_{TEF1}, P_{DCD1}-PEST, and P_{TEF1}-PEST, each requiring their own model parameters (table 3.1, chapter 3.6). The model parameter values and rate equations were identical apart from the promoter strength global parameter value and the IRP degradation rate. For the

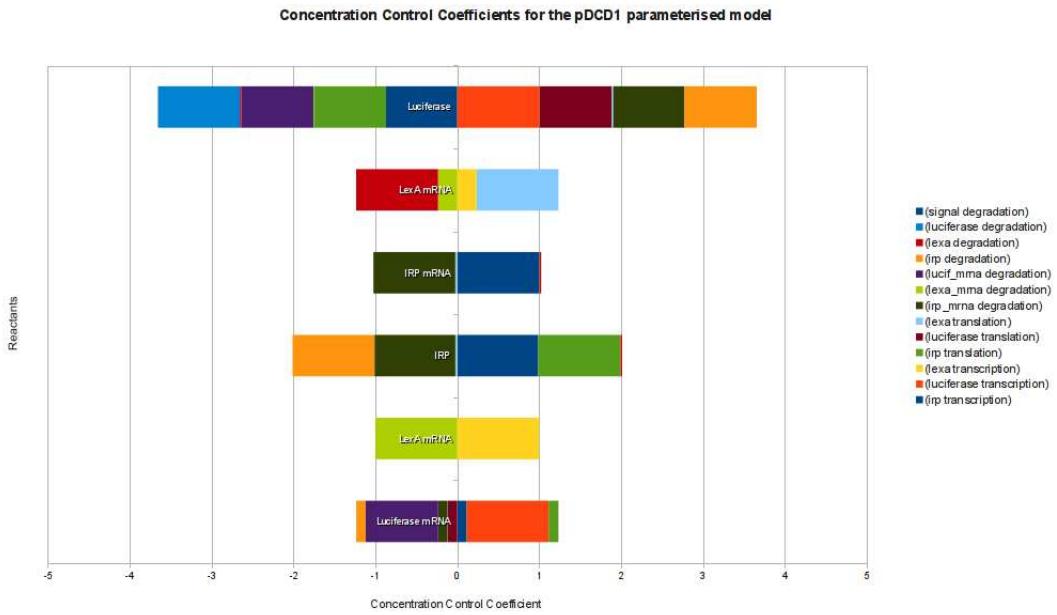

FIGURE 5.13: Metabolic control analysis results for the P_{DCD1} parameterised model. The graph shows the reactants in the model on the x axis, and the y axis shows the concentration control coefficients for each of the reactions in terms of the control exerted by the concentration of the reactants. The concentration control coefficient illustrates the effect of perturbing the rate of a reactions on the steady state concentration of a reactant in the network. A negative control coefficient has a negative effect on the reactant steady state concentration, and a positive control coefficient a positive effect.

P_{DCD1} and P_{TEF1} circuits, the IRP degradation rate was based on the published IRP half life of approximately 16 hours, which was input as 0.001 particles per minute (at this stage in the development of the model, dilution of the mRNA and protein species by cell doubling was not incorporated into the rate laws.). For the short half-life IRP_{PEST} models the IRP degradation rate was increased by 10 fold to 0.01 particles per minute, based on the estimated reduction in half-life of eGFP by Mateus and Avery [216].

Time course simulations were performed on the P_{DCD1}-PEST and P_{TEF1}-PEST models to investigate the effect of the IRP_{PEST} species on the luciferase expression level in the circuit (figures 5.15 and 5.16). Time course simulation of the P_{DCD1}-PEST model revealed a higher level of luciferase compared with earlier simulations with the wild-type half-life IRP (figure 5.9). IRP levels were maintained at a lower level, enabling the up-regulation of luciferase during circuit activation. The circuit required the same time to reach maximal luciferase expression, at around 300 minutes. The model did not predict shortening the half-life of the IRP would increase the rate of expression of

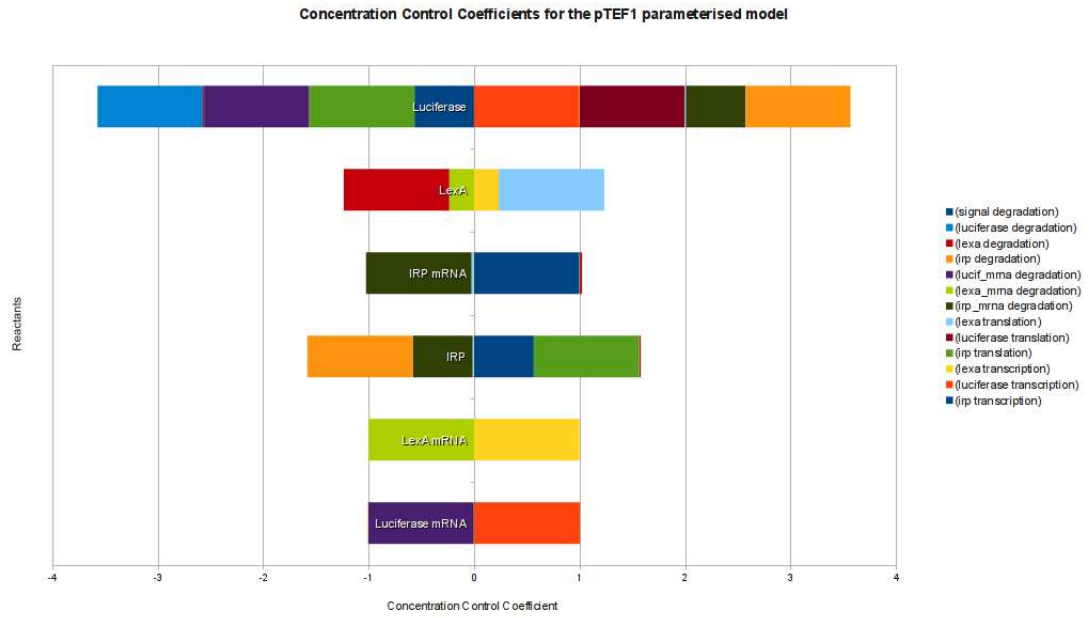

FIGURE 5.14: Metabolic control analysis results for the P_{TEF1} parameterised model. The graph shows the reactants in the model on the x axis, and the y axis shows the concentration control coefficients for each of the reactions in terms of the control exerted by the concentration of the reactants. The concentration control coefficient illustrates the effect of perturbing the rate of a reactions on the steady state concentration of a reactant in the network. A negative control coefficient has a negative effect on the reactant steady state concentration, and a positive control coefficient a positive effect.

luciferase for the P_{DCD1}-PEST circuit. The simulations were repeated with the P_{TEF1}-PEST model (figure 5.16). The data showed a lower expression level of the IRP and a higher expression level of luciferase compared with the simulations of the wild-type IRP (figure 5.10). Simulations with the stronger *TEF1* promoter still showed a reduction in luciferase expression compared with the weaker promoter in the P_{DCD1} model, indicating the abundance of IRP in the cell, despite the short half-life was still sufficiently high to negate repression by LexA (figure 5.11). The luciferase expression levels were re-plotted for the P_{DCD1}-PEST and P_{TEF1}-PEST models for comparison (figure 5.17 A and B). The simulations showed approximately the same rate in the increase of luciferase expression for both models, reaching maximum expression after approximately 300 minutes. Comparison with the luciferase expression levels from the wild type IRP models (P_{DCD1} and P_{TEF1}) (figure 5.11) showed a significantly increased level of luciferase for both PEST models, due to the lower expression levels of IRP_{PEST}.

The P_{DCD1}-PEST and P_{TEF1}-PEST models result in a higher expression level of

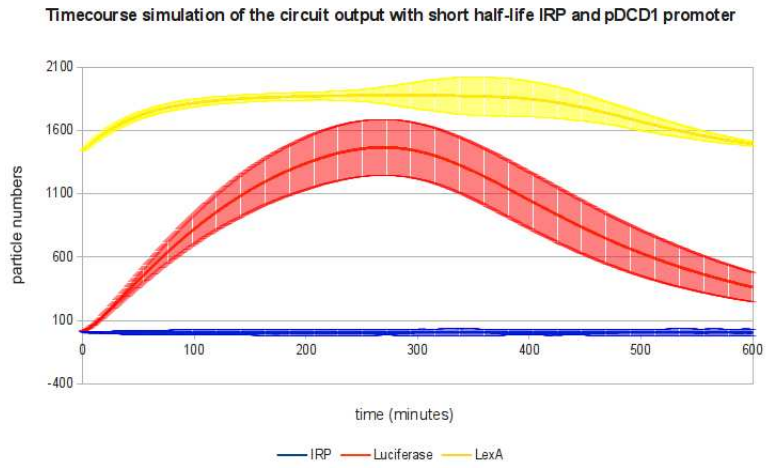

FIGURE 5.15: Time course simulation of the gene circuit with the short half-life IRP_{PEST} and the *DCD1* promoter. Protein species are plotted against time.

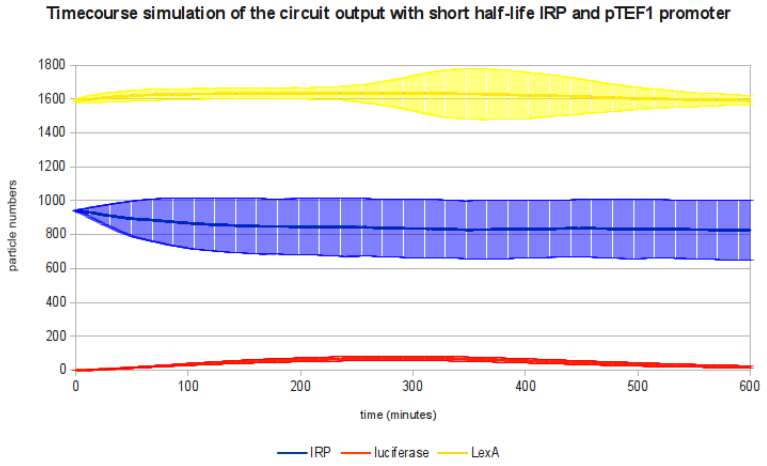

FIGURE 5.16: Time course simulation of the gene circuit with IRP_{PEST} and the P_{TEF1} promoter. Protein species are plotted against time.

luciferase, and lower expression level of IRP, as expected from the reduction in the half-life of the IRP protein. The shorter half-life IRP species however remains at a sufficiently high expression level negate the repression of IRP transcription by LexA.

The SNR data for the P_{DCD1}-PEST and P_{TEF1}-PEST models (figure 5.18) demonstrated a significantly lower SNR compared with the P_{DCD1} and P_{TEF1} models for the IRP protein species (figure 5.12). The expression of IRP_{PEST} is more noisy than the wild-type IRP. It was noted that the SNR for luciferase did not change significantly for the P_{TEF1}-PEST model compared with the P_{TEF1} model, however this is most likely be due to the low expression of luciferase in both of these models, despite the shorter half-life IRP (figure 5.17). The P_{DCD1}-PEST luciferase SNR dropped more rapidly

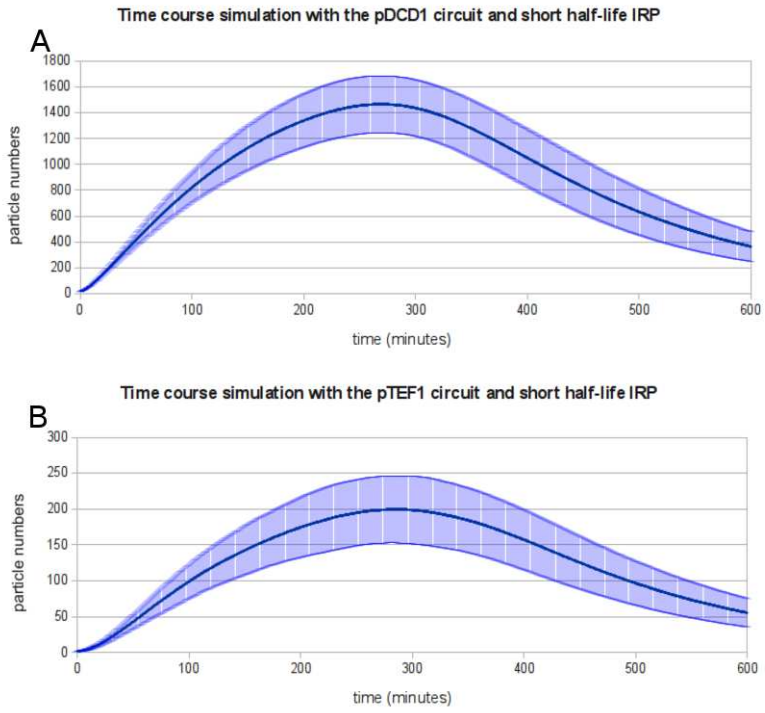

FIGURE 5.17: Time course simulation of the gene circuit P_{DCD1} and P_{TEF1} simulated promoters and IRP_{PEST}. A. Luciferase expression from the P_{DCD1}-PEST model. B. Luciferase expression from the P_{TEF1}-PEST model.

than the P_{DCD1} circuit in the first 100 minutes of the simulation, despite the higher particle numbers of luciferase from the reduced repression by IRP (which would be expected to increase the SNR). The degradation rate of the IRP protein with the PEST degradation tag was speculative and required further validation with experimental data to provide an accurate kinetic parameter for the model. However, The lower SNR in the PEST models indicates the shorter half-life IRP modification to the PEST circuits may introduce additional noise into the expression of luciferase, possibly through increased noise in expression of the IRP repressor.

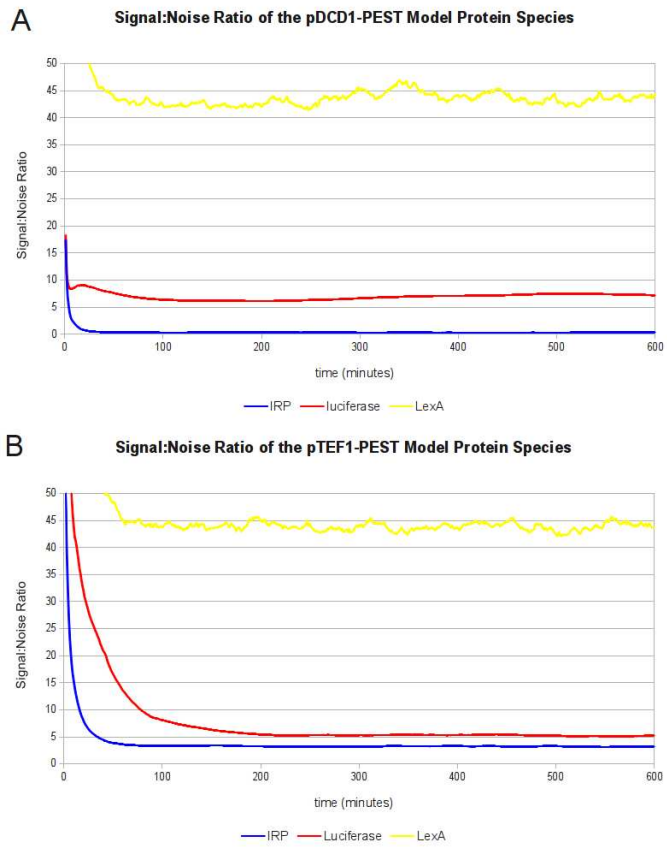

FIGURE 5.18: A) Is the signal to noise ratio for the P_{DCD1} model, and B) is the signal to noise ratio for the P_{TEF1} model. Graphs show the signal to noise ratio of the luciferase, LexA, and IRP protein components of the circuit during the time course simulation.

5.6 Parameter Estimation

The final stage in the construction of the model was to "fit" experimental data to the parameter values and attempt to refine the model with a more accurate time course simulation that represented the behaviour of the *in vivo* circuits. luciferase data had been collected from the lab, and provided data for the behaviour of each circuit strain. The particle numbers from the simulation data are not related to the relative luminescence units from the luminometer, therefore an additional parameter value was added to convert the luciferase particle numbers to the same scale as the relative luminescence data. Firstly, a global quantity was created for the transient number of luciferase particles minus the initial number of luciferase particles. This number was then multiplied by a second global quantity named "gain". The gain value could be tuned to scale the particle numbers to the relative luminescence units

(RLU) recorded experimentally, and the luciferase particle numbers from the model are baseline subtracted, as with the lab data being fitted. Copasi was used to perform the parameter fitting as it is pre-programmed with a number of parameter estimation algorithms [148]. The luciferase data for each luciferase time course experiment was fitted against the promoter strength in the model and the luminescence global variable. Parameter estimation was performed using the Hooke and Jeeves algorithm and the simulated annealing algorithm. The models were then compared with the experimental time course data for each circuit strain. The parameter values obtained from the parameter estimation algorithms are presented in table 5.13.

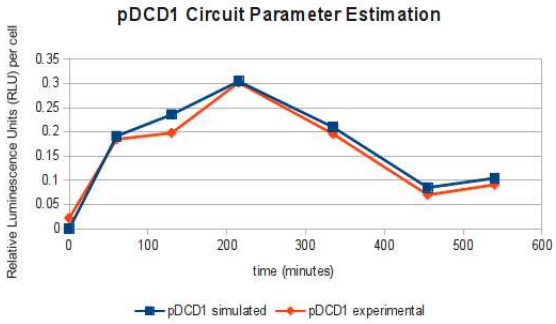

FIGURE 5.19: Graph of the P_{DCD1} model fitted to experimental luciferase time course measurements. The x axis contains the simulation time (in minutes) and the y axis contains the relative luminescence units (RLU) per cell, as calculated experimentally.

For the P_{DCD1} model fitting, the simulation closely tracked the experimental data and replicated the dynamic range of the circuit response (figure 5.19). Peak luciferase expression was observed at approximately 200 minutes and then slowly declined over 200 - 500 minutes. There was a margin of approximately 5% error between the

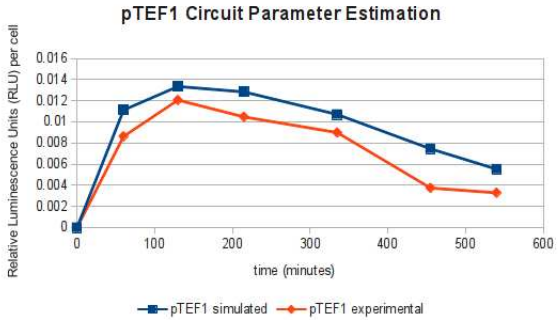

FIGURE 5.20: Graph of the P_{TEF1} model fitted to experimental luciferase time course measurements. The x axis contains the simulation time (in minutes) and the y axis contains the relative luminescence units (RLU) per cell, as calculated experimentally.

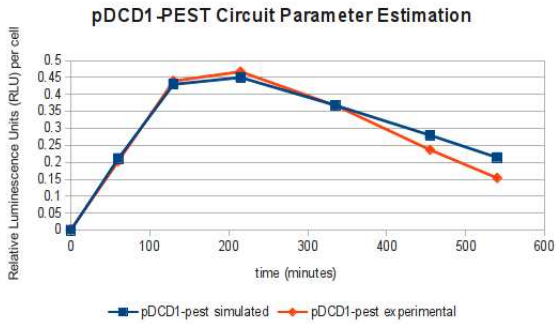

FIGURE 5.21: Graph of the P_{DCD1}-PEST model fitted to experimental luciferase time course measurements. The x axis contains the simulation time (in minutes) and the y axis contains the relative luminescence units (RLU) per cell, as calculated experimentally

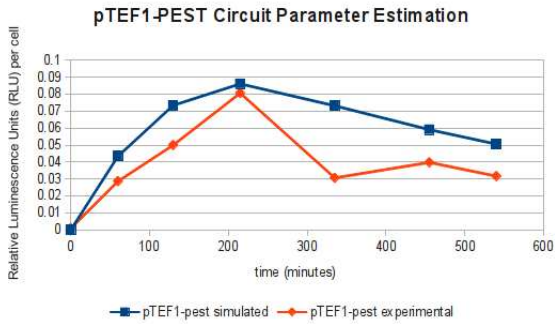

FIGURE 5.22: Graph of the P_{TEF1}-PEST model fitted to experimental luciferase time course measurements. The x axis contains the simulation time (in minutes) and the y axis contains the relative luminescence units (RLU) per cell, as calculated experimentally.

experimental and simulation data (figure 5.19). The parameter estimation algorithms have made small changes to all of the parameters in the model, and made a large reduction in the basal expression level of IRP (table 5.13).

For the P_{TEF1} model, the simulation also closely tracked the experimental data, with an error of approximately 20% (figure 5.20). The model replicated the lower level of luciferase expression from the P_{TEF1} circuit, and the slower decline in luciferase expression compared with the P_{DCD1} circuit. The algorithms made large reductions in the mRNA degradation rate of the IRP and LexA, and increased the affinity of the FUS1 promoter for the LexA and luciferase genes.

For the P_{DCD1}-PEST model the simulation tracked the experimental data with approximately 6% error. The P_{TEF1}-PEST model however was more difficult to fit. The simulation tracked the experimental data with an error of approximately 35%. This

P_{TEF1}-PEST model was the most difficult to fit due to the sharp decline in luminescence between 200 and 300 minutes, presenting a problem for curve fitting algorithms to map the parameters to. The model however predicted the maximum luciferase expression at 200 minutes with only a 6% over-shoot and tracked the reduction in luminescence over the 200 to 500 minute duration. The algorithms have also decreased the basal transcription rate of IRP, as well as the transcription and translation rates.

The fitting algorithms provided a good fit to the experimental data, however sharp increases and decreases in expression levels were problematic. It was not possible to obtain a more accurate curve fit for the P_{TEF1}-PEST circuit using additional genetic, particle swarm, and Levenberg-Marquardt parameter estimation algorithms, and a set of parameters to produce such steep decent after the 200 minutes measurement (figure 5.22) could not be found. The parameter estimation algorithms reduced the LexA translation rate, and reduced the basal transcription rate of LexA and the affnity of LexA for the IRP gene. The algorithms made additional small changes across all of the model parameters, as observed from the other models.

For a more accurate model of the P_{TEF1}-PEST circuit, additional measurements around the 200 minutes time point needed to be collected to obtain a smoother curve from the experimental data. Attempts to fit the model with additional data from RT-qPCR and western blotting (chapter 4) resulted in models that did not accurately track the luciferase experimental data. As the "output" of the circuit was the luciferase reporter gene, fitting the model with luminometer data provided sufficient predictive accuracy. More quantitative data is required from RT-qPCR and western blot techniques, with less variability for fitting with the luminometer data.

In terms of the modelling objective of the project, this approach demonstrated how to build a model of the biological circuit, parameterise it with literature data to obtain representative results, and then refine the model behaviour with specific experimental data. Sensitivity and metabolic control analysis also provided continuous experimental direction and hypothesis generation, throughout the construction phase in the absence of experimental data. The process demonstrated the value of using modelling as an

investigative tool prior to building synthetic gene circuits, even in the absence of accurate time course simulations. The model was used to refine the experimental approach during the construction of the circuit, and the data then used to refine the model into an accurate simulation of the *in vivo* circuit response. This process demonstrates a complete cycle of model design, hypothesis generation, and experimental investigation that is characteristic of synthetic biology. [13, 14].

Model Parameter	generic model	DCD1	DCD1-PEST	TEF1	TEF1-PEST
IRP degradation k	0.001	0.001	0.03	0.001	0.03
IRP mRNA degradation k	0.04	0.08	0.001	$1.6e^{-9}$	0.04
IRP transcription K_i	0.001	0.001	0.004	0.02	0.003
IRP transcription K_M	0.1	0.18	0.19	0.2	0.2
IRP transcription P	N/A	60	60	1	1
IRP transcription s	0.003	$4.9x10^{-12}$	0.002	0.002	0.005
IRP transcription V_{MAX}	0.3	0.6	0.01	0.6	0.7
IRP translation K_M	0.1	0.2	0.01	0.006	0.1
IRP translation V	0.75	1.5	0.7	1.5	1.5
LexA degradation k	0.02	0.03	0.03	0.04	0.04
LexA mRNA degradation k	0.04	0.03	0.02	$4.5x10^{-5}$	0.04
LexA transcription K_M	0.1	0.08	0.15	$4.3e^{-6}$	0.18
LexA transcription s	0.013	0.14	0.01	0.0003	0.0008
LexA transcription V_{MAX}	1.3	0.94	0.3	$5.3x10^{-6}$	0.05
LexA translation K_M	0.1	0.13	0.05	0.2	0.2
LexA translation V	1.3	2.5	2	$4x10^{-8}$	$5x10-8$
luciferase degradation k	0.006	0.01	0.01	0.003	0.002
luciferase transcription K_M	0.1	0.001	0.09	$6.8x10^{-5}$	0.06
luciferase transcription s	0.005	0.01	0.006	$3.2x10^{-7}$	0.01
luciferase transcription V_{MAX}	0.5	0.01	0.9	0.94	0.94
luciferase translation K_i	0.01	0.02	0.005	0.01	0.01
luciferase translation K_M	0.1	0.04	0.13	0.16	0.04
luciferase translation V	0.5	0.07	0.06	1	0.2
luciferase mRNA degradation k	0.04	0.07	0.05	0.08	0.08
signal degradation k	0.02	0.04	0.01	0.04	0.04

TABLE 5.13: Circuit model parameters, following parameter estimation using experimental luminescence time course data. Generic model refers to the model of the circuit prior to parameter estimation.

CHAPTER

SIX

DISCUSSION

6.1 Introduction

The gene circuit was designed to utilize components that had been characterized previously in yeast [4, 203, 208, 238], as a circuit that would function as a module, enabling cells to produce a quantifiable output in response to an extra-cellular stimulus. The project utilized a synthetic biology approach to building the circuit, through the application of computer modelling to augment molecular biology techniques, with computer aided design to understand the behaviour of the system.

The application of mathematical modelling enabled the early prediction of the interactions within the circuit while it was under construction. Sensitivity and metabolic control analysis provided insight into the key control elements of the circuit, such as the promoter strength and degradation rate of the IRP repressor. Through modelling and simulation, the circuit was modified from a single design into multiple versions of the circuit that enabled tuning of the luminescence output. This form of design and development is iconic of synthetic biology and differentiates the field from classical molecular biology and microbiology [8, 11, 14].

6.2 Design and Development

The primary strategy for measuring the circuit behaviour was to measure the luciferase reporter gene using luminescence. Additional measurements of the mRNA transcription, and protein expression levels were performed using RT-qPCR and SDS-PAGE and western blotting, in order to gather data to parameterise an *in vivo*-relevant, *in silico* model of the gene circuit. The modelling approach provided an insight into the dynamics of the circuit, and modelling could be integrated into the construction, as opposed to having distinct phases of construction and modelling. The two phases of the project were complementary and inter-dependent rather than successive. During construction, a simple mechanistic model of the interactions was sufficient to enable sensitivity and metabolic control analysis to be performed on an unparameterised model (figure 5.7).

Metabolic control analysis provided an early indication of the control that the IRP could assert over the expression of the luciferase reporter gene. The model indicated the abundance of the IRP repressor, and its rate of degradation, would exert a high level of influence on the expression of the reporter gene. To investigate the influence of IRP abundance, two alternative promoters were incorporated into the circuit design: the strong P_{TEF1} promoter and the weak P_{DCD1} promoter. P_{TEF1} had been shown in the McCarthy lab to be approximately 60 times stronger than P_{DCD1}, which may only produce one or two mRNA transcripts per cell (Malys N. and Pietroni P. (2011) McCarthy Lab, personal communication).

For perturbing the degradation rate of the IRP, the exact half-life of the wild type IRP had not been reported, however published research indicated it is over 12 hours, [205, 251, 252], which would make it difficult to repress with LexA within the two hour time-scale of the pheromone response pathway. A short half-life GFP, yEGFP3$_{PEST}$ was provided by Simon Avery, from published work by Mateus and Avery [216]. The yEGFP protein was fused with a PEST-rich 178 C-terminal residue of the G_1 cyclin Cln2, which was shown to reduce the half-life of the protein by 10 fold [216]. The C-terminal residue was cloned to the C-terminus of the IRP in an attempt to reduce the half-life by the same 10 fold, which would be approximately 1.2 hours and within the

2 hours time period of the yeast pheromone response [126]. Preliminary western blot data indicated IRP_{PEST} did have a reduced half-life, and expression appeared to have been reduced by approximately 10 fold (figures 4.19 and 4.21).

The combination of modelling and experimental investigation during the construction phase resulted in four versions of the circuit that enabled the tuning of the reporter gene expression; P_{DCD1}, P_{TEF1}, P_{DCD1}-PEST, and P_{TEF1}-PEST. The four circuits enabled investigating high and low expression levels of the IRP, combined with short and wild-type half-life of the IRP.

6.3 Characterisation

The circuits were first assayed using the luminometer to observe their behaviour in terms of expression of luciferase. Direct comparison of the raw data from each experiment (figures 4.2, 4.5, 4.3, 4.6) showed unexplainable results, such as the P_{DCD1}-PEST circuit demonstrating higher expression of luciferase than the P_{DCD1} circuit, which intuitively should not be so, as the circuit also demonstrated a higher base line expression compared with the P_{DCD1} circuit and control (figure 4.3). An explanation for this was that the *DCD1* promoter expression of IRP is very low in the cell and consequently, is not inhibiting the translation of luciferase mRNA. However this does not explain an *increase* in expression from cells harbouring the P_{DCD1} and P_{DCD1}-PEST circuit plasmids. Comparison across experiments, however is unreliable due to variation between cultures and experimental conditions. Therefore, relative fold-change in expression was calculated for each experiment that enabled comparison of the ratio of baseline expression of each experiment to the induced level of expression. It was also hypothesized that a true negative control for this work may not be possible, as this would require a promoter that is not induced by the pheromone response but provides equivalent basal expression as the wild type P_{FUS1} promoter. Finally, the baseline expression of the gene circuits was compared with yeast transformed with the reporter plasmid alone. The expression of the reporter gene would be expressed under different cellular conditions, compared with cells transformed with three plasmids, and

grown in different auxotrophic drop out media. Fluctuating basal expression from the repressor and de-repressor plasmids would affect expression of the reporter plasmid, as well as the metabolic differences between control and circuit cells. If the work was repeated, the luciferase control should have been run with each experiment, however logistical limitations of batch culture of three biological replicates, with three technical replicates precluded additional cell cultures, if measuring optical density, cell number, and luminescence simultaneously, whilst sampling for western blot and RT-qPCR, from cultures with and without pheromone induction over a nine hour period. For these reasons it was decided that calculating the increase in luciferase expression in induced cells as a function of luciferase expression in un-induced cells of the same culture (the fold-change induction), provided a reasonable comparison of the level of induction between the circuits.

From the circuit variants constructed during the project, the P_{DCD1} promoter expressed the most effective level of IRP repression when compared with the stronger P_{TEF1} promoter, producing a higher fold-change increase in the luciferase output of the circuit, compared with the P_{TEF1} circuit.

The time required to reach maximum expression of luciferase was extended for both the P_{DCD1} and P_{TEF1} circuits when compared with the control strain (figures 4.4 and 4.7). Signal transduction through the pheromone response pathway requires approximately two hours to reach maximum expression [55], as observed from the control strain, transformed with only the reporter plasmid (figures 4.4 and 4.7). Both the P_{DCD1} and P_{TEF1} circuits required approximately 3.5 hours to reach maximum expression. This delay can be explained by the system used for de-repressing the system. LexA transcription is also up-regulated by the yeast pheromone response and therefore requires approximately two hours to reach maximum expression levels, at which time it is able to repress transcription of the IRP. The system then requires the existing IRP mRNA transcripts to be degraded by the cell (a process that takes approximately 23 minutes [280]), and for the existing IRP protein to degrade.

The wild-type IRP has a half-life of over 12 hours [205] therefore overall expression of luciferase would be inhibited for significantly longer than the period of the pheromone

response. Using western blot analysis the short half-life IRP_{PEST} was estimated to have a 10 fold reduced level of expression in the P_{TEF1}-PEST circuit (figure 4.20), indicating the PEST residues fused to the C terminus where functioning to target the IRP for degradation.

The effect on luciferase expression in the P_{TEF1}-PEST circuit was to increase the level of luminescence to a similar level as the control within 3 hours, compared with the wild-type IRP in the P_{TEF1} circuit which remained lower than the control throughout the experiment (figure 4.7). For the P_{DCD1}-PEST circuit, the IRP_{PEST} resulted in a longer period of luminescence following pheromone induction compared with the wild-type IRP in the P_{DCD1} circuit, due to reduced repression of luciferase mRNA translation (figure 4.4). The data contained a high level of variability however, and estimation of protein half-life and luminescence could not be measured with sufficient accuracy to make significant conclusions.

Control data obtained for the circuits without the de-repressor plasmid demonstrate the interaction of the LexA protein in the full circuits. In the control circuits, maximum luciferase expression was observed increasing between zero and 240 minutes in both the circuits and the control, but luminescence was maintained at a lower level throughout the experiments (figures 4.10 and 4.11). At the maximum level of luciferase expression for the control (approximately 240 minutes), the P_{DCD1} and P_{DCD1}-PEST circuits demonstrates an increase of approximately 50% of the control, while the P_{TEF1} circuits where approximately 30% of the control, increased to approximately 70% of the control for P_{TEF1}-PEST. The LexA de-repressor restores luciferase output of the P_{DCD1} circuit to 120% and P_{DCD1}-PEST to 160% of the control (figure 4.4). For the stronger P_{TEF1} promoter, repression of IRP transcription by LexA is insufficient to restore luciferase expression to the same level as the control, and it remains at approximately 75% of the control (figure 4.7).

Further characterisation was possible through SDS-PAGE and western blot. There was a problem obtaining antibodies sufficiently sensitive to detect low levels of luciferase and IRP protein in yeast cell extracts by means of western blot. Antibodies were obtained from Santa Cruz, Sigma Aldrich, and Abcam and tested at 1:200

to 1:100 dilutions with multiple repeats. No signal could be obtained for the luciferase antibodies using fluorescein Isothiocyanate (FITC), Horseradish peroxidase (HRP), alkaline phosphatase, and infra-red conjugated secondary antibodies. However, luminometer data was available to quantify luciferase expression. For the IRP, neither the Sigma Aldrich or Santa Cruz primary antibodies could detected the protein, however the Abcam antibodies produced a signal when using the alkaline phosphatase detection method, but only for the P_{TEF1} circuit with the higher IRP expression level (figure 4.20). The alkaline phosphatase method is an enzyme-based assay, and therefore amplified the weak signal from the IRP primary antibodies, but required a high level of expression to produce a signal. The lack of western blot signal for the P_{DCD1} circuit correlated with the hypothesis that the P_{DCD1} promoter produces a low level of transcription (Malys N. and Paola P. (2011). McCarthy Lab. Personal communication). Further work is ongoing at the McCarthy lab to characterize the expression level of the P_{DCD1} promoter. The increase in fold-induction observed from the luminescence data for the P_{DCD1} circuit could be explained by the possible high affinity of the IRP for the IRE [4] and the long half-life of the protein [252], resulting in sufficient expression of IRP to repress luciferase translation, but sufficiently weak to enable repression of transcription by LexA, resulting in the higher fold-change observed in the luminescence data, whereas the stronger P_{TEF1} promoter results in the IRP quenching the signal from the circuit. However, the antibodies for IRP are not sensitive as they do not provide a strong signal at low dilution, and high concentration of the protein from the P_{TEF1} promoter.

Preliminary western blot data could be collected from the signal obtained from the P_{TEF1} and P_{TEF1}-PEST circuits. The western blot data for the P_{TEF1} circuit showed an 70% reduction in IRP expression during the pheromone response, which was attributed to repression by LexA (figure 4.22). Inhibition increased throughout the duration of the yeast pheromone response and remained for the duration of the experiment (figure 4.20), correlating with an increase in LexA expression between two and five hours after pheromone stimulation (figure 4.25). The short half-life IRP_{PEST} produced a weaker signal than the wild-type IRP, and was difficult to obtain for the expression of the PEST tagged IRP, however preliminary western blot data indicated a reduction

of approximately 60% compared with the wild type at the initial measurement (figure 4.20).

The western blot data for LexA showed a high basal expression, and a smaller fold-change in expression, compared with the luciferase luminescence data (figures 4.24 and 4.25). The lower fold-change in LexA (three to five-fold increase) compared with luciferase (five to seven- fold increase) could indicate LexA was not being expressed correctly. There was a high level of variation in the data however,making accurate quantification of fold-change difficult. The higher fold-change in luciferase protein expression compared with LexA protein expression may also highlight the effect of the circuit design, in that constitutive repression of luciferase by IRP decreased baseline expression, increasing the fold-change under pheromone induction. This conclusion is speculative however due to the variation in the data, and comparing expression levels across different assays.

The quantitative data from the western blot using the alkaline phosphatase method were very noisy. The reaction is enzyme-based, as with the luciferase method therefore it is difficult to relate the signal measured from the film in terms of numbers of molecules in the protein extract. Variation in the amount of protein in the original protein extract, the western blot transfer, and the enzyme reaction on the membrane make quantification of fold-changes in protein levels speculative. The data provides an indication as to the directional change in the protein levels expressed by the circuits during the pheromone response. As such, the western blot confirms the expression of the LexA and IRP components of the circuit and indicates a relative reduction in IRP compared to the non-induced circuit during the pheromone response that correlates with a similar increase in LexA signal that can be inferred as repression of IRP by the LexA repressor.

Finally, RT-qPCR was used to attempt to quantify the fold-change in mRNA expression that was occurring in the circuit during the pheromone response. RT-qPCR had not previously been used in the McCarthy lab for quantification of mRNA, since calibrated northern blotting had been the preferred method. A RotorGene RT-qPCR machine was sourced from QiaGen and a series of house keeping genes validated for use in yeast, based on published data by Teste *et al.* (table 4.2) [257]. House keeping genes

were confirmed using bioinformatics data from the SGD database [258], based on micro-array data of gene expression during the pheromone response [85], and primers designed for the components of the luciferase, IRP, and LexA components of the gene circuit (table 2.2).

The mRNA data obtained from the time-course samples were also very noisy, and a larger number of replicates was required to obtain statistically significant data. The data obtained by RT-qPCR however provided an indication of directional changes in the relative expression levels of components of the gene circuit, comparing induced and non-induced cells. The data indicated that during the pheromone response, the level of luciferase and LexA mRNA increased, as expected from the P_{FUS1} pheromone-induced promoter. mRNA expression increased from approximately 1 hour after induction for luciferase and LexA (figures 4.29, 4.32, 4.31, and 4.32) for all circuits. IRP mRNA levels began to reduce within 2 hours of pheromone stimulation. The P_{TEF1} circuit demonstrates a higher rate of recovery with approximately 5 fold increase in expression of IRP after 5.5 hours of pheromone stimulation (figure 4.30, compared with a 2 fold increase after 7.5 hours after stimulation for P_{DCD1} (figure 4.29). The data supported the western blot and luciferase data showing higher IRP expression level from the stronger P_{TEF1} promoter, and mRNA levels correlate with the increases in protein and luciferase expression levels. The data from the RT-qPCR experiments however were extremely noisy and statistical analysis to determine significant differences was not possible. A larger number of replicates is required for reliable quantification of the mRNA levels in the circuit, also screened against a number of house-keeping genes simultaneously to normalise the relative expression levels of the components. The RT-qPCR MIQE guidelines provide recommendations for performing RT-qPCR studies and recommend a minimum of 10 replicates for each time point screened with at least 10 house keeping genes [259]. It was not possible with the resources available to this project to perform RT-qPCR in accordance with all MIQE guidelines. However, the data does provide information on the directional changes in mRNA expression levels in the circuit and can be used to confirm the interactions of the LexA and IRP components and to confirm the plasmids are being expressed in the host cell.

The inhibition of luciferase observed from the luminescence data confirms the interaction of the IRP with the IRE, cloned upstream of the luciferase reporter gene, as published by Koloteva *et al* [4]. The inhibition of IRP expression during the pheromone response, observed by RT-qPCR and western blot, and the observed up-regulation in LexA also by RT-qPCR and western blot, confirms the interaction of the LexA repressor with the IRP gene, and the function of LexA in yeast as published by Brent [5]. The data confirms the LexA operators are functioning in their position immediately upstream of the promoter on the repressor plasmid by the repression of IRP. The RT-qPCR data indicates the LexA protein is capable of blocking transcription of IRP from the P_{DCD1} promoter, and for a shorter period of time from the P_{TEF1} promoter. The half-life of the IRP however can negate the inhibition of transcription, by providing a long-lived protein that can continue to repress the reporter gene, preventing de-repression of the circuit output, as observed from the P_{TEF1} circuits. The IRP western blot data demonstrated the half-life of the IRP protein can be reduced through the addition of the Cln2-PEST C-terminal residue to the protein. The western blot data conforms with the published work by Mateus and Avery [216], and western blot data indicated the abundance of the IRP was reduced by approximately 10-fold (figure 4.21). Further work, and more effective antibodies are required to determine the exact half-life of IRP_{PEST} species in yeast. From the preliminary data however, the development of the short half-life IRP_{PEST} provides additional tools for synthetic biology circuits with a high affinity, short-lived translational repressor. The PEST degradation tag could also be cloned into additional proteins such as luciferase, to obtain a short-half reporter gene that would permit higher response times for the circuit output.

6.4 Noise

Signal:Noise ratio (SNR) is of interest to this project, as one of the objectives of the circuit design was to reduce the noise from the reporter gene by repressing basal expression levels. The SNR was calculated for the luciferase data to explore the change in noise in the circuit output during the time course experiments (figures 4.12 B, and 4.13 B). The data showed that as the expression of luciferase increases in the first 200

minutes of pheromone-induction, the SNR increases for all of the circuits, and the control (the reporter gene with no repression) (figure 4.14). However, although the reduction in the expression of luciferase for the P_{DCD1} circuits occurs at approximately the same rate as the control (figure 4.4), the SNR reduces sharply for the circuits, with a lower SNR for all of the circuits after 200 minutes compared with the control (figures 4.12 and 4.13 B). The SNR decreases at a sharper rate than the luminescence measurements decrease, while the control remains constant during the decrease in luminescence, after the two hour time scale of the pheromone response pathway.

Noise propagation throughout the experiments is high however (as observed from the luminescence data), which makes it impossible to draw statistically significant conclusions from the data (figures 4.12 A, and 4.13 A). Observations of the effect of the circuit design and function on the level of noise from the circuit output is speculative. From the data obtained from this project, the SNR data indicated the circuits are less noisy when the circuit is at maximum luciferase expression, compared with the pheromone-induced reporter gene alone, but the circuits are more noisy than the control when the circuit returns to the OFF-state (from 200 minutes to the end of the time course experiments). The increased SNR in the circuit, compared with the control however indicated the design of the circuit may be working. As the luminescence measurement is a measurement of a population of cells, some would be responding to pheromone and some would not, in a stochastic response to pheromone stimulation. As the luciferase reporter gene is expressed from a wild type P_{FUS1} promoter and therefore only capable of the same maximum luciferase expression level as the control, the higher SNR can be attributed to lower basal expression level in the circuit cells.

Interaction with the IRP, and fluctuations in IRP expression though interaction with LexA may add noise to the luciferase protein expression, after the maximum period of activation by the pheromone response pathway has passed. This is a hypothesis however, and requires more quantitative data than was obtained during this project, such as quantitative mass spectrometry to accurately quantify the expression levels of each of the protein species.

It is possible that in addition to the small fold increase in luciferase expression, the circuit design may have additional effects on noise in the expression of the reporter gene when incorporating low level constitutive repression. The model of the P_{DCD1} circuit predicted bursting in transcription, and consequently bursting in protein expression levels, from the low level of activity of the *DCD1* promoter. The *DCD1* promoter therefore, may contribute additional noise to the luminescence measurement (figure 5.8). More data is required to support the hypothesis made regarding noise in the circuit. Noise propagation and the effects of noise on luciferase expression could be further investigated using high-throughput single cell measurements using instrumentation such as flow cytometry, as has been done with oscillating circuits by Elowitz and Liebler, and later by Atkinson, and attempt to determine the source and contribution of intrinsic and extrinsic noise in expression of the circuit components. [165, 179, 296, 297].

6.5 Modelling

The modelling component of the project has generated a set of tools and an approach that can be used for building circuits from interacting genetic components. Previous published research attempts to build fully parameterised models of biological systems that incorporate all of the details of the involved reactions [128], or even all of the reactions in the cell [298–300]. While fully parameterised models theoretically provide a quantitative, predictive tool for biology, mechanistic models that represent what is known about the system being studied can be interrogated using tools from systems biology such as metabolic control analysis [149] and sensitivity analysis [301]. The mechanistic model of the system built at the start of this project provided a framework, from which MCA was used to determine the most influential reactions of the circuit (figure 5.7, 5.13, and 5.14), and investigate the contributions of each component to the systems level behaviour of the circuit. MCA has been used in studies of large models of yeast metabolism [302, 303], however it has not been used in synthetic biology applications where published models focus on predictive time course simulation [8, 128, 178–180]. The application of MCS in this study was a novel approach to

using a model as a design tool to *build* the circuit, rather than as a tool to confirm the understanding of experimental observations [126, 128, 178]. In many synthetic and systems biology studies, data is obtained from biological systems in order to refine the model, whereas in this study the model was used to refine the biology, and then the experimental data used to refine the model to enable additional functionality, such as time course simulation.

The MCA data highlighted the control of the IRP, leading to the development of additional versions of the circuit with varying strength promoters for the repressor. The circuit combined repression at transcription and translation and throughout the simulations, the repression of luciferase translation remained the most significant component in changing the level of luciferase output. This indicated that tuning repression of the IRP by LexA would not have as large an effect as tuning the IRP expression levels and half-life. It is expected that translational inhibition would be more effective than transcriptional inhibition, due to the abundance of mRNA template and repressor protein after the initiation of repression by LexA. Therefore, while the model was created using Michaelis-Menten kinetics, the predictions conformed with expected biological understanding of the system, and validated the approach taken to modelling the gene circuit.

The half-life of the repressor was also highlighted and the PEST tag modified IRP_{PEST} was developed. Therefore, while a molecular biology understanding of the interactions of the components guided construction of the circuit experimentally, the model provided data on tuning the behaviour of the circuit which could not be investigated empirically. These hypotheses could not have been made during the construction phase of the project, and would have required testing followed by revision of the circuit with further construction and testing. Using the modelling approach, construction could be expanded prior to generating experimental data, providing rounds of modelling and construction, followed by rounds of data acquisition and model development.

Populating the model with parameter values from the published literature enabled refining the model, such that it began to provide time-course simulation predictions of the behaviour of the circuit (table 5.7 and figure 5.11), and predicted an output that

correlated with the 2 hour time-scale of the yeast pheromone response pathway, and could be fitted to experimental data.

Deterministic modelling of the system using a set of coupled ordinary differential equations provided an overview of the behaviour of the circuit, however these simulations rely on a well mixed, homogeneous environment with a large number of molecules. Many biological processes have been shown to be stochastic in nature, with small numbers of randomly distributed molecules [288]. Stochastic simulation of the circuit revealed, for the P_{DCD1} promoter using a rate law 60 times lower than the P_{TEF1} circuits, produced mRNA levels at 1 or 2 molecules per time course simulation, which correlated with the estimated number of mRNA per cell produced from this promoter (Malys N. (2011), personal communication) (see figure 5.8 A). Although there was no kinetic data available for the exact rate of transcription from the P_{DCD1} and P_{TEF1} promoters, the model was accurately simulating an appropriate level of mRNA expression for the circuit. The stochastic events in mRNA transcription in the model were also producing bursting behaviour in the protein translation level in the model (figure 5.8 B). The model also provides a tool, therefore for investigating noise in transcription and translation in the circuit. It is likely from the model prediction that the P_{DCD1} circuit will have more noise than the P_{TEF1} circuit, due to stochastic noise gene expression of the IRP repressor. The model provides a potential tool for the further investigation of the SNR data in the luminescence data (figure 4.13), using stochastic modelling to investigate more detailed interactions such as bursting, which can be key to determining the origin and influence of noise in a biological system [295, 304].

At the end of the project, the model parameter values were fitted to the luminescence data, to produce a model output that represents the observed experimental behaviour. Parameter fitting provided final models for each of the circuits that simulated time course experiments in terms of luciferase expression, with a close fit to the experimental data (figures 5.19, 5.20, 5.21, and 5.22).

The fitting algorithms are not able to make changes that are based on biological understanding, and the changes are not consistent across all of the models (chapter 5, table 5.13). The rates of transcription and translation is consistent across the

gene circuits as they use the same promoters, however these rates are different across models following parameter estimation. Further work was required on the strategy for parameter estimation, to standardize the rates of transcription and translation and attempt to fit the model based on the known changes that had been made to the circuit. Allowing the parameter estimation algorithms to fit the experimental data to all of the model parameters provided a first pass at refining the models so that they provide an *in vivo* relevant simulation of luciferase expression for each of the circuits. The models can be further developed with additional experimental and literature investigation for parameter values for transcription and translation for the specific promoters, as well as the binding affinities for the repressors and transcription factors.

The MAP kinase model developed in the early stage of this project was not utilized for the modelling of the gene circuit, as the circuit was activated with a saturating concentration of pheromone, and the data was collected from the pheromone response pathway at steady-state, therefore fluctuation in the MAPK cascade was not expected to influence the dynamic behaviour of the circuit. The MAPK model also incorporated a large number of reactions and would have added significant computational time to stochastic simulations of the circuit, which were used for the time course simulations. Fitting experimental data to a large model would also require a large amount of computational time, which was not available towards the end of the project. If unexpected experimental behaviour had been observed, then the circuit model could have been coupled with the reactions from the MAPK model to increase the detail of the simulations. Also, with regard to modelling circuits coupled with the mating response pathway, smaller mechanistic models such as the model by Huang and Ferrell can simulate the dynamic behaviour of the MAPK cascade without requiring all of the detail of the pheromone response pathway [122]. This approach of using a simplified model of the cascade was recently used by O'Shaughnessy *et al* to model a synthetic signalling cascade in yeast [46].

It should not be ignored however, that the circuit is activated by a complex cascade of upstream events. The MAPK cascade provides a number of benefits and issues for investigating the circuit. Using the circuit to study noise in cell signalling is not possible, as the cascade functions to dampen fluctuations in the external signal, generating a

binary response once the signal achieves a threshold concentration (hyper-sensitivity). Varying the level of pheromone therefore would have no effect on the output of the level of circuit behaviour, except at threshold activation levels where activation may burst, however this is purely hypothetical and has not been observed experimentally [120]. The MAPK cascade does however provide a native signal processing module for the circuit, and perturbation of the cascade components could produce more complex behaviours in the circuit output such as oscillations [124].

6.6 Summary and Further Work

For modelling, the number of molecules per cell of the reporter gene would provide quantitative data for parameter fitting, and performing predictive stochastic time course simulations. luminescence can be measured in terms of luciferase molecules per cell if it is properly calibrated, however the reaction is enzyme-based and is therefore non-linear, requiring understanding of the kinetics of the enzyme reaction, in the specific conditions of the experiment [204]. A more suitable reporter gene would be a fluorescence protein, such as GFP, that enables direct quantification of fluorescence that can be coupled with a simultaneous cell count, such as is possible with flow cytometry [159, 178, 305]. High throughput instrumentation is not currently widely available to measure luminescence, and requires additional steps of substrate addition prior to measurement, and then maintenance of a saturating substrate concentration if measured over time. Therefore, reporter genes encoding fluorescent proteins would be more appropriate for time course measurements of gene circuits. However, GFP can also be very bright, and produce a high basal fluorescence that would reduce any observed fold-increase during induction with pheromone. Also, GFP is a very stable protein, and would require the short half-life species constructed by Mateus and Avery to enable observations of the dynamic range of the circuit behaviour within the time-scale of the pheromone response pathway [216]. Further refinement of the *in vivo* luminescence assay to reduce the variation between measurements, and transferring to a micro-plate based, high-throughput methodology would benefit future work by retaining the sensitivity of the luminescence reporter gene, compared with moving to fluorescent proteins.

A number of potential experimental investigations of the circuit were outstanding at the end of the project. Firstly, this study focused on the components behaving as a system of interactions. It would be useful to characterise the behaviour of the components individually, and obtain specific Ki values for the IRP and LexA repressor in the sst2Δ yeast strain that were hosting the circuit. This would provide additional parameter data to further refine the model. It would also be of use to investigate the dynamics of the circuit in terms of returning to the OFF-state. The data collected during the project characterised the global behaviour of the circuit, in terms of the output in response to activation signal. The cells should be stimulated with pheromone, observed for maximum luciferase activity, and then the pheromone washed out of the system, and the rate of decrease in the luciferase reporter measured. This would provide data on the efficiency of the IRP repressor for inhibiting the reporter gene, as well as the ability of the circuit to reset after activation. This is another important feature of the combination of repressors in the circuit: their ability to control the dynamic range of the circuit and increase the sensitivity of the pheromone-induced reporter gene to activation and de-activation. The data collected from this work indicates there is a gradual decline in the pheromone-induced reporter gene. The P_{TEF1}-PEST circuit provided an indication that the circuit is capable of initiating a faster return to the OFF-state compared to the control strain. The P_{DCD1} circuits did not sufficiently repress the reporter gene to effect a rapid transition to the OFF-state upon the de-activation of the pheromone response. Also, the interactions of the LexA de-repressor will be influential in this reaction as following the pheromone response, LexA will continue to de-repress the IRP from the weaker P_{DCD1} promoter for a longer period of time than the stronger P_{TEF1} promoter.

The design of the circuit incorporated repression at the transcriptional and translational levels. This is unique in the design of repressor interactions in current published gene circuits, where researchers focus on the repression of transcription. The circuit presents the opportunity to investigate the timing difference between translational and transcriptional repression. Swapping the level of repression on the IRP and luciferase components would enable comparison of the effect of different hierarchical levels of control on the output of the circuit; this could be achieved by reversing the repressors in the circuit. The LexA repressor could be cloned into the repressor plasmid in place

of the IRP, and the IRP cloned into the pheromone-induced de-repressor plasmid. The IRE in the reporter plasmid could be replaced with the LexA operators and the IRE placed in the repressor plasmid upstream of LexA. This would create constitutive repression of the transcription of luciferase and pheromone-induced repression of LexA mRNA translation. Theoretically, this would enable the repression of luciferase mRNA transcription, which would be effective within the time frame of the luciferase degradation rate (three hours) [285] (which could also be PEST tagged for a shorter half-life). Upon pheromone stimulation, the LexA repressor would be inhibited at translation by the IRP, which could be faster than the current circuit design, and enable a faster up-regulation of luciferase. A faster induction time, and a short half-life luciferase could enable a faster transition between the ON and OFF-states of the circuit and a more sensitive circuit response to pheromone activation. Short half-life variants of all of the components would enable tuning of the repressor and reporter gene expression, and the dynamic range of the circuit output.

Chromosomal integration of the circuit would have been beneficial during the project as, although the plasmids are believed to be single copy, there is no confirmation of this in the data from this project. The plasmids also require an auxotrophic amino acid biosynthesis marker to be retained by the host cell, which leads to a burden on cell metabolism. The growth rate of the cells transformed with the circuit plasmids was checked, and the circuit did not appear to have a detrimental effect on the growth rate, however the cells are grown in an environment lacking essential amino acids which are substituted on the plasmids, but requiring synthesizing rather than uptake from the media. This may have an additional effect on the metabolism of the host cell that could be avoided by integrating the circuit components into the chromosome. Also, following chromosomal integration, the cells cannot lose the gene circuit, unlike with plasmids. Chromosomal integration of the gene circuit would also enable the addition of new components to the circuit, using additional plasmids.

For quantification of the components, addition of a C or N-terminal Myc or HA tag to the IRP would enable the quantification of expression levels at a higher resolution than was achieved during this project. Alternatively, specific custom antibodies could be created for the components, however modification with a myc or HA tag would

enable more accurate protein quantification by western blot, without expensive custom antibodies. The luciferase, IRP, and LexA components are sufficiently different in size that tagging all of the components with myc or HA would enable quantification of all of the components in a single western blot. Purified protein controls for each of the components is required however for further western blot analysis. A positive control is required to confirm the western blot signal is correct and also a calibration curve could be created from a known quantity of purified protein, enabling quantification of the components instead of the relative fold-change, which would be more informative and also more appropriate for model parametrization.

Blotting techniques are only semi-quantitative however. Variation can occur during the protein extraction, gel loading, transfer, and imaging stages, increasing the separation between the original cell culture and the final data. A collaboration was being undertaken at the end of the project between University of Warwick, University of Manchester, and University of Liverpool to use mass spectrometry to quantify components of the pheromone response pathway with QconCAT. This technique, once validated would provide quantitative data on the expression level of the protein species in the circuit, to an accuracy significantly greater than that achievable by western blotting or luminometer measurement. Quantification of the yeast pheromone response pathway, especially Ste11, Ste7, and Fus3, would also enable parametrization of the MAPK model and development of phospho-proteomics, which would enable a much more accurate model of the phosphorylation states of the components. This would allow for the investigation into the signal flux through the cascade and allow the propagation of the signal at a much greater resolution and understanding than has been currently published. It has been shown that a MAPK cascade can be transplanted between organisms [46], therefore this data and the model would provide a signal transduction input module for synthetic circuits that could be transplanted between eukaryotic organisms, and understanding of the dynamic behaviour of the phosphorylation states would enable tuning of the signal dynamics to customise the response to environmental stimuli.

Quantitative data on the dynamics of the MAPK cascade, as well the changes in the levels of the circuit components, would also enable a more detailed understanding of the

interactions that make up the whole process of translating pheromone-binding at the cell membrane, through to induction of the pheromone response genes and the circuit output. This data may reveal complex behaviours at different stages in the signal cascade, such as bursting or oscillation. Understanding of these various dynamic behaviours within the overall behaviour observed during this project would enable more accurate models to be constructed. The mechanisms of the interactions is reasonably well understood, enabling mechanistic models to be made, however a lack of understanding of the dynamic behaviours results in rate laws that do not accurately represent the behaviour of the reaction, and result in a model that is correct in terms of the reactions but does not simulate the behaviour observed experimentally. Understanding the rate of phosphorylation of the MAPK components, or the stochastic bursting in transcription and translation [304] will ultimately lead to a greater resolution for modelling and consequently synthetic circuits and systems, with predictable and robust behaviour.

Additional loops could be added into the circuit for more complex dynamic behaviours. For example, a feed-forward loop could be integrated through a pheromone inducible plasmid containing the α-factor pheromone gene, as in the study by Gross *et al*, where pheromone-induced pheromone expression was used as an amplifier in a gene circuit [1]. Addition of such an "amplifier plasmid" could lead to a toggle-switch type behaviour with a continuous signal after activation. Inducible promoters from other pathways, such as the Hog pathway for osmotic stress could be used to induce secondary circuits such as Hog-induced expression of Bar1, a protease which degrades α-factor, could be used as a second toggle-switch to switch the amplifier circuit off. Adding plasmids to the circuit for constitutive expression of pheromone and pheromone-induced expression of Bar1 could also induce oscillatory behaviour in the circuit. These reporter genes could also be tagged with different fluorescent proteins such as green and red fluorescent proteins, to enable dual reporter systems to be developed, similar to the work by Ajo-Franklin [190]. Simultaneous application of different signal pathway induced promoters, such as the pheromone response pathway and the Hog osmotic shock pathway would enable the investigation of cross talk between the pathways, as well as investigating the ability of cells to respond to multiple environmental stimuli.

The host cell is also of interest for studying the behaviour of the circuit. The strain of yeast used in this study were *S. cerevisiae sst2*Δ, with a mutation in the negative feedback of the pheromone response pathway (the re-association of the α subunit of the membrane bound G-protein to the βγ subunits to stop the phosphorylation of the MAPK cascade). This mutation makes the cells hyper-sensitive to pheromone, enabling lower concentrations to be used to stimulate the pheromone response, and a longer period of stimulation to be achieved. This strain is therefore optimized for studying the activation of the circuit and obtaining the maximum possible fold-change in the activated circuit. The cells will continue to express Bar1, a protease that degrades α-factor pheromone and Far1 which growth arrests the cells in preparation for mating. A mutant strain containing additional mutations of the *FAR1* and *BAR1* genes was under construction at the time of writing and would provide a better host strain for the circuit, as the cells would not undergo growth arrest during the pheromone response, and continue to grow and express the circuit components following stimulation, which may further amplify the luciferase circuit output. Ultimately, mutation of all of the yeast pheromone response genes would result in a cell that directs the pheromone response exclusively to the gene circuit, however this requires the mutation of over 200 genes (3% of the genome) [55] and the systems level effect of this is unknown, as is the degree of complexity of the interaction between the components of the mating response with other metabolic pathways [71].

The design of the "chassis" for gene circuits is an important consideration for synthetic biology, and is as important as the design of the circuit. Systems biology programs are attempting to understand the metabolism of *S. cerevisiae* [299], however this research is still emerging. Other researchers are attempting to construct a "minimal cell" by deleting all the "non-essential" genes that are not required for sustaining a viable cell using "bottom-up synthetic biology" [306, 307]. A minimal cell chassis provides a base from which to study a synthetic gene circuit where all the interactions can be understood. Building such a system, even in prokaryotic organisms remains an enormously complex task [202].

6.7 Conclusion

The project represents a complete cycle of design and development in synthetic biology (figure 1.6) [14]. The model and circuit were constructed, and the experimental data fed back into the model for parameter estimation and model refinement. The model provided an *in silico* relevant simulation of the circuit that can be used in a computer-aided design role to build additional functions, or further tune the circuit. The IRP repressor, previously characterised in the McCarthy Lab functions alongside the LexA repressor, characterised by Brent in the context of a gene circuit in yeast [4, 5]. The application of a constitutive repressor can be used to "boost" the ratio of expression of a reporter gene between the induced and un-induced states, however the interactions create additional noise in the output, which can be tuned through the application of varying strengths of the constitutive repressor and the half-life of the repressor protein. The development of a short half-life PEST tag enabled the tuning the dynamic range of the circuit output, and also provides an additional tool for constructing synthetic biology circuit components. The project provides a novel combination of transcriptional and translational repression that can be used to study noise, and to control the expression of genes at different hierarchical levels within the cell. Differentially-regulating transcription and translation also enables the investigation of the temporal separation of these events as well as noise at each level of gene expression.

The mathematical modelling approach undertaken in this study provided a truly novel approach to investigating the behaviour of the gene circuit *in silico*, as a predictive design tool for building circuits *in vivo*. The approach identified key control points within the circuit without requiring a fully parameterised model, or accurate time course simulations, prompting the construction of multiple versions of the circuit, modifications of the components, and expanding the scope of the project. These hypotheses would not have been formulated from empirical observation of the circuit design, and would have required multiple rounds of experimentation to obtain the data from this project. Parameterisation of the model then provided additional predictive capabilities such as time course simulations, that could be further refined with experimental data, evolving the model alongside the circuit, and completing the cycle

of modelling, hypothesis generation, experimentation, and further model refinement [14]. The project therefore demonstrates the benefit of combining computational and biological sciences in future biotechnology and genetic engineering studies.

BIBLIOGRAPHY

[1] A Gross, G Rödel, and K Ostermann. Application of the yeast pheromone system for controlled cell-cell communication and signal amplification. *Lett Appl Microbiol*, 52(5):521–526, May 2011. ISSN 1472-765X. doi: 10.1111/j.1472-765X.2011.03035.x.

[2] Sergi Regot, Javier Macia, Núria Conde, Kentaro Furukawa, Jimmy Kjellén, Tom Peeters, Stefan Hohmann, Eulàlia de Nadal, Francesc Posas, and Ricard Solé. Distributed biological computation with multicellular engineered networks. *Nature*, 469(7329):207–211, January 2011. ISSN 1476-4687. doi: 10.1038/nature09679.

[3] Jun Ishii, Shizuka Matsumura, Sakurako Kimura, Kenji Tatematsu, Shun'ichi Kuroda, Hideki Fukuda, and Akihiko Kondo. Quantitative and dynamic analyses of G protein-coupled receptor signaling in yeast using Fus1, enhanced green fluorescence protein (EGFP), and His3 fusion protein. *Biotechnol Prog*, 22(4):954–960, 2006. ISSN 8756-7938. doi: 10.1021/bp0601387.

[4] N Koloteva, P P Müller, and J E McCarthy. The position dependence of translational regulation via RNA-RNA and RNA-protein interactions in the 5'-untranslated region of eukaryotic mRNA is a function of the thermodynamic competence of 40 S ribosomes in translational initiation. *J Biol Chem*, 272:16531–16539, June 1997. ISSN 0021-9258.

[5] R Brent. Repression of transcription in yeast. *Cell*, 42:3–4, August 1985. ISSN 0092-8674.

[6] K Fujita, M Nagaoka, Y Komatsu, and H Iwahashi. Yeast pheromone signaling pathway as a bioassay to assess the effect of chemicals on mammalian peptide hormones. *Ecotoxicol Environ Saf*, 56:358–366, November 2003. URL http://view.ncbi.nlm.nih.gov/pubmed/14575675.

[7] Edith Pajot-Augy, Mark Crowe, Grégoire Levasseur, Roland Salesse, and Ian Connerton. Engineered yeasts as reporter systems for odorant detection. *J Recept Signal Transduct Res*, 23:155–171, 2003. URL http://view.ncbi.nlm.nih.gov/pubmed/14626444.

[8] Ernesto Andrianantoandro, Subhayu Basu, David K Karig, and Ron Weiss. Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol*, 2:2006.0028, 2006. URL http://view.ncbi.nlm.nih.gov/pubmed/16738572.

[9] Hans V Westerhoff, Alexey Kolodkin, Riaan Conradie, Stephen J Wilkinson, Frank J Bruggeman, Klaas Krab, Jan H van Schuppen, Hanna Hardin, Barbara M Bakker, Martijn J Moné, Katja N Rybakova, Marco Eijken, Hans J P van Leeuwen, and Jacky L Snoep. Systems biology towards life in silico: mathematics of the control of living cells. *J Math Biol*, 58(1-2):7–34, January 2009. ISSN 0303-6812. doi: 10.1007/s00285-008-0160-8. URL http://dx.doi.org/10.1007/s00285-008-0160-8.

[10] D Endy and R Brent. Modelling cellular behavior. *Nature*, 409:391–395, 2001.

[11] Drew Endy. Foundations for engineering biology. *Nature*, 438:449–453, 2005. URL http://dx.doi.org/10.1038/nature04342.

[12] Ryan McDaniel and Ron Weiss. Advances in synthetic biology: on the path from prototypes to applications. *Curr Opin Biotechnol*, 16(4):476–483, August 2005. ISSN 0958-1669. doi: 10.1016/j.copbio.2005.07.002.

[13] Tessa E Pronk, Andy D Pimentel, Marco Roos, and Timo M Breit. Taking the example of computer systems engineering for the analysis of biological cell systems. *Biosystems*, 90(3):623–635, 2007. ISSN 0303-2647. doi: 10.1016/j.biosystems.2007.02.002.

[14] Hiroaki Kitano. Systems biology: a brief overview. *Science*, 295(5560):1662–1664, March 2002. ISSN 1095-9203. doi: 10.1126/science.1069492. URL http://dx.doi.org/10.1126/science.1069492.

[15] Mario Andrea Marchisio and Fabian Rudolf. Synthetic biosensing systems. *Int J Biochem Cell Biol*, 43(3):310–319, March 2011. ISSN 1878-5875. doi: 10.1016/j.biocel.2010.11.012.

[16] M A Marchisio and J Stelling. Computational design of synthetic gene circuits with composable parts. *Bioinformatics*, 24:1903–1910, September 2008. ISSN 1460-2059. doi: 10.1093/bioinformatics/btn330.

[17] R Daniel Gietz and Robin A Woods. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol*, 350:87–96, 2002. URL http://view.ncbi.nlm.nih.gov/pubmed/12073338.

[18] Fred Sherman. Getting started with yeast. *Methods Enzymol*, 350:3–41, 2002. URL http://view.ncbi.nlm.nih.gov/pubmed/12073320.

[19] M A Romanos, C A Scorer, and J J Clare. Foreign gene expression in yeast: a review. *Yeast*, 8(6):423–488, June 1992. ISSN 0749-503X. doi: 10.1002/yea.320080602.

[20] Bernard Dujon. Yeast evolutionary genomics. *Nat Rev Genet*, 11(7):512–524, July 2010. ISSN 1471-0064. doi: 10.1038/nrg2811.

[21] A J Fijalkowski. yeast mating. Yeast Mating. Available: http://commons.wikimedia.org/wiki/File:Yeast_mating_scheme.svg Last accessed 23rd September 2011., 2006.

[22] J Dolan and S Fields. Cell-type-specific transcription in yeast. *Biochimica et Biophysica Acta*, 1088:155–169, 1991.

[23] C C Lindegren. Yeast genetics; life cycles, cytology, hybridization, vitamin synthesis, and adaptive enzymes. *Bacteriol Rev*, 9:111–170, 1945. ISSN 0005-3678.

[24] R Akada, L Kallal, D I Johnson, and J Kurjan. Genetic Relationships Between the G Protein Py Complex, Ste5p, Ste20p and Cdc42p: Investigation of Effector Roles in the Yeast Pheromone Response Pathway. *Genetics*, 143:103–117, 1996.

[25] S Erdman and M Snyder. A Filamentous Growth Response Mediated by the Yeast Mating Pathway. *Genetics*, 159:919–928, 2001.

[26] J R Strazdiz and V L MacKay. Induction of yeast mating pheromone a-factor by alpha cells. *Nature*, 305:543–545, 1983.

[27] S Fields, D T Chaleff, and G F Sprague. Yeast STE7, STE11, and STE12 genes are required for expression of cell-type-specific genes. *Mol Cell Biol*, 8:551–556, 1988. ISSN 0270-7306.

[28] H Pi, C T Chien, and S Fields. Transcriptional activation upon pheromone stimulation mediated by a small domain of Saccharomyces cerevisiae Ste12p. *Mol Cell Biol*, 17:6410–6418, 1997. ISSN 0270-7306.

[29] John Chant. Cell Polarity in Yeast. *Annu. Rev. Cell Dev. Biol*, 15:365–391, 1999.

[30] Peter Houston, Peter J Simon, and James R Broach. The Saccharomyces cerevisiae recombination enhancer biases recombination during interchromosomal mating-type switching but not in interchromosomal homologous recombination. *Genetics*, 166(3):1187–1197, March 2004. ISSN 0016-6731.

[31] H Dohlman and J Slessareva. Pheromone Signaling Pathways in Yeast. *Science Signaling*, 364, 2006.

[32] F Chang and I Herskowitz. Identification of a Gene Necessary for Cell Cycle Arrest by a Negative Growth Factor of Yeast: FAR1 Is an Inhibitor of a Gl Cyclin, CLN2. *Cell*, 63:999–1011, 1990.

[33] G M Cole, D E Stone, and S I Reed. Stoichiometry of G Protein Subunits Affects the Saccharomyces cerevisiae Mating Pheromone Signal Transduction Pathway. *Molecular and Cellular Biology*, 10:510–517, 1990.

[34] A C Butty, P M Pryciak, L S Huang, I Herskowitz, and M Peter. The role of Far1p in linking the heterotrimeric G protein to polarity establishment proteins during yeast mating. *Science*, 282:1511–1516, 1998. ISSN 0036-8075.

[35] Laura Kawasaki, Mar\'ia Castañeda Bueno, Edith Sánchez-Paredes, Nancy Velázquez-Zavala, Francisco Torres-Quiroz, Laura Ongay-Larios, and Roberto Coria. Protein kinases involved in mating and osmotic stress in the yeast Kluyveromyces lactis. *Eukaryot Cell*, 7:78–85, 2008. ISSN 1535-9786.

[36] M L Hairfield, A B Ayers, and J W Dolan. Phospholipase D1 is required for efficient mating projection formation in Saccharomyces cerevisiae. *FEMS Yeast Res*, 1:225–232, 2001. ISSN 1567-1356.

[37] Gustav Ammerer. Sex, stress and integrity: the importance of IbAP kinases in yeast. *Current Opinion in Genetics and Development*, 4:90–95, 1994.

[38] Elaine Elion. Pheromone response, mating and cell biology. *Current Opinions in Microbiology*, 3:573–581, 2000.

[39] M Gustin, J Albertyn, M Alexander, and K Davenport. MAP Kinase Pathways in the Yeast Saccharomyces cerevisiae. *Microbiol Mol Biol Rev.*, 62:1264–1300, 1998.

[40] F Posas, M Takeawa, and H Saito. Signal transduction by MAP kinase cascades in budding yeast. *Current Opinion in Microbiology*, 1:175–182, 1998.

[41] Anasua B Kusari, Douglas M Molina, Walid Sabbagh, Chang S Lau, and Lee Bardwell. A conserved protein interaction network involving the yeast MAP kinases Fus3 and Kss1. *J Cell Biol*, 164:267–277, 2004. ISSN 0021-9525.

[42] L Bardwell, Z Xiufen, N Qing, and N Komarova. Mathematical Models of SpeciïÂñï£¡city in Cell Signaling. *Biophysical Journal*, 92:3425–3441, 2007.

[43] K Madden and M Snyder. Cell Polarity and Morphogenesis in budding yeast. *Annu. Rev. Microbiol.*, 52:687–744, 1998.

[44] H Dohlman. G PROTEINS AND PHEROMONE SIGNALING. *Annu. Rev. Physiol.*, 64:129–152, 2002.

[45] Tarun B Patel. Single transmembrane spanning heterotrimeric g protein-coupled receptors and their signaling cascades. *Pharmacol Rev*, 56(3):371–385, September 2004. ISSN 0031-6997. doi: 10.1124/pr.56.3.4.

[46] Ellen C O'Shaughnessy, Santhosh Palani, James J Collins, and Casim A Sarkar. Tunable signal processing in synthetic MAP kinase cascades. *Cell*, 144(1):119–131, January 2011. ISSN 1097-4172. doi: 10.1016/j.cell.2010.12.014.

[47] A Neiman and I Herskowitz. Reconstitution of a yeast protein kinase cascade in vitro: Activation of the yeast MEK homologue STE7 by STE11. *Proc. Natl. Acad. Sci. USA*, 91:3398–3402, 1994.

[48] D Levin and B Errede. The proliferation of MAP kinase signaling pathways in yeast. *Current Opinion in Cell Biology*, 7:197–202, 1995.

[49] H Dohlman and J Thorner. REGULATION OF G PROTEIN–INITIATED SIGNAL TRANSDUCTION IN YEAST: Paradigms and Principles. *Annu. Rev. Biochem*, 70:703–754, 2001.

[50] L Bardwell. G-Protein Signaling: A New Branch in an Old Pathway. *Current Biology*, 16:53–55, 2006.

[51] F Drogen and P Matthias. MAP Kinase Cascades: Scaffolding Signal Specificity. *Current Biology*, 12:53–55, 2002.

[52] R Ballon, P Flanary, P Gladue, J Konopka, H Dohlman, and J Thomer. DEP-Domain-Mediated Regulation of GPCR Signaling Responses. *Cell*, 126:1079–1093, 2006.

[53] C E Ford, N P Skiba, H Bae, Y Daaka, E Reuveny, L R Shekter, R Rosal, G Weng, C S Yang, R Iyengar, R J Miller, L Y Jan, R J Lefkowitz, and H E Hamm. Molecular basis for interactions of G protein betagamma subunits with effectors. *Science*, 280:1271–1274, 1998. ISSN 0036-8075.

[54] J E Hirschman and D D Jenness. Dual Lipid Modification of the Yeast Gg Subunit Ste18p Determines Membrane Localization of Gbg. *MOLECULAR AND CELLULAR BIOLOGY*, 19:7705–7711, 1999.

[55] L Bardwell. A walk-through of the yeast mating pheromone response pathway. *Peptides*, 26:339–350, 2005.

[56] K C Chen, A Nagy-Csikasz, B Gyorffy, J Val, B Novak, and J J Tyson. Kinetic Analysis of a Molecular Model of the Budding Yeast Cell Cycle. *Molecular Biology of the Cell*, 11:369–391, 2000.

[57] Melinda Hauser, Sarah Kauffman, Byung-Kwon Lee, Fred Naider, and Jeffrey M Becker. The first extracellular loop of the Saccharomyces cerevisiae G protein-coupled receptor Ste2p undergoes a conformational change upon ligand binding. *J Biol Chem*, 282:10387–10397, April 2007. URL http://view.ncbi.nlm.nih.gov/pubmed/17293349.

[58] L Bardwell, J G Cook, C J Inouye, and J Thorner. Signal propagation and regulation in the mating pheromone response pathway of the yeast Saccharomyces cerevisiae. *Dev Biol*, 166:363–379, 1994. ISSN 0012-1606.

[59] G Sprague. Control of MAP kinase signaling specificity or how not to go HOG wild. *Genes Dev*, 12:2817–2820, 1998.

[60] M S Whiteway and D Y Thomas. Site-directed mutations altering the CAAX box of Ste18, the yeast pheromone-response pathway G gamma subunit. *Genetics*, 137:967–976, 1994. ISSN 0016-6731.

[61] Thompson T. and D Endy. yeast pheromone model, 2011. URL http://yeastpheromonemodel.org.

[62] M Ziman, J M O'Brien, L A Ouellette, W R Church, and D I Johnson. Mutational analysis of CDC42Sc, a Saccharomyces cerevisiae gene that encodes a putative GTP-binding protein involved in the control of cell polarity. *Mol Cell Biol*, 11: 3537–3544, 1991. ISSN 0270-7306.

[63] P M Pryciak and F A Huntress. Membrane recruitment of the kinase cascade scaffold protein Ste5 by the Gbetagamma complex underlies activation of the yeast pheromone response pathway. *Genes Dev*, 12:2684–2697, 1998. ISSN 0890-9369.

[64] K Y Choi, B Satterberg, D M Lyons, and E A Elion. Ste5 tethers multiple protein kinases in the MAP kinase cascade required for mating in S. cerevisiae. *Cell*, 78: 499–512, 1994. ISSN 0092-8674.

[65] I Herskowitz. MAP kinase pathways in yeast: for mating and more. *Cell*, 80: 187–197, 1995. ISSN 0092-8674.

[66] E Leberer, D Y Thomas, and M Whiteway. Pheromone signalling and polarized morphogenesis in yeast. *Curr Opin Genet Dev*, 7:59–66, 1997. ISSN 0959-437X.

[67] Matthew Good, Grace Tang, Julie Singleton, Attila Reményi, and Wendell A Lim. The Ste5 scaffold directs mating signaling by catalytically unlocking the Fus3 MAP kinase for activation. *Cell*, 136(6):1085–1097, March 2009. ISSN 1097-4172. doi: 10.1016/j.cell.2009.01.049.

[68] Elaine A Elion, Maosong Qi, and Weidong Chen. Signal transduction. Signaling specificity in yeast. *Science*, 307:687–688, 2005. ISSN 1095-9203.

[69] Bo Hu, Wouter-Jan Rappel, and Herbert Levine. Mechanisms and constraints on yeast MAPK signaling specificity. *Biophys J*, 96(12):4755–4763, June 2009. ISSN 1542-0086. doi: 10.1016/j.bpj.2009.02.065.

[70] J Schaber, B Kofahl, A Kowald, and E Klipp. A modelling approach to quantify dynamic crosstalk between the pheromone and the starvation pathway in bakers yeast. *FEBS Journal*, 273:3520–3533, 2006.

[71] Haruo Saito. Regulation of cross-talk in yeast MAPK signaling pathways. *Curr Opin Microbiol*, 13(6):677–683, December 2010. ISSN 1879-0364. doi: 10.1016/j.mib.2010.09.001.

[72] A Gartner, A Jovanović, D I Jeoung, S Bourlat, F R Cross, and G Ammerer. Pheromone-dependent G1 cell cycle arrest requires Far1 phosphorylation, but may not involve inhibition of Cdc28-Cln2 kinase, in vivo. *Mol Cell Biol*, 18: 3681–3691, 1998. ISSN 0270-7306.

[73] J Erickson, J Wi, G Goddard, G Tifyi, K Kawanishi, D Tomei, and C Kiefer. Edg-2/Vzg-1 Couples to the Yeast Pheromone Response Pathway Selectively in

Response to Lysophosphatidic Acid. *The Journal of Biological Chemistry*, 3: 1506–1510, 1998.

[74] L Bardwell, J G Cook, E C Chang, B R Cairns, and J Thorner. Signaling in the yeast pheromone response pathway: specific and high-affinity interaction of the mitogen-activated protein (MAP) kinases Kss1 and Fus3 with the upstream MAP kinase kinase Ste7. *Mol Cell Biol*, 16:3637–3650, 1996. ISSN 0270-7306.

[75] G C Brown, J B Hoek, and B N Kholodenko. Why do protein kinase cascades have more than one level? *Trends Biochem Sci*, 22:288, August 1997. ISSN 0968-0004.

[76] Nick I Markevich, Jan B Hoek, and Boris N Kholodenko. Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *J Cell Biol*, 164:353–359, February 2004. ISSN 0021-9525. doi: 10.1083/jcb.200308060.

[77] Fernando Ortega, José L Garcés, Francesc Mas, Boris N Kholodenko, and Marta Cascante. Bistability from double phosphorylation in signal transduction. Kinetic and structural requirements. *FEBS J*, 273(17):3915–3926, September 2006. ISSN 1742-464X. doi: 10.1111/j.1742-4658.2006.05394.x.

[78] M A Teague, D T Chaleff, and B Errede. Nucleotide sequence of the yeast regulatory gene STE7 predicts a protein homologous to protein kinases. *Proc Natl Acad Sci U S A*, 83:7371–7375, 1986. ISSN 0027-8424.

[79] R Roberts and G Fink. Elements of a single MAP kinase cascade in Saccharomyces cerevisiae mediate two developmental programs in the same cell type: mating and invasive growth. *Genes \& Dev.*, 8:2974–2985, 1994.

[80] Mohan K Malleshaiah, Vahid Shahrezaei, Peter S Swain, and Stephen W Michnick. The scaffold protein Ste5 directly controls a switch-like mating decision in yeast. *Nature*, 465(7294):101–105, May 2010. ISSN 1476-4687. doi: 10.1038/nature08946.

[81] D C Hagen, G McCaffrey, and G F Jr Sprague. Pheromone response elements are necessary and sufficient for basal and pheromone-induced transcription of the FUS1 gene of Saccharomyces cerevisiae. *Mol Cell Biol*, 11:2952–2961, June 1991. URL `http://view.ncbi.nlm.nih.gov/pubmed/1903837`.

[82] B Errede and G Ammerer. STE12, a protein involved in cell-typespecific transcription and signal transduction in yeast, is part of protein- DNA complexes. *Genes \& Dev*, 3:1349–1361, 1989.

[83] Anthony R Borneman, Tara A Gianoulis, Zhengdong D Zhang, Haiyuan Yu, Joel Rozowsky, Michael R Seringhaus, Lu Yong Wang, Mark Gerstein, and Michael Snyder. Divergence of transcription factor binding sites across related yeast species. *Science*, 317(5839):815–819, August 2007. ISSN 1095-9203. doi: 10.1126/science.1140748.

[84] Itay Tirosh, Adina Weinberger, Dana Bezalel, Mark Kaganovich, and Naama Barkai. On the relation between promoter divergence and gene expression evolution. *Mol Syst Biol*, 4:159, January 2008. ISSN 1744-4292. doi: 10.1038/msb4100198.

[85] J Roberts, N Bryce, M Marton, R Stoughton, M Meyer, H Bennett, Y He, H Dai, W Walker, T Highes, M Tyers, C Boone, and S Friend. Signaling and Circuitry ofMultiple MAPK Pathways Revealed by a Matrix of Global Gene Expression Profiles. *Science*, 287:873–880, 2000.

[86] Y O Yuan and S Fields. Properties of the DNA-Binding Domain of the Saccharomyces cerevisiae STE12 Protein. *Molecular and Cellular Biology*, 11:12:5910–5918, 1991.

[87] C Kirkman-Correia, I Stoke, and S Fields. Functional Domains of the Yeast STE12 Protein, a Pheromone-Responsive Transcriptional Activator. *Molecular and Cellular Biology*, 13:7:3765–3772, 1993.

[88] J M White and M D Rose. Yeast mating: getting close to membrane merger. *Curr Biol*, 11:R16—-20, 2001. ISSN 0960-9822.

[89] S Chou, S Lane, and H Liu. Regulation of Mating and Filamentation Genes by Two Distinct Ste12 Complexes in Saccharomyces cerevisiaeââĆň. *Molecular and Cellular Biology*, 26:4794–4805, 2006.

[90] B Ren, F Robert, J J Wyrick, O Aparicio, E G Jennings, I Simon, J Zeitlinger, J Schreiber, N Hannett, E Kanin, T L Volkert, C J Wilson, S P Bell, and R A Young. Genome-wide location and function of DNA binding proteins. *Science*, 290:2306–2309, 2000. ISSN 0036-8075.

[91] Henrik G Dohlman, Jianping Song, Donald M Apanovitch, Paul R DiBello, and Kathy M Gillen. Regulation of G protein signalling in yeast. *Seminars in Cell & Developmental Biology*, 9:135–141, April 1998. ISSN 1084-9521.

[92] H F Stratton, J Zhou, S I Reed, and D E Stone. The mating-specific G(alpha) protein of Saccharomyces cerevisiae downregulates the mating signal by a mechanism that is dependent on pheromone and independent of G(beta)(gamma) sequestration. *Mol Cell Biol*, 16:6325–6337, 1996. ISSN 0270-7306.

[93] H G Dohlman, D Apaniesk, Y Chen, J Song, and D Nusskern. Inhibition of G-protein signaling by dominant gain-of-function mutations in Sst2p, a pheromone desensitization factor in Saccharomyces cerevisiae. *Mol Cell Biol*, 15:3635–3643, July 1995. ISSN 0270-7306.

[94] W P Hausdorff, M G Caron, and R J Lefkowitz. Turning off the signal: desensitization of beta-adrenergic receptor function [published erratum appears in FASEB J 1990 Sep;4(12):3049]. *FASEB J.*, 4:2881–2889, 1990. URL `http://www.fasebj.org/cgi/content/abstract/4/11/2881`.

[95] J D Jordan, E M Landau, and R Iyengar. Signaling networks: the origins of cellular multitasking. *Cell*, 103:193–200, October 2000. ISSN 0092-8674.

[96] Y Feng and N G Davis. Feedback phosphorylation of the yeast a-factor receptor requires activation of the downstream signaling pathway from G protein through mitogen-activated protein kinase. *Mol Cell Biol*, 20:563–574, January 2000.

[97] J Slessareva, S Routt, B Temple, V Bankaitis, and H Dohlman. Activation of the Phosphatidylinositol 3-Kinase Vps34 by a G Protein a Subunit at the Endosome. *Cell*, 126:191–203, 2006.

[98] R K Chan and C A Otte. Isolation and genetic analysis of Saccharomyces cerevisiae mutants supersensitive to G1 arrest by a factor and alpha factor pheromones. *Mol Cell Biol*, 2:11–20, January 1982. ISSN 0270-7306.

[99] R K Chan and C A Otte. Physiological characterization of Saccharomyces cerevisiae mutants supersensitive to G1 arrest by a factor and alpha factor pheromones. *Mol Cell Biol*, 2:21–29, January 1982. ISSN 0270-7306.

[100] V L MacKay, S K Welch, M Y Insley, T R Manney, J Holly, G C Saari, and M L Parker. The Saccharomyces cerevisiae BAR1 gene encodes an exported protein with homology to pepsin. *Proc Natl Acad Sci U S A*, 85:55–59, January 1988. ISSN 0027-8424.

[101] G M Cole and S I Reed. Pheromone-induced phosphorylation of a G protein beta subunit in S. cerevisiae is associated with an adaptive response to mating pheromone. *Cell*, 64:703–716, February 1991. ISSN 0092-8674.

[102] K A Schandel and D D Jenness. Direct evidence for ligand-induced internalization of the yeast alpha-factor pheromone receptor. *Mol Cell Biol*, 14: 7245–7255, November 1994.

[103] T Wieland and C K Chen. Regulators of G-protein signalling: a novel protein family involved in timely deactivation and desensitization of signalling via heterotrimeric G proteins. *Naunyn Schmiedebergs Arch Pharmacol*, 360:14–26, July 1999. ISSN 0028-1298.

[104] Janna E Slessareva and Henrik G Dohlman. G protein signaling in yeast: new components, new connections, new compartments. *Science*, 314(5804):1412–1413, December 2006. ISSN 1095-9203. doi: 10.1126/science.1134041.

[105] J Kurjan, J P Hirsch, and C Dietzel. Mutations in the guanine nucleotide-binding domains of a yeast G alpha protein confer a constitutive or uninducible state to

the pheromone response pathway. *Genes Dev*, 5:475–483, March 1991. ISSN 0890-9369.

[106] Ming Guo, Christopher Aston, Scott A Burchett, Christine Dyke, Stanley Fields, S Johannes R Rajarao, Peter Uetz, Yuqi Wang, Kathleen Young, and Henrik G Dohlman. The yeast G protein alpha subunit Gpa1 transmits a signal through an RNA binding effector protein Scp160. *Mol Cell*, 12:517–524, August 2003. ISSN 1097-2765.

[107] Ernest Blackwell, Izabel M Halatek, Hye-Jin N Kim, Alexis T Ellicott, Andrey A Obukhov, and David E Stone. Effect of the pheromone-responsive G(alpha) and phosphatase proteins of Saccharomyces cerevisiae on the subcellular localization of the Fus3 mitogen-activated protein kinase. *Mol Cell Biol*, 23:1135–1150, February 2003.

[108] K Doi, A Gartner, G Ammerer, B Errede, H Shinkawa, K Sugimoto, and K Matsumoto. MSG5, a novel protein phosphatase promotes adaptation to pheromone response in S. cerevisiae. *EMBO J*, 13:61–70, January 1994. ISSN 0261-4189.

[109] Jianlong Zhou, Michelle Arora, and David Stone. The yeast pheromone-responsive GÎ\textpm protein stimulates recovery from chronic pheromone treatment by two mechanisms that are activated at distinct levels of stimulus. *Cell Biochemistry and Biophysics*, 30:193–212, June 1999. URL `http://dx.doi.org/10.1007/BF02738067`.

[110] Metodi V Metodiev, Dina Matheos, Mark D Rose, and David E Stone. Regulation of MAPK function by direct interaction with the mating-specific Galpha in yeast. *Science*, 296:1483–1486, 2002. ISSN 1095-9203.

[111] Danying Shao, Wen Zheng, Wenjun Qiu, Qi Ouyang, and Chao Tang. Dynamic studies of scaffold-dependent mating pathway in yeast. *Biophys J*, 91:3986–4001, 2006. ISSN 0006-3495.

[112] Frank J Bruggeman and Hans V Westerhoff. The nature of systems biology. *Trends Microbiol*, 15(1):45–50, January 2007. ISSN 0966-842X. doi: 10.1016/j.tim.2006.11.003. URL http://dx.doi.org/10.1016/j.tim.2006.11.003.

[113] Frank J Bruggeman, Jorrit J Hornberg, Fred C Boogerd, and Hans V Westerhoff. Introduction to systems biology. *EXS*, 97:1–19, 2007.

[114] Alain Friboulet and Daniel Thomas. Systems Biology-an interdisciplinary approach. *Biosens Bioelectron*, 20(12):2404–2407, June 2005. ISSN 0956-5663. doi: 10.1016/j.bios.2004.11.014. URL http://dx.doi.org/10.1016/j.bios.2004.11.014.

[115] John J Tyson, Katherine C Chen, and Bela Novak. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr Opin Cell Biol*, 15:221–231, April 2003. ISSN 0955-0674.

[116] A L Hodgkin and A F Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*, 117(4):500–544, August 1952. ISSN 0022-3751.

[117] Denis Noble. Systems biology and the heart. *Biosystems*, 83(2-3):75–80, 2006. ISSN 0303-2647. doi: 10.1016/j.biosystems.2005.05.013.

[118] Patrick B Warren and Pieter Rein ten Wolde. Chemical models of genetic toggle switches. *J Phys Chem B*, 109(14):6812–6823, April 2005. ISSN 1520-6106. doi: 10.1021/jp045523y.

[119] H V Westerhoff. Signalling control strength. *Journal of Theoretical Biology*, 252:555–567, 2008.

[120] Xiao Wang, Nan Hao, Henrik G Dohlman, and Timothy C Elston. Bistability, stochasticity, and oscillations in the mitogen-activated protein kinase cascade. *Biophys J*, 90:1961–1978, 2006. ISSN 0006-3495.

[121] R C Yu, O Resnekov, A P Abola, S S Andrews, K R Benjamin, J Bruck, I E Burbulis, A Colman-Lerner, D Endy, A Gordon, M Holl, L Lok, C G Pesce, E Serra, R D Smith, T M Thomson, A E Tsong, and R Brent. The Alpha

Project: a model system for systems biology research. *IET Syst Biol*, 2(5):222–233, September 2008. ISSN 1751-8849. doi: 10.1049/iet-syb:20080127. URL http://dx.doi.org/10.1049/iet-syb:20080127.

[122] C Huang and J E Ferrell. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc. Natl. Acad. Sci.*, 93:10078–10083, 1996.

[123] A Goldbeter and D E Koshland. Ultrasensitivity in biochemical systems controlled by covalent modification. Interplay between zero-order and multistep effects. *J Biol Chem*, 259(23):14441–14447, December 1984. ISSN 0021-9258.

[124] B N Kholodenko. Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *Eur J Biochem*, 267:1583–1588, March 2000. ISSN 0014-2956.

[125] Bluthgen N. and Hanspeter H. MAP-Kinase-Cascade: Switch, Amplifier or Feedback Controller? *2nd Workshop on Computation of Biochemical Pathways and Genetic Networks*, Logos-Verl:55–62, 2001.

[126] T Yi, H Kitano, and M I Simon. A quantitative characterization of the yeast heterotrimeric G protein cycle. *PNAS*, 19:10764–10769, 2003.

[127] Nan Hao, Necmettin Yildirim, Yuqi Wang, Timothy C Elston, and Henrik G Dohlman. Regulators of G protein signaling and transient activation of signaling: experimental and computational analysis reveals negative and positive feedback controls on G protein activity. *J Biol Chem*, 278(47):46506–46515, November 2003. ISSN 0021-9258. doi: 10.1074/jbc.M308432200. URL http://dx.doi.org/10.1074/jbc.M308432200.

[128] Bente Kofahl and Edda Klipp. Modelling the dynamics of the yeast pheromone pathway. *Yeast*, 21:831–850, July 2004. URL http://view.ncbi.nlm.nih.gov/pubmed/15300679.

[129] A Goldbeter and L Wolpert. Covalent modification of proteins as a threshold mechanism in development. *J Theor Biol*, 142(2):243–250, January 1990. ISSN 0022-5193.

[130] A Goldbeter and D E Koshland. An amplified sensitivity arising from covalent modification in biological systems. *Proc Natl Acad Sci U S A*, 78(11):6840–6844, November 1981. ISSN 0027-8424.

[131] Boris N Kholodenko and Marc R Birtwistle. Four-dimensional dynamics of MAPK information processing systems. *Wiley Interdiscip Rev Syst Biol Med*, 1(1):28–44, 2009. ISSN 1939-005X. doi: 10.1002/wsbm.16.

[132] James E Ferrell. How responses get more switch-like as you move down a protein kinase cascade. *Trends in Biochemical Sciences*, 22(8):288–289, 1997. ISSN 0968-0004. doi: DOI:10.1016/S0968-0004(97)82217-7.

[133] Eils R. and Kriete A. *Computational systems biology*. Elsevier science and technology, May 2005.

[134] Stefan Hoops, Sven Sahle, Ralph Gauges, Christine Lee, Jürgen Pahle, Natalia Simus, Mudita Singhal, Liang Xu, Pedro Mendes, and Ursula Kummer. COPASI–a COmplex PAthway SImulator. *Bioinformatics*, 22(24):3067–3074, December 2006. ISSN 1460-2059. doi: 10.1093/bioinformatics/btl485.

[135] Mario A Marchisio and Jörg Stelling. Computational design tools for synthetic biology. *Curr Opin Biotechnol*, 20(4):479–485, August 2009. ISSN 1879-0429. doi: 10.1016/j.copbio.2009.08.007.

[136] Bard Ermentrout. *Simulating, Analyzing, and Animating Dynamical Systems: A Guide to XPPAUT for Researchers and Students (Software, Environments and Tools)*. Society for Industrial Mathematic (siam), 1987.

[137] J Nowacki. XPPy, 2011. URL `http://seis.bris.ac.uk/~enxjn/xppy/`.

[138] Paul Shannon, Andrew Markiel, Owen Ozier, Nitin S Baliga, Jonathan T Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, 13(11):2498–2504, November 2003. ISSN 1088-9051. doi: 10.1101/gr.1239303.

[139] A Quarteroni and S Fausto. *Scientific Computing with Matlab and Octave*. Springer, 2006.

[140] Sarah M Keating, Benjamin J Bornstein, Andrew Finney, and Michael Hucka. SBMLToolbox: an SBML toolbox for MATLAB users. *Bioinformatics*, 22(10): 1275–1277, May 2006. ISSN 1367-4803. doi: 10.1093/bioinformatics/btl111.

[141] Bruce E Shapiro, Michael Hucka, Andrew Finney, and John Doyle. MathSBML: a package for manipulating SBML-based biological models. *Bioinformatics*, 20(16):2829–2831, November 2004. ISSN 1367-4803. doi: 10.1093/bioinformatics/bth271.

[142] G V Rossum and J D Boer. Interactively Testing Remote Servers Using the Python Programming Language. *CWI Quarterly*, 4(4):283–303, 1991.

[143] H P Langtangen. *A Primer on Scientific Programming with Python*. Springer, 1st edition, 2009.

[144] A B Downey. *Python for Software Design*. 2009.

[145] M Hucka, A Finney, H M Sauro, H Bolouri, J C Doyle, H Kitano, A P Arkin, B J Bornstein, D Bray, A Cornish-Bowden, A A Cuellar, S Dronov, E D Gilles, M Ginkel, V Gor, I I Goryanin, W J Hedley, T C Hodgman, J-H Hofmeyr, P J Hunter, N S Juty, J L Kasberger, A Kremling, U Kummer, N Le Novère, L M Loew, D Lucio, P Mendes, E Minch, E D Mjolsness, Y Nakayama, M R Nelson, P F Nielsen, T Sakurada, J C Schaff, B E Shapiro, T S Shimizu, H D Spence, J Stelling, K Takahashi, M Tomita, J Wagner, and J Wang. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, March 2003. ISSN 1367-4803.

[146] A Finney and M Hucka. Systems biology markup language: Level 2 and beyond. *Biochem Soc Trans*, 31(Pt 6):1472–1473, December 2003. ISSN 0300-5127. doi: 10.1042/.

[147] Brett G Olivier, Johann M Rohwer, and Jan-Hendrik S Hofmeyr. Modelling cellular systems with PySCeS. *Bioinformatics*, 21(4):560–561, February 2005. ISSN 1367-4803. doi: 10.1093/bioinformatics/bti046.

[148] P Mendes, S Hoops, S Sahle, R Gauges, J Dada, and U Kummer. Systems Biology. *Methods in Molecular Biology*, 500(2):17–59, 2009.

[149] H Kacser and J A Burns. The control of flux. *Symposia of the Society for Experimental Biology*, 27:65–104, 1973.

[150] J A Burns, A Cornish-Bowden, A K Groen, R Heinrich, H Kacser, J W Porteous, S M Rapoport, T A Rapoport, J W Stucki, J M Tager, R J A Wanders, and H V Westerhoff. Control analysis of metabolic systems. *Trends in Biochemical Sciences*, 10(1):16, 1985. ISSN 0968-0004. doi: DOI:10.1016/0968-0004(85)90008-8.

[151] R Heinrich, B G Neel, and T A Rapoport. Mathematical Models of Protein Kinase Signal Transduction. *Molecular Cell*, 9:957–970, 2002.

[152] C Giersch. Control analysis of metabolic networks. 1. Homogeneous functions and the summation theorems for control coefficients. *Eur J Biochem*, 174(3):509–513, June 1988. ISSN 0014-2956.

[153] P Mendes, S Hoops, S Sahle, R Gauges, J Dada, and U Kummer. Computational Modeling of Biochemical Networks Using COPASI. *Methods Mol. Biol.*, 500:17–59, 2009.

[154] Pedro Mendes, Wei Sha, and Keying Ye. Artificial gene networks for objective comparison of analysis algorithms. *Bioinformatics*, 19 Suppl 2:ii122—-ii129, October 2003. ISSN 1367-4811.

[155] Robert Hooke and T A Jeeves. Direct Search Solution of Numerical and Statistical Problems. *J. ACM*, 8(2):212–229, April 1961. ISSN 0004-5411.

[156] S Kirkpatrick, C D Gelatt, and M P Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598):pp. 671–680, 1983. ISSN 00368075. URL `http://www.jstor.org/stable/1690046`.

[157] Uri Alon. *Introduction to Systems Biology*. Chapman and Hall / CRC, 2007.

[158] Chris. Voigt. *Synthetic Biology: Methods for Part/Device Characterization and Chassis*. Academic Press, 2011.

[159] Michael B Elowitz, Arnold J Levine, Eric D Siggia, and Peter S Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, August 2002. ISSN 1095-9203. doi: 10.1126/science.1070919.

[160] Peter S Swain, Michael B Elowitz, and Eric D Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci U S A*, 99(20):12795–12800, October 2002. ISSN 0027-8424. doi: 10.1073/pnas.162041399.

[161] Priscilla E M Purnick and Ron Weiss. The second wave of synthetic biology: from modules to systems. *Nat Rev Mol Cell Biol*, 10(6):410–422, June 2009. ISSN 1471-0080. doi: 10.1038/nrm2698.

[162] S A Benner and A M Sismour. Synthetic Biology. *Nature Reviews Genetics*, 6: 533–543, 2005.

[163] Walter Kolch, Muffy Calder, and David Gilbert. When kinases meet mathematics: the systems biology of MAPK signalling. *FEBS Lett*, 579(8): 1891–1895, March 2005. ISSN 0014-5793. doi: 10.1016/j.febslet.2005.02.002.

[164] The Royal Academy of Engineering. *Synthetic Biology: scope, applications and implications*. The Royal Academy of Engineering, May 2009. URL `www.raeng.org.uk/synbio`.

[165] M Heinemann and S Panke. Synthetic biologyââĆňâĂİputting engineering into biology. *Bioinformatics*, 22:2790–2799, 2006.

[166] P Marguet, F Balagadde, and C Tan. Biology by design: reduction and synthesis of cellular components and behaviour. *J. R. Soc. Interface*, 4:607–623, 2007.

[167] Szybalski Waclaw. *In Vivo and in Vitro Initiation of Transcription*, chapter Control of. New York: Plenum Press, 1974.

[168] D A Drubin, J C Way, and P A Silver. Designing biological systems. *Genes & Dev*, 21:242–254, 2007.

[169] V D Lorenzo and A Danchin. Synthetic biology: discovering new worlds and new words. *EMBO reports*, 9:9:822–827, 2008.

[170] R Brent. A partnership between biology and engineering. *Nature Biotechnology*, 22:1211–1214, 2004.

[171] Nicholas T Ingolia and Jonathan S Weissman. Systems biology: Reverse engineering the cell. *Nature*, 454(7208):1059–1062, August 2008. ISSN 1476-4687. doi: 10.1038/4541059a.

[172] Silvia D M Santos and James E Ferrell. Systems biology: On the cell cycle and its switches. *Nature*, 454(7202):288–289, July 2008. ISSN 1476-4687. doi: 10.1038/454288a.

[173] L Serrano. Synthetic biology: promises and challenges. *Molecular Systems Biology*, 3:158, 2007.

[174] S A Benner. Act natural. *Nature*, 421:118, 2003.

[175] Michael L Simpson. Rewiring the cell: synthetic biology moves towards higher functional complexity. *Trends Biotechnol*, 22(11):555–557, November 2004. ISSN 0167-7799. doi: 10.1016/j.tibtech.2004.09.003.

[176] Barry Canton, Anna Labno, and Drew Endy. Refinement and standardization of synthetic biological parts and devices. *Nat Biotechnol*, 26(7):787–793, July 2008. ISSN 1546-1696. doi: 10.1038/nbt1413.

[177] J Peccoud, M F Blauvert, Y Cai, K L Cooper, O Crasta, E C DeLalla, C Evans, O Folkerts, B M Lyons, S P Mane, R Shelton, M A Sweede, and S A Waldon. Targeted Development of Registries of Biological Parts. *PLoS ONE*, 3:7:e2671, 2008.

[178] J Stricker, S Cookson, M R Bennett, W H Mather, L S Tsimring, and J Hasty. A fast, robust and tunable synthetic gene oscillator. *nature letters*, 456:516–520, 2008.

[179] M B Elowitz and S Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, January 2000. ISSN 0028-0836. doi: 10.1038/35002125.

[180] Jeff Hasty, David McMillen, and J J Collins. Engineered gene circuits. *Nature*, 420:224–230, 2002. URL `http://dx.doi.org/10.1038/nature01257`.

[181] Chris Grilly, Jesse Stricker, Wyming Lee Pang, Matthew R Bennett, and Jeff Hasty. A synthetic gene network for tuning protein degradation in Saccharomyces cerevisiae. *Mol Syst Biol*, 3:127, July 2007. ISSN 1744-4292. doi: 10.1038/msb4100168.

[182] T S Bayer and C D Smolke. Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nature Biotechnology*, 23:3:337–343, 2005.

[183] M N Win and C D Smolke. Higher-Order Cellular Information Processing with Synthetic RNA Devices. *Science*, 322:456–460, 2008.

[184] N Rosenfeld, J W Young, U Alon, P S Swain, and M B Elowitz. Accurate prediction of gene feedback circuit behavior from component properties. *Molecular Systems Biology*, 3:143, 2007.

[185] S Hooshangi, S Thiberge, and R Weiss. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *PNAS*, 102:3581–3586, 2005.

[186] N Barkai and S Leibler. Circadian clocks limited by noise. *Nature*, 403(6767): 267–268, January 2000. ISSN 0028-0836. doi: 10.1038/35002258.

[187] Tony Yu-Chen Tsai, Yoon Sup Choi, Wenzhe Ma, Joseph R Pomerening, Chao Tang, and James E Ferrell. Robust, tunable biological oscillations from interlinked positive and negative feedback loops. *Science*, 321(5885):126–129, July 2008. ISSN 1095-9203. doi: 10.1126/science.1156951.

[188] T S Gardner, C R Cantor, and J J Collins. Construction of a genetic toggle switch in Escherichia coli. *Nature*, 403:339–342, 2000.

[189] A Becskei, B Seraphin, and L Serrano. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *The EMBO Journal*, 20:2528–2535, 2001.

[190] Caroline M Ajo-Franklin, David A Drubin, Julian A Eskin, Elaine P S Gee, Dirk Landgraf, Ira Phillips, and Pamela A Silver. Rational design of memory in eukaryotic cells. *Genes Dev*, 21(18):2271–2276, 2007. ISSN 0890-9369. doi: 10.1101/gad.1586107.

[191] S Basu, R Mehreja, S Thiberge, M Chen, and R Weiss. Spatiotemporal control of gene expression with pulse-generating networks. *PNAS*, 101:17:6355–6360, 2004.

[192] Parliamentary Office of Science and Technology. Synthetic Biology. *Postnote*, number 298, 2008.

[193] Vincent J J Martin, Douglas J Pitera, Sydnor T Withers, Jack D Newman, and Jay D Keasling. Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nat Biotechnol*, 21(7):796–802, July 2003. ISSN 1087-0156. doi: 10.1038/nbt833.

[194] J Christopher Anderson, Elizabeth J Clarke, Adam P Arkin, and Christopher A Voigt. Environmentally controlled invasion of cancer cells by engineered bacteria. *J Mol Biol*, 355(4):619–627, January 2006. ISSN 0022-2836. doi: 10.1016/j.jmb.2005.10.076.

[195] Travis I Moore, Ching-Shan Chou, Qing Nie, Noo Li Jeon, and Tau-Mu Yi. Robust spatial sensing of mating pheromone gradients by yeast cells. *PLoS One*, 3(12):e3865, December 2008. ISSN 1932-6203. doi: 10.1371/journal.pone.0003865.

[196] F van Drogen and M Peter. MAP kinase dynamics in yeast. *Biol Cell*, 93:63–70, September 2001.

[197] Frank. Boogerd, Frank J Bruggeman, Jan-Hendrik S Hofmeyr, and Hans V Westerhoff. *Systems Biology: Philosophical Foundations*. Elsevier science and technology, 1st edition, May 2007.

[198] Ferric C Fang and Arturo Casadevall. Reductionistic and holistic science. *Infect Immun*, 79(4):1401–1404, April 2011. ISSN 1098-5522. doi: 10.1128/IAI.01343-10.

[199] J Christopher Anderson, John E Dueber, Mariana Leguia, Gabriel C Wu, Jonathan A Goler, Adam P Arkin, and Jay D Keasling. BglBricks: A flexible standard for biological part assembly. *J Biol Eng*, 4(1):1, January 2010. ISSN 1754-1611. doi: 10.1186/1754-1611-4-1.

[200] Christina D Smolke. Building outside of the box: iGEM and the BioBricks Foundation. *Nat Biotechnol*, 27(12):1099–1102, December 2009. ISSN 1546-1696. doi: 10.1038/nbt1209-1099.

[201] David Greber and Martin Fussenegger. Mammalian synthetic biology: engineering of sophisticated gene networks. *J Biotechnol*, 130(4):329–345, July 2007. ISSN 0168-1656. doi: 10.1016/j.jbiotec.2007.05.014.

[202] Daniel G Gibson, John I Glass, Carole Lartigue, Vladimir N Noskov, Ray-Yuan Chuang, Mikkel A Algire, Gwynedd A Benders, Michael G Montague, Li Ma, Monzia M Moodie, Chuck Merryman, Sanjay Vashee, Radha Krishnakumar, Nacyra Assad-Garcia, Cynthia Andrews-Pfannkoch, Evgeniya A Denisova, Lei Young, Zhi-Qing Qi, Thomas H Segall-Shapiro, Christopher H Calvey, Prashanth P Parmar, Clyde A Hutchison, Hamilton O Smith, and J Craig Venter. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, 329(5987):52–56, July 2010. ISSN 1095-9203. doi: 10.1126/science.1190719.

[203] Roger Brent and Mark Ptashne. A bacterial repressor protein or a yeast transcriptional terminator can block upstream activation of a yeast gene. *Nature*, 312:612–615, December 1984. URL `http://dx.doi.org/10.1038/312612a0`.

[204] J M Vieites, F Navarro-Garc\'ia, R Pérez-D\'iaz, J Pla, and C Nombela. Expression and in vivo determination of firefly luciferase as gene reporter in Saccharomyces cerevisiae. *Yeast*, 10:1321–1327, October 1994. ISSN 0749-503X. doi: 10.1002/yea.320101009.

[205] K Pantopoulos, N K Gray, and M W Hentze. Differential regulation of two related RNA-binding proteins, iron regulatory protein (IRP) and IRPB. *RNA*, 1 (2):155–163, April 1995. ISSN 1355-8382.

[206] Jeremy B Goforth, Sheila A Anderson, Christopher P Nizzi, and Richard S Eisenstein. Multiple determinants within iron-responsive elements dictate iron regulatory protein binding and regulatory hierarchy. *RNA*, 16(1):154–169, January 2010. ISSN 1469-9001. doi: 10.1261/rna.1857210.

[207] Alessandro Campanella, Sonia Levi, Gaetano Cairo, Giorgio Biasiotto, and Paolo Arosio. Blotting analysis of native IRP1: a novel approach to distinguish the different forms of IRP1 in cells and tissues. *Biochemistry*, 43(1):195–204, January 2004. ISSN 0006-2960. doi: 10.1021/bi035386f.

[208] J R de Wet, K V Wood, M DeLuca, D R Helinski, and S Subramani. Firefly luciferase gene: structure and expression in mammalian cells. *Mol Cell Biol*, 7 (2):725–737, February 1987. ISSN 0270-7306.

[209] J R de Wet, K V Wood, D R Helinski, and M DeLuca. Cloning of firefly luciferase cDNA and the expression of active luciferase in Escherichia coli. *Proc Natl Acad Sci U S A*, 82(23):7870–7873, December 1985. ISSN 0027-8424.

[210] E A Meighen. Molecular biology of bacterial bioluminescence. *Microbiol Rev*, 55(1):123–142, March 1991. ISSN 0146-0749.

[211] J Trueheart, J D Boeke, and G R Fink. Two genes required for cell fusion during yeast conjugation: evidence for a pheromone-induced surface protein. *Mol Cell Biol*, 7(7):2316–2328, July 1987. ISSN 0270-7306.

[212] Tobias von der Haar. A quantitative estimation of the global translational activity in logarithmically growing yeast cells. *BMC Syst Biol*, 2:87, October 2008. ISSN 1752-0509. doi: 10.1186/1752-0509-2-87.

[213] S P Sikorski and P Hieter. A System of Shuttle Vectors and Yeast Host Strains Designed for Efficient Manipulation of DNA in Saccharomyces ceratisiae. *Genetics*, 122:19–27, 1989.

[214] C C Oliveira, J J van den Heuvel, and J E McCarthy. Inhibition of translational initiation in Saccharomyces cerevisiae by secondary structure: the roles of the stability and position of stem-loops in the mRNA leader. *Mol Microbiol*, 9:521–532, August 1993. ISSN 0950-382X.

[215] C C Oliveira, B Goossen, N I Zanchin, J E McCarthy, M W Hentze, and R Stripecke. Translational repression by the human iron-regulatory factor (IRF) in Saccharomyces cerevisiae. *Nucleic acids research*, 21(23):5316–22, November 1993. ISSN 0305-1048.

[216] C Mateus and S V Avery. Destabilized green fluorescent protein for monitoring dynamic changes in yeast gene expression with flow cytometry. *Yeast*, 16(14): 1313–1323, October 2000. ISSN 0749-503X. doi: 10.1002/1097-0061(200010) 16:14<1313::AID-YEA626>3.0.CO;2-O.

[217] C B Brachmann, A Davies, G J Cost, E Caputo, J Li, P Hieter, and J D Boeke. Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*, 14(2):115–132, January 1998. ISSN 0749-503X. doi: 10.1002/(SICI)1097-0061(19980130)14:2<115::AID-YEA204>3.0.CO;2-2.

[218] D B Clewell and D R Helinski. Supercoiled circular DNA-protein complex in Escherichia coli: purification and induced conversion to an opern circular DNA form. *Proc Natl Acad Sci U S A*, 62:1159–1166, April 1969. ISSN 0027-8424.

[219] Sambrook J. and Russel D W. *Molecular Cloning - A laboratory Manual*. Cold Spring Harbor Laboratory Press, 2001.

[220] R M Cranenburgh. An equation for calculating the volumetric ratios required in a ligation reaction. *Appl Microbiol Biotechnol*, 65(2):200–202, August 2004. ISSN 0175-7598. doi: 10.1007/s00253-004-1577-7.

[221] U Guldener, S Heck, T Fielder, J Beinhauer, and J H Hegemann. A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res*, 24:2519–2524, July 1996. ISSN 0305-1048.

[222] Tobias von der Haar. Optimized protein extraction for quantitative proteomics of yeasts. *PLoS One*, 2(10):e1078, October 2007. ISSN 1932-6203. doi: 10.1371/journal.pone.0001078.

[223] M A Larkin, G Blackshields, N P Brown, R Chenna, P A McGettigan, H McWilliam, F Valentin, I M Wallace, A Wilm, R Lopez, J D Thompson, T J Gibson, and D G Higgins. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21):2947–2948, November 2007. ISSN 1367-4811. doi: 10.1093/bioinformatics/btm404.

[224] Nicolas LeNovre, Benjamin Bornstein, Alexander Broicher, Mélanie Courtot, Marco Donizelli, Harish Dharuri, Lu Li, Herbert Sauro, Maria Schilstra, Bruce Shapiro, Jacky L Snoep, and Michael Hucka. BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Res*, 34:D689—-D691, January 2006. ISSN 1362-4962. doi: 10.1093/nar/gkj092.

[225] R Milo, S Shen-Orr, S Itzkovitz, N Kashtan, D Chklovskii, and U Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, October 2002. ISSN 1095-9203. doi: 10.1126/science.298.5594.824.

[226] Stephen A Chapman and Anand R Asthagiri. Quantitative effect of scaffold abundance on signal propagation. *Mol Syst Biol*, 5:313, October 2009. ISSN 1744-4292. doi: 10.1038/msb.2009.73.

[227] Roger Brent. Cell signaling: what is the signal and what information does it carry? *FEBS Lett*, 583(24):4019–4024, December 2009. ISSN 1873-3468. doi: 10.1016/j.febslet.2009.11.029.

[228] Jian Wang and Kostas Pantopoulos. Regulation of cellular iron metabolism. *Biochem J*, 434(3):365–381, February 2011. ISSN 1470-8728. doi: 10.1042/BJ20101825.

[229] Tracey A Rouault. The role of iron regulatory proteins in mammalian iron homeostasis and disease. *Nat Chem Biol*, 2(8):406–414, August 2006. ISSN 1552-4450. doi: 10.1038/nchembio807.

[230] Jérôme Dupuy, Anne Volbeda, Philippe Carpentier, Claudine Darnault, Jean-Marc Moulis, and Juan Carlos Fontecilla-Camps. Crystal structure of human iron regulatory protein 1 as cytosolic aconitase. *Structure*, 14(1):129–139, January 2006. ISSN 0969-2126. doi: 10.1016/j.str.2005.09.009.

[231] R S Eisenstein. Iron regulatory proteins and the molecular control of mammalian iron metabolism. *Annu Rev Nutr*, 20:627–662, 2000. ISSN 0199-9885. doi: 10.1146/annurev.nutr.20.1.627.

[232] William E Walden, Anna I Selezneva, Jérôme Dupuy, Anne Volbeda, Juan C Fontecilla-Camps, Elizabeth C Theil, and Karl Volz. Structure of dual function iron regulatory protein 1 complexed with ferritin IRE-RNA. *Science*, 314(5807):1903–1908, December 2006. ISSN 1095-9203. doi: 10.1126/science.1133116.

[233] J W Little, D W Mount, and C R Yanisch-Perron. Purified lexA protein is a repressor of the recA and lexA genes. *Proc Natl Acad Sci U S A*, 78:4199–4203, July 1981. ISSN 0027-8424.

[234] O G Berg. Selection of DNA binding sites by regulatory proteins: the LexA protein and the arginine repressor use different strategies for functional specificity. *Nucleic Acids Res*, 16(11):5089–5105, June 1988. ISSN 0305-1048.

[235] M Butala, D Zgur-Bertok, and S J W Busby. The bacterial LexA transcriptional repressor. *Cell Mol Life Sci*, 66(1):82–93, January 2009. ISSN 1420-9071. doi: 10.1007/s00018-008-8378-6.

[236] K F Wertman and D W Mount. Nucleotide sequence binding specificity of the LexA repressor of Escherichia coli K-12. *J Bacteriol*, 163:376–384, July 1985. ISSN 0021-9193.

[237] G M Smith, K A Mileham, S E Cooke, S J Woolston, H K George, A D Charles, and W J Brammar. The Escherichia coli LexA repressor-operator system works in mammalian cells. *EMBO J*, 7:3975–3982, December 1988. ISSN 0261-4189.

[238] Roger Brent and Mark Ptashne. Mechanism of action of the lexA gene product. *Proc. Natl. Acad. Sci, USA*, 78:4204–4208, 1981.

[239] R Brent. Regulation and autoregulation by lexA protein. *Biochimie*, 64:565–569, 1982. ISSN 0300-9084.

[240] Adrianna P P Zhang, Ying Z Pigli, and Phoebe A Rice. Structure of the LexA-DNA complex and implications for SOS box measurement. *Nature*, 466(7308): 883–886, August 2010. ISSN 1476-4687. doi: 10.1038/nature09200.

[241] J Zhu and M Q Zhang. SCPD: a promoter database of the yeast Saccharomyces cerevisiae. *Bioinformatics*, 15(7-8):607–611, 1999. ISSN 1367-4803.

[242] E M McIntosh and R H Haynes. Sequence and expression of the dCMP deaminase gene (DCD1) of Saccharomyces cerevisiae. *Mol Cell Biol*, 6(5):1711–1721, May 1986. ISSN 0270-7306.

[243] J Mellor, C Midgely, A J Kingsman, S M Kingsman, and S Kim. Transcriptional activation by upstream activator sequences requires distinct interactions with downstream elements in the yeast TRP1 promoter. *Mol Gen Genet*, 225:217–224, February 1991. ISSN 0026-8925.

[244] Elke Nevoigt, Jessica Kohnke, Curt R Fischer, Hal Alper, Ulf Stahl, and Gregory Stephanopoulos. Engineering of promoter replacement cassettes for fine-tuning of gene expression in Saccharomyces cerevisiae. *Appl Environ Microbiol*, 72(8): 5266–5273, August 2006. ISSN 0099-2240. doi: 10.1128/AEM.00530-06.

[245] Ertugrul M Ozbudak, Mukund Thattai, Iren Kurtser, Alan D Grossman, and Alexander van Oudenaarden. Regulation of noise in the expression of a single gene. *Nat Genet*, 31(1):69–73, May 2002. ISSN 1061-4036. doi: 10.1038/ng869.

[246] Jonathan M Raser and Erin K O'Shea. Noise in gene expression: origins, consequences, and control. *Science*, 309(5743):2010–2013, September 2005. ISSN 1095-9203. doi: 10.1126/science.1105891.

[247] H H McAdams and A Arkin. It's a noisy business! Genetic regulation at the nanomolar scale. *Trends Genet*, 15(2):65–69, February 1999. ISSN 0168-9525.

[248] S R Salama, K B Hendricks, and J Thorner. G1 cyclin degradation: the PEST motif of yeast Cln2 is necessary, but not sufficient, for rapid protein turnover. *Mol Cell Biol*, 14(12):7953–7966, December 1994. ISSN 0270-7306.

[249] M Rechsteiner and S W Rogers. PEST sequences and regulation by proteolysis. *Trends Biochem Sci*, 21(7):267–271, July 1996. ISSN 0968-0004.

[250] A Varshavsky. Ubiquitin fusion technique and its descendants. *Methods Enzymol*, 327:578–593, 2000. ISSN 0076-6879.

[251] L Oliveira and J C Drapier. Down-regulation of iron regulatory protein 1 gene expression by nitric oxide. *Proc Natl Acad Sci U S A*, 97(12):6550–6555, June 2000. ISSN 0027-8424. doi: 10.1073/pnas.120571797.

[252] Stephen L Clarke, Aparna Vasanthakumar, Sheila A Anderson, Corinne Pondarre, Cheryl M Koh, Kathryn M Deck, Joseph S Pitula, Charles J Epstein, Mark D Fleming, and Richard S Eisenstein. Iron-responsive degradation of iron-regulatory protein 1 does not require the Fe-S cluster. *EMBO J*, 25(3):544–553, February 2006. ISSN 0261-4189.

[253] Toru Nakatsu, Susumu Ichiyama, Jun Hiratake, Adrian Saldanha, Nobuyuki Kobashi, Kanzo Sakata, and Hiroaki Kato. Structural basis for the spectral difference in luciferase bioluminescence. *Nature*, 440(7082):372–376, March 2006. ISSN 1476-4687. doi: 10.1038/nature04542.

[254] E Conti, N P Franks, and P Brick. Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure*, 4(3):287–298, March 1996. ISSN 0969-2126.

[255] S Kim, J Mellor, A J Kingsman, and S M Kingsman. Multiple control elements in the TRP1 promoter of Saccharomyces cerevisiae. *Mol Cell Biol*, 6:4251–4258, December 1986. ISSN 0270-7306.

[256] G Braus, G Paravicini, and R Hütter. A consensus transcription termination sequence in the promoter region is necessary for efficient gene expression of the TRP1 gene of Saccharomyces cerevisiae. *Mol Gen Genet*, 212:495–504, June 1988. ISSN 0026-8925.

[257] Marie-Ange Teste, Manon Duquenne, Jean M François, and Jean-Luc Parrou. Validation of reference genes for quantitative expression analysis by real-time RT-PCR in Saccharomyces cerevisiae. *BMC Mol Biol*, 10:99, October 2009. ISSN 1471-2199. doi: 10.1186/1471-2199-10-99.

[258] C A Ball, H Jin, G Sherlock, S Weng, J C Matese, R Andrada, G Binkley, K Dolinski, S S Dwight, M A Harris, L Issel-Tarver, M Schroeder, D Botstein, and J M Cherry. Saccharomyces Genome Database provides tools to survey gene expression and functional analysis data. *Nucleic Acids Res*, 29(1):80–81, January 2001. ISSN 1362-4962.

[259] Stephen A Bustin, Vladimir Benes, Jeremy A Garson, Jan Hellemans, Jim Huggett, Mikael Kubista, Reinhold Mueller, Tania Nolan, Michael W Pfaffl, Gregory L Shipley, Jo Vandesompele, and Carl T Wittwer. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem*, 55(4):611–622, April 2009. ISSN 1530-8561. doi: 10.1373/clinchem.2008.112797.

[260] Sean Taylor, Michael Wakem, Greg Dijkman, Marwan Alsarraj, and Marie Nguyen. A practical approach to RT-qPCR-Publishing data that conform to the MIQE guidelines. *Methods*, 50(4):S1—-S5, April 2010. ISSN 1095-9130. doi: 10.1016/j.ymeth.2010.01.005.

[261] J Hasty, J Pradines, M Dolnik, and J J Collins. Noise-based switches and amplifiers for gene expression. *PNAS*, 97:2075–2080, 2000.

[262] F JACOB and J MONOD. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol*, 3:318–356, June 1961. ISSN 0022-2836.

[263] Y N Kaznessis. Models for synthetic biology. *BMC Systems Biology*, 1:47, 2007.

[264] J E Ferrell. Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs. *Trends Biochem Sci*, 21(12):460–466, December 1996. ISSN 0968-0004.

[265] Sina Ghaemmaghami, Won-Ki Huh, Kiowa Bower, Russell W Howson, Archana Belle, Noah Dephoure, Erin K O'Shea, and Jonathan S Weissman. Global analysis of protein expression in yeast. *Nature*, 425(6959):737–741, October 2003.

[266] Brian D Slaughter, Joel W Schwartz, and Rong Li. Mapping dynamic protein interactions in MAP kinase signaling using live-cell fluorescence fluctuation spectroscopy and imaging. *Proc Natl Acad Sci U S A*, 104(51):20320–20325, December 2007. ISSN 1091-6490. doi: 10.1073/pnas.0710336105.

[267] Celine I Maeder, Mark A Hink, Ali Kinkhabwala, Reinhard Mayr, Philippe I H Bastiaens, and Michael Knop. Spatial regulation of Fus3 MAP kinase activity through a reaction-diffusion mechanism in yeast pheromone signalling. *Nat Cell Biol*, 9(11):1319–1326, November 2007. ISSN 1465-7392. doi: 10.1038/ncb1652.

[268] S Park. Rewiring MAP Kinase Pathways Using Alternative Scaffold Assembly Mechanisms. *Science*, 299:1061–1064, 2003.

[269] Nicholas T Ingolia and Andrew W Murray. Positive-feedback loops as a flexible biological module. *Curr Biol*, 17(8):668–677, April 2007. ISSN 0960-9822. doi: 10.1016/j.cub.2007.03.016.

[270] J E Ferrell and R R Bhatt. Mechanistic studies of the dual phosphorylation of mitogen-activated protein kinase. *J Biol Chem*, 272:19008–19016, July 1997.

[271] L Michaelis and L Menten M. Die Kinetik der Invertinwirkung. *Biochem. Z.*, 49: 333–369, 1913.

[272] A G Hinnebusch. Mechanisms of gene regulation in the general control of amino acid biosynthesis in Saccharomyces cerevisiae. *Microbiol Rev*, 52(2):248–273, June 1988. ISSN 0146-0749.

[273] A Cornish-Bowden. *Basic Mathematics for Biochemists*. Oxford University Press, 2nd edition, 1999.

[274] Benjamin Lewin. *Genes IX*. Jones and Bartlett Publishers Inc, 9 edition, 2007.

[275] X Zhang, P Dennis, M Ehrenberg, and H Bremer. Kinetic properties of rrn promoters in Escherichia coli. *Biochimie*, 84(10):981–996, October 2002. ISSN 0300-9084.

[276] Patrick P Dennis, Mans Ehrenberg, and Hans Bremer. Control of rRNA synthesis in Escherichia coli: a systems biology approach. *Microbiol Mol Biol Rev*, 68 (4):639–668, December 2004. ISSN 1092-2172. doi: 10.1128/MMBR.68.4. 639-668.2004.

[277] Nicolas E Buchler, Ulrich Gerland, and Terence Hwa. Nonlinear protein degradation and the function of genetic circuits. *Proc Natl Acad Sci U S A*, 102 (27):9559–9564, July 2005. ISSN 0027-8424. doi: 10.1073/pnas.0409553102.

[278] Daniel Zenklusen, Daniel R Larson, and Robert H Singer. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat Struct Mol Biol*, 15 (12):1263–1271, December 2008. ISSN 1545-9985. doi: 10.1038/nsmb.1514.

[279] B Bonven and K Gullv. Peptide chain elongation rate and ribosomal activity in Saccharomyces cerevisiae as a function of the growth rate. *Mol Gen Genet*, 170 (2):225–230, February 1979. ISSN 0026-8925.

[280] Yulei Wang, Chih Long Liu, John D Storey, Robert J Tibshirani, Daniel Herschlag, and Patrick O Brown. Precision and functional specificity in mRNA decay. *Proc Natl Acad Sci U S A*, 99(9):5860–5865, April 2002. ISSN 0027-8424. doi: 10.1073/pnas.092538799.

[281] Archana Belle, Amos Tanay, Ledion Bitincka, Ron Shamir, and Erin K O'Shea. Quantification of protein half-lives in the budding yeast proteome. *Proc Natl Acad Sci U S A*, 103(35):13004–13009, August 2006. ISSN 0027-8424. doi: 10.1073/pnas.0605420103.

[282] J Hakenberg, Sebastian Schmeier, Axel Kowald, Edda Klipp, and Ulf Leser. Finding kinetic parameters using text mining. *OMICS*, 8(2):131–152, 2004. ISSN 1536-2310. doi: 10.1089/1536231041388366.

[283] Peter Li, Joseph O Dada, Daniel Jameson, Irena Spasic, Neil Swainston, Kathleen Carroll, Warwick Dunn, Farid Khan, Naglis Malys, Hanan L Messiha, Evangelos Simeonidis, Dieter Weichart, Catherine Winder, Jill Wishart, David S Broomhead, Carole A Goble, Simon J Gaskell, Douglas B Kell, Hans V Westerhoff, Pedro Mendes, and Norman W Paton. Systematic integration of experimental data and models in systems biology. *BMC Bioinformatics*, 11:582, November 2010. ISSN 1471-2105. doi: 10.1186/1471-2105-11-582.

[284] M Sassanfar and J W Roberts. Nature of the SOS-inducing signal in Escherichia coli. The involvement of DNA replication. *J Mol Biol*, 212(1):79–96, March 1990. ISSN 0022-2836. doi: 10.1016/0022-2836(90)90306-7.

[285] G M Leclerc, F R Boockfor, W J Faught, and L S Frawley. Development of a destabilized firefly luciferase enzyme for measurement of gene expression. *Biotechniques*, 29(3):590—-1, 594–6, 598 passim, September 2000. ISSN 0736-6205.

[286] R Keith Esch, Yuqi Wang, and Beverly Errede. Pheromone-induced degradation of Ste12 contributes to signal attenuation and the specificity of developmental fate. *Eukaryot Cell*, 5(12):2147–2160, December 2006. ISSN 1535-9778. doi: 10.1128/EC.00270-06.

[287] M A Gibson and J Bruck. Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels. *J. Phys. Chem*, A104(9):1876–1889, 2000.

[288] D J Wilkinson. *Stochastic Modelling for Systems Biology*. Chapman and Hall/CRC, April 2006.

[289] Florian Hartig, Justin M Calabrese, Björn Reineking, Thorsten Wiegand, and Andreas Huth. Statistical inference for stochastic simulation models - theory and application. *Ecol Lett*, 14(8):816–827, August 2011. ISSN 1461-0248. doi: 10.1111/j.1461-0248.2011.01640.x.

[290] Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403–434, 1976. ISSN 0021-9991. doi: DOI:10.1016/0021-9991(76)90041-3.

[291] Guido Klingbeil, Radek Erban, Mike Giles, and Philip K Maini. STOCHSIMGPU: parallel stochastic simulation for the Systems Biology Toolbox 2 for MATLAB. *Bioinformatics*, 27(8):1170–1171, April 2011. ISSN 1367-4811. doi: 10.1093/bioinformatics/btr068.

[292] D T Gillespie. Exact Stochastic Simulation of Coupled Chemical Reactions. *The Journal of Physical Chemistry*, 81:25:2340–2361, 1977.

[293] W H Press, S A Teukolsky, T A Vetterling, and B P Flannery. *Numerical Recipes: The Art of Scientific Computing.* New York: Cambridge University Press., 3rd edition, 2007.

[294] Mads Kaern, Timothy C Elston, William J Blake, and James J Collins. Stochasticity in gene expression: from theories to phenotypes. *Nat Rev Genet*, 6(6):451–464, June 2005. ISSN 1471-0056. doi: 10.1038/nrg1615.

[295] Avigdor Eldar and Michael B Elowitz. Functional roles for noise in genetic circuits. *Nature*, 467(7312):167–173, September 2010. ISSN 1476-4687. doi: 10.1038/nature09326.

[296] Mariette R Atkinson, Michael A Savageau, Jesse T Myers, and Alexander J Ninfa. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli. *Cell*, 113(5):597–607, May 2003. ISSN 0092-8674.

[297] Chris D Cox, James M McCollum, Derek W Austin, Michael S Allen, Roy D Dar, and Michael L Simpson. Frequency domain analysis of noise in simple gene circuits. *Chaos*, 16(2):26102, June 2006. ISSN 1054-1500. doi: 10.1063/1.2204354.

[298] Andrew R Joyce and Bernhard ØPalsson. The model organism as a system: integrating 'omics' data sets. *Nat Rev Mol Cell Biol*, 7(3):198–210, March 2006. ISSN 1471-0072. doi: 10.1038/nrm1857.

[299] Paul D Dobson, Kieran Smallbone, Daniel Jameson, Evangelos Simeonidis, Karin Lanthaler, PÄśnar Pir, Chuan Lu, Neil Swainston, Warwick B Dunn, Paul Fisher, Duncan Hull, Marie Brown, Olusegun Oshota, Natalie J Stanford, Douglas B Kell, Ross D King, Stephen G Oliver, Robert D Stevens, and Pedro Mendes. Further developments towards a genome-scale metabolic model of yeast. *BMC Syst Biol*, 4:145, October 2010. ISSN 1752-0509. doi: 10.1186/1752-0509-4-145.

[300] Jean-Marc Schwartz, Claire Gaugain, Jose C Nacher, Antoine de Daruvar, and Minoru Kanehisa. Observing metabolic functions at the genome scale. *Genome Biol*, 8(6):R123, 2007. ISSN 1465-6914. doi: 10.1186/gb-2007-8-6-r123.

[301] H Rabitz, M Kramer, and D Dacol. Sensitivity Analysis in Chemical Kinetics. *Ann. Rev. Phys. Chern.*, 34:419–461, 1983.

[302] S G Oliver, M K Winson, D B Kell, and F Baganz. Systematic functional analysis of the yeast genome. *Trends Biotechnol*, 16(9):373–378, September 1998. ISSN 0167-7799.

[303] D A Fell. Metabolic control analysis: a survey of its theoretical and experimental development. *Biochem J*, 286 (Pt 2:313–330, September 1992. ISSN 0264-6021.

[304] Maciej Dobrzynski and Frank J Bruggeman. Elongation dynamics shape bursty transcription and translation. *Proc Natl Acad Sci U S A*, 106(8):2583–2588, February 2009. ISSN 1091-6490. doi: 10.1073/pnas.0803507106.

[305] Andrew Gordon, Alejandro Colman-Lerner, Tina E Chin, Kirsten R Benjamin, Richard C Yu, and Roger Brent. Single-cell quantification of molecules and rates using open-source microscope-based cytometry. *Nat Methods*, 4(2):175–181, February 2007. ISSN 1548-7091. doi: 10.1038/nmeth1008.

[306] Petra Schwille and Stefan Diez. Synthetic biology of minimal systems. *Crit Rev Biochem Mol Biol*, 44:223–242, 2009. ISSN 1549-7798. doi: 10.1080/10409230903074549.

[307] Ahmed Fadiel, Kenneth D Eichenbaum, Nermin El Semary, and Brittiny Epperson. Mycoplasma genomics: tailoring the genome for minimal life requirements through reductive evolution. *Front Biosci*, 12:2020–2028, January 2007. ISSN 1093-4715.

APPENDIX

A

APPENDIX

A.1 Sequences

A.1.1 Iron Response Element (IRE) Nucleotide Sequence

```
  1 CCAATTATCT ACTTAAGCTT CAACAGTGCT TGAACTTAAG AACACAAAAC
 51 TCGAGA
```

Wild type Iron Response Element, from Koloteva *et al* [4].

A.1.2 P_{FUS1}-IRE-Luciferase Nucleotide Sequence

```
  1 GGATCCGGCA CCAGAACCGC TACTGAACGA TGATTCAGTT CGCCTTCTAT
 51 CCTTTGTTTA CGTATTTGTT TATATATATA ACTTTATTTT TTTTTATTAA
101 TTGGGCTGCA AGACAATTTT GTTGTCAGTG ATGCCTCAAT CCTTCTTTTG
151 CTTCCATATT TACCATGTGG ACCCTTTCAA AACAGAGTTG TATCTCTGCA
201 GGATGCCCTT TTTGACGTAT TGAATGGCAT AATTGCACTG TCACTTTTCG
251 CGCTGTCTCA TTTTGGTGCG ATGATGAAAC AAACATGAAA CGTCTGTAAT
```

```
 301 TTGAAACAAA TAACGTAATT CTCGGGATTG GTTTTATTTA AATGACAATG
 351 TAGGAGTGGC TTTGTAAGGT ATGTGTTGCT CTTAAAATAT TTGGATACGA
 401 CATCCTTTAT CTTTTTTCCT TTAAGAGCAG GATATAAGCC ATCAAGTTTC
 451 TGAAAATCCA TACCAATTAT CTACTTAAGC TTCAACAGTG CTTGAACTTA
 501 AGAACACAAA ACTCGAGAAG ATATGGAAGA CGCCAAAAAC ATAAAGAAAG
 551 GCCCGGCGCC ATTCTATCCA CTAGAGGATG GAACCGCTGG AGAGCAACTG
 601 CATAAGGCTA TGAAGAGATA CGCCCTGGTT CCTGGAACAA TTGCTTTTAC
 651 AGATGCACAT ATCGAGGTGA ACATCACGTA CGCGGAATAC TTCGAAATGT
 701 CCGTTCGGTT GGCAGAAGCT ATGAAACGAT ATGGGCTGAA TACAAATCAC
 751 AGAATCGTCG TATGCAGTGA AAACTCTCTT CAATTCTTTA TGCCGGTGTT
 801 GGGCGCGTTA TTTATCGGAG TTGCAGTTGC GCCCGCGAAC GACATTTATA
 851 ATGAACGTGA ATTGCTCAAC AGTATGAACA TTTCGCAGCC TACCGTAGTG
 901 TTTGTTTCCA AAAAGGGGTT GCAAAAAATT TTGAACGTGC AAAAAAAATT
 951 ACCAATAATC CAGAAAATTA TTATCATGGA TTCTAAAACG GATTACCAGG
1001 GATTTCAGTC GATGTACACG TTCGTCACAT CTCATCTACC TCCCGGTTTT
1051 AATGAATACG ATTTTGTACC AGAGTCCTTT GATCGTGACA AAACAATTGC
1101 ACTGATAATG AATTCCTCTG GATCTACTGG GTTACCTAAG GGTGTGGCCC
1151 TTCCGCATAG AACTGCCTGC GTCAGATTCT CGCATGCCAG AGATCCTATT
1201 TTTGGCAATC AAATCATTCC GGATACTGCG ATTTTAAGTG TTGTTCCATT
1251 CCATCACGGT TTTGGAATGT TTACTACACT CGGATATTTG ATATGTGGAT
1301 TTCGAGTCGT CTTAATGTAT AGATTTGAAG AAGAGCTGTT TTTACGATCC
1351 CTTCAGGATT ACAAAATTCA AAGTGCGTTG CTAGTACCAA CCCTATTTTC
1401 ATTCTTCGCC AAAAGCACTC TGATTGACAA ATACGATTTA TCTAATTTAC
1451 ACGAAATTGC TTCTGGGGGC GCACCTCTTT CGAAAGAAGT CGGGGAAGCG
1501 GTTGCAAAAC GCTTCCATCT TCCAGGGATA CGACAAGGAT ATGGGCTCAC
1551 TGAGACTACA TCAGCTATTC TGATTACACC CGAGGGGGAT GATAAACCGG
1601 GCGCGGTCGG TAAAGTTGTT CCATTTTTTG AAGCGAAGGT TGTGGATCTG
1651 GATACCGGGA AAACGCTGGG CGTTAATCAG AGAGGCGAAT TATGTGTCAG
1701 AGGACCTATG ATTATGTCCG GTTATGTAAA CAATCCGGAA GCGACCAACG
1751 CCTTGATTGA CAAGGATGGA TGGCTACATT CTGGAGACAT AGCTTACTGG
1801 GACGAAGACG AACACTTCTT CATAGTTGAC CGCTTGAAGT CTTTAATTAA
```

1851 ATACAAAGGA TATCAGGTGG CCCCCGCTGA ATTGGAATCG ATATTGTTAC

1901 AACACCCCAA CATCTTCGAC GCGGGCGTGG CAGGTCTTCC CGACGATGAC

1951 GCCGGTGAAC TTCCCGCCGC CGTTGTTGTT TTGGAGCACG GAAAGACGAT

2001 GACGGAAAAA GAGATCGTGG ATTACGTGGC CAGTCAAGTA ACAACCGCGA

2051 AAAAGTTGCG CGGAGGAGTT GTGTTTGTGG ACGAAGTACC GAAAGGTCTT

2101 ACCGGAAAAC TCGACGCAAG AAAAATCAGA GAGATCCTCA TAAAGGCCAA

2151 GAAGGGCGGA AAGTCCAAAT TGTAA

FUS1 promoter (green), with iron response element (red), and the Luciferase gene (yellow)

A.1.3 P$_{FUS1}$-LexA Nucleotide Sequence

1 GGATCCGGCA CCAGAACCGC TACTGAACGA TGATTCAGTT CGCCTTCTAT

51 CCTTTGTTTA CGTATTTGTT TATATATATA ACTTTATTTT TTTTTATTAA

101 TTGGGCTGCA AGACAATTTT GTTGTCAGTG ATGCCTCAAT CCTTCTTTTG

151 CTTCCATATT TACCATGTGG ACCCTTTCAA AACAGAGTTG TATCTCTGCA

201 GGATGCCCTT TTTGACGTAT TGAATGGCAT AATTGCACTG TCACTTTTCG

251 CGCTGTCTCA TTTTGGTGCG ATGATGAAAC AAACATGAAA CGTCTGTAAT

301 TTGAAACAAA TAACGTAATT CTCGGGATTG GTTTTATTTA AATGACAATG

351 TAGGAGTGGC TTTGTAAGGT ATGTGTTGCT CTTAAAATAT TTGGATACGA

401 CATCCTTTAT CTTTTTTCCT TTAAGAGCAG GATATAAGCC ATCAAGTTTC

451 TGAAAATCCA TATGATGAAA GCGTTAACGG CCAGGCAACA AGAGGTGTTT

501 GATCTCATCC GTGATCACAT CAGCCAGACA GGTATGCCGC CGACGCGTGC

551 GGAAATCGCG CAGCGTTTGG GGTTCCGTTC CCCAAACGCG GCTGAAGAAC

601 ATCTGAAGGC GCTGGCACGC AAAGGCGTTA TTGAAATTGT TTCCGGCGCA

651 TCACGCGGGA TTCGTCTGTT GCAGGAAGAG GAAGAAGGGT TGCCGCTGGT

701 AGGTCGTGTG GCTGCCGGTG AACCACTTCT GGCGCAACAG CATATTGAAG

751 GTCATTATCA GGTCGATCCT TCCTTATTCA AGCCGAATGC TGATTTCCTG

801 CTGCGCGTCA GCGGGATGTC GATGAAAGAT ATCGGCATTA TGGATGGTGA

851 CTTGCTGGCA GTGCATAAAA CTCAGGATGT ACGTAACGGT CAGGTCGTTG

901 TCGCACGTAT TGATGACGAA GTTACCGTTA AGCGCCTGAA AAAACAGGGC

```
 951 AATAAAGTCG AACTGTTGCC AGAAAATAGC GAGTTTAAAC CAATTGTCGT
1001 TGACCTTCGT CAGCAGAGCT TCACCATTGA AGGGCTGGCG GTTGGGGTTA
1051 TTCGCAACGG CGACTGGCTG TAA
```

P_{FUS1} promoter (green) and the *E. coli* LexA gene (yellow)

A.1.4 Yeast Cln2 Protein Sequence (Uniprot P20438)

```
   1 MASAEPRPRM GLVINAKPDY YPIELSNAEL LSHFEMLQEY HQEISTNVIA
  51 QSCKFKPNPK LIDQQPEMNP VETRSNIITF LFELSVVTRV TNGIFFHSVR
 101 LYDRYCSKRI VLRDQAKLVV ATCLWLAAKT WGGCNHIINN VVIPTGGRFY
 151 GPNPRARIPR LSELVHYCGD GQVFDESMFL QMERHILDTL NWNIYEPMIN
 201 DYVLNVDENC LMQYELYENQ VTYDKQCSEK RQSQLSQDSD ATVDERPYQN
 251 EEEEEEDLKL KIKLINLKKF LIDVSAWQYD LLRYELFEVS HGIFSIINQF
 301 TNQDHGPFLM TPMTSESKNG EILSTLMNGI VSIPNSLMEV YKTVNGVLPF
 351 INQVKEYHLD LQRKLQIASN LNISRKLTIS TPSCSFENSN STSIPSPASS
 401 SQSHTPMRNM SSLSDNSVFS RNMEQSSPIT PSMYQFGQQQ SNSICGSTVS
 451 VNSLVNTNNK QRIYEQITGP NSNNAINDYI DLLNLNESNK ENQNPATAHY
 501 LNGGPPKTSF INHGMFPSPT GTINSGKSSS ASFFNFFWYG QYPSNIVPVL
 551 FMINECQFCQ TYQKETNIIR FMKRKAFIFC GKRETTYSVY RYNFNGIYLY
 601 FCKSMKQPTK FCVELKKGH
```

The carboxy-terminal PEST-rich region is highlighted in red.

A.1.5 Nucleotide Sequence of the PEST-rich Carboxy-terminal region of Cln2

```
   1 GAATTGTACA AAGCATCCAA CTTGAACATT TCGAGAAAGC TTACCATATC
  51 AACCCCATCA TGCTCTTTCG AAAATTCAAA TAGCACATCC ATTCCTTCGC
 101 CCGCTTCCTC ATCTCAAAGC CACACTCCAA TGAGAAACAT GAGCTCACTC
 151 TCTGATAACA GCGTTTTCAG CCGGAATATG GAACAATCAT CACCAATCAC
 201 TCCAAGTATG TACCAATTTG GTCAGCAGCA GTCAAACAGT ATATGTGGTA
```

```
 251 GCACCGTTAG TGTGAATAGT CTGGTGAATA CAAATAACAA ACAAAGGATC
 301 TACGAACAAA TCACGGGTCC TAACAGCAAT AACGCAACCA ATGATTATAT
 351 TGATTTGCTA AACCTAAATG AGTCTAACAA GGAAAACCAA AATCCCGCAA
 401 CGGCGCATTA CCTCAATGGG GGCCCACCCA AGACAAGCTT CATTAACCAT
 451 GGAATGTTCC CCTCGCCAAC TGGGACCATA AATAGCGGTA AATCTAGCAG
 501 TGCCTCATCT TTAATTTCTT TTGGTATGGG CAATACCCAA GTAATATAGG
```

A.1.6 Iron Response Protein (IRP) Nucleotide Sequence

```
   1 ATGAGCAACC CATTCGCACA CCTTGCTGAG CCATTGGATC CTGTACAACC
  51 AGGAAAGAAA TTCTTCAATT TGAATAAATT GGAGGATTCA AGATATGGGC
 101 GCTTACCATT TTCGATCAGA GTTCTTCTGG AAGCAGCCAT TCGGAATTGT
 151 GATGAGTTTT TGGTGAAGAA ACAGGATATT GAAAATATTC TACATTGGAA
 201 TGTCACTCAG CACAAGAACA TAGAAGTGCC ATTTAAGCCT GCTCGTGTCA
 251 TCCTGCAGGA CTTTACGGGT GTGCCCGCTG TGGTTGACTT TGCTGCAATG
 301 CGTGATGCTG TGAAAAAGTT AGGAGGAGAT CCAGAGAAAA TAAACCCTGT
 351 CTGCCCTGCT GATCTTGTAA TAGATCATTC CATCCAGGTT GATTTCAACA
 401 GAAGGGCAGA CAGTTTACAG AAGAATCAAG ACCTGGAATT TGAAAGAAAT
 451 AGAGAGCGAT TTGAATTTTT AAAGTGGGGT TCCCAGGCTT TTCACAACAT
 501 GCGGATTATT CCCCCTGGCT CAGGAATCAT CCACCAGGTG AATTTGGAAT
 551 ATTTGGCAAG AGTGGTATTT GATCAGGATG GATATTATTA CCCAGACAGC
 601 CTCGTGGGCA CAGACTCGCA CACTACCATG ATTGATGGCT TGGGCATTCT
 651 TGGTTGGGGT GTCGGTGGTA TTGAAGCAGA AGCTGTCATG CTGGGTCAGC
 701 CAATCAGTAT GGTGCTTCCT CAGGTGATTG GCTACAGGCT GATGGGGAAG
 751 CCCCACCCTC TGGTAACATC CACTGACATC GTGCTCACCA TTACCAAGCA
 801 CCTCCGCCAG GTTGGGGTAG TGGGCAAATT TGTCGAGTTC TTCGGGCCTG
 851 GAGTAGCCCA GTTGTCCATT GCTGACCGAG CTACGATTGC TAACATGTGT
 901 CCAGAGTACG GAGCAACTGC TGCCTTTTTC CCAGTTGATG AAGTTAGTAT
 951 CACGTACCTG GTGCAAACAG GTCGTGATGA AGAAAAATTA AAGTATATTA
1001 AAAAATATCT TCAGGCTGTA GGAATGTTTC GAGATTTCAA TGACCCTTCT
1051 CAAGACCCAG ACTTCACCCA GGTTGTGGAA TTAGATTTGA AAACAGTAGT
```

```
1101 GCCTTGCTGT AGTGGACCCA AAAGGCCTCA GGACAAAGTT GCTGTGTCCG
1151 ACATGAAAAA GGACTTTGAG AGCTGCCTTG GAGCCAAGCA AGGATTTAAA
1201 GGATTCCAAG TTGCTCCTGA ACATCATAAT GACCATAAGA CCTTTATCTA
1251 TGATAACACT GAATTCACCC TTGCTCATGG TTCTGTGGTC ATTGCTGCCA
1301 TTACTAGCTG CACAAACACC AGTAATCCGT CTGTGATGTT AGGGGCAGGA
1351 TTGTTAGCAA AGAAAGCTGT GGATGCTGGC CTGAACGTGA TGCCTTACAT
1401 CAAAACTAGC CTGTCTCCTG GGAGTGGCGT GGTCACCTAC TACCTACAAG
1451 AAAGCGGAGT CATGCCTTAT CTGTCTCAGC TTGGGTTTGA CGTGGTGGGC
1501 TATGGCTGCA TGACCTGCAT TGGCAACAGT GGGCCTTTAC CTGAACCTGT
1551 GGTAGAAGCC ATCACACAGG GAGACCTTGT AGCTGTTGGA GTACTATCTG
1601 GAAACAGGAA TTTTGAAGGT CGAGTTCACC CCAACACCCG GGCCAACTAT
1651 TTAGCCTCTC CCCCCTTAGT AATAGCATAT GCAATTGCTG GAACCATCAG
1701 AATCGACTTT GAGAAAGAGC CATTGGGAGT AAATGCAAAG GGACAGCAGG
1751 TATTTCTGAA AGATATCTGG CCGACTAGAG ACGAGATCCA GGCAGTGGAG
1801 CGTCAGTATG TCATCCCGGG GATGTTTAAG GAAGTCTATC AGAAAATAGA
1851 GACTGTGAAT GAAAGCTGGA ATGCCTTAGC AACCCCATCA GATAAGCTGT
1901 TTTTCTGGAA TTCCAAATCT ACGTATATCA AATCACCACC ATTCTTTGAA
1951 AACCTGACTT TGGATCTTCA GCCCCCTAAA TCTATAGTGG ATGCCTATGT
2001 GCTGCTAAAT TTGGGAGATT CGGTAACAAC TGACCACATC TCCCCAGCTG
2051 GAAATATTGC AAGAAACAGT CCTGCTGCTC GCTACTTAAC TAACAGAGGC
2101 CTAACTCCAC GAGAATTCAA CTCCTATGGC TCCCGCCGAG GTAATGACGC
2151 CGTCATGGCA CGGGGAACAT TTGCCAACAT TCGCTTGTTA AACAGATTTT
2201 TGAACAAGCA GGCACCACAG ACTATCCATC TGCCTTCTGG GGAAATCCTT
2251 GATGTGTTTG ATGCTGCTGA GCGGTACCAG CAGGCAGGCC TTCCCCTGAT
2301 CGTTCTGGCT GGCAAAGAGT ACGGTGCAGG CAGCTCCCGA GACTGGGCAG
2351 CTAAGGGCCC TTTCCTGCTG GGAATCAAAG CCGTCCTGGC CGAGAGCTAC
2401 GAGCGCATTC ACCGCAGTAA CCTGGTTGGG ATGGGTGTGA TCCCACTTGA
2451 ATATCTCCCT GGTGAGAATG CAGATGCCCT GGGGCTCACA GGGCAAGAAC
2501 GATACACTAT CATTATTCCA GAAAACCTCA AACCACAAAT GAAAGTCCAG
2551 GTCAAGCTGG ATACTGGCAA GACCTTCCAG GCTGTCATGA GGTTTGACAC
2601 TGATGTGGAG CTCACTTATT TCCTCAACGG GGGCATCCTC AACTACATGA
```

```
2651 TCCGCAAGAT GGCCAAGTAG GAGACGTGCA CTTGGTCGTG CGCCCAGGGA
2701 GGAAGCCGCA CCACCAGCCA GCGCAGGCCC TGGTGGAGAG GCCTCCCTGG
2751 CTGCCTCTGG GAGGGGTGCT GCCTTGTAGA TGGAGCAAGT GAGCACTGAG
2801 GGTCTGGTGC CAATCCTGTA GGCACAAAAC CAGAAGTTTC TACATTCTCT
2851 ATTTTTGTTA ATCATCTTCT CTTTTTCCAG AATTTGGAAG CTAGAATGGT
2901 GGGAATGTCA GTAGTGCCAG AAAGAGAGAA CCAAG
```

A.1.7 IRP$_{PEST}$ Nucleotide Sequence

```
   1 ATGAGCAACC CATTCGCACA CCTTGCTGAG CCATTGGATC CTGTACAACC
  51 AGGAAAGAAA TTCTTCAATT TGAATAAATT GGAGGATTCA AGATATGGGC
 101 GCTTACCATT TTCGATCAGA GTTCTTCTGG AAGCAGCCAT TCGGAATTGT
 151 GATGAGTTTT TGGTGAAGAA ACAGGATATT GAAAATATTC TACATTGGAA
 201 TGTCACTCAG CACAAGAACA TAGAAGTGCC ATTTAAGCCT GCTCGTGTCA
 251 TCCTGCAGGA CTTTACGGGT GTGCCCGCTG TGGTTGACTT TGCTGCAATG
 301 CGTGATGCTG TGAAAAAGTT AGGAGGAGAT CCAGAGAAAA TAAACCCTGT
 351 CTGCCCTGCT GATCTTGTAA TAGATCATTC CATCCAGGTT GATTTCAACA
 401 GAAGGGCAGA CAGTTTACAG AAGAATCAAG ACCTGGAATT TGAAAGAAAT
 451 AGAGAGCGAT TTGAATTTTT AAAGTGGGGT TCCCAGGCTT TTCACAACAT
 501 GCGGATTATT CCCCCTGGCT CAGGAATCAT CCACCAGGTG AATTTGGAAT
 551 ATTTGGCAAG AGTGGTATTT GATCAGGATG GATATTATTA CCCAGACAGC
 601 CTCGTGGGCA CAGACTCGCA CACTACCATG ATTGATGGCT TGGGCATTCT
 651 TGGTTGGGGT GTCGGTGGTA TTGAAGCAGA AGCTGTCATG CTGGGTCAGC
 701 CAATCAGTAT GGTGCTTCCT CAGGTGATTG GCTACAGGCT GATGGGGAAG
 751 CCCCACCCTC TGGTAACATC CACTGACATC GTGCTCACCA TTACCAAGCA
 801 CCTCCGCCAG GTTGGGGTAG TGGGCAAATT TGTCGAGTTC TTCGGGCCTG
 851 GAGTAGCCCA GTTGTCCATT GCTGACCGAG CTACGATTGC TAACATGTGT
 901 CCAGAGTACG GAGCAACTGC TGCCTTTTTC CCAGTTGATG AAGTTAGTAT
 951 CACGTACCTG GTGCAAACAG GTCGTGATGA AGAAAAATTA AAGTATATTA
1001 AAAAATATCT TCAGGCTGTA GGAATGTTTC GAGATTTCAA TGACCCTTCT
1051 CAAGACCCAG ACTTCACCCA GGTTGTGGAA TTAGATTTGA AAACAGTAGT
```

```
1101 GCCTTGCTGT AGTGGACCCA AAAGGCCTCA GGACAAAGTT GCTGTGTCCG
1151 ACATGAAAAA GGACTTTGAG AGCTGCCTTG GAGCCAAGCA AGGATTTAAA
1201 GGATTCCAAG TTGCTCCTGA ACATCATAAT GACCATAAGA CCTTTATCTA
1251 TGATAACACT GAATTCACCC TTGCTCATGG TTCTGTGGTC ATTGCTGCCA
1301 TTACTAGCTG CACAAACACC AGTAATCCGT CTGTGATGTT AGGGGCAGGA
1351 TTGTTAGCAA AGAAAGCTGT GGATGCTGGC CTGAACGTGA TGCCTTACAT
1401 CAAAACTAGC CTGTCTCCTG GGAGTGGCGT GGTCACCTAC TACCTACAAG
1451 AAAGCGGAGT CATGCCTTAT CTGTCTCAGC TTGGGTTTGA CGTGGTGGGC
1501 TATGGCTGCA TGACCTGCAT TGGCAACAGT GGGCCTTTAC CTGAACCTGT
1551 GGTAGAAGCC ATCACACAGG GAGACCTTGT AGCTGTTGGA GTACTATCTG
1601 GAAACAGGAA TTTTGAAGGT CGAGTTCACC CCAACACCCG GGCCAACTAT
1651 TTAGCCTCTC CCCCCTTAGT AATAGCATAT GCAATTGCTG GAACCATCAG
1701 AATCGACTTT GAGAAAGAGC CATTGGGAGT AAATGCAAAG GGACAGCAGG
1751 TATTTCTGAA AGATATCTGG CCGACTAGAG ACGAGATCCA GGCAGTGGAG
1801 CGTCAGTATG TCATCCCGGG GATGTTTAAG GAAGTCTATC AGAAAATAGA
1851 GACTGTGAAT GAAAGCTGGA ATGCCTTAGC AACCCCATCA GATAAGCTGT
1901 TTTTCTGGAA TTCCAAATCT ACGTATATCA AATCACCACC ATTCTTTGAA
1951 AACCTGACTT TGGATCTTCA GCCCCCTAAA TCTATAGTGG ATGCCTATGT
2001 GCTGCTAAAT TTGGGAGATT CGGTAACAAC TGACCACATC TCCCCAGCTG
2051 GAAATATTGC AAGAAACAGT CCTGCTGCTC GCTACTTAAC TAACAGAGGC
2101 CTAACTCCAC GAGAATTCAA CTCCTATGGC TCCCGCCGAG GTAATGACGC
2151 CGTCATGGCA CGGGGAACAT TTGCCAACAT TCGCTTGTTA AACAGATTTT
2201 TGAACAAGCA GGCACCACAG ACTATCCATC TGCCTTCTGG GGAAATCCTT
2251 GATGTGTTTG ATGCTGCTGA GCGGTACCAG CAGGCAGGCC TTCCCCTGAT
2301 CGTTCTGGCT GGCAAAGAGT ACGGTGCAGG CAGCTCCCGA GACTGGGCAG
2351 CTAAGGGCCC TTTCCTGCTG GGAATCAAAG CCGTCCTGGC CGAGAGCTAC
2401 GAGCGCATTC ACCGCAGTAA CCTGGTTGGG ATGGGTGTGA TCCCACTTGA
2451 ATATCTCCCT GGTGAGAATG CAGATGCCCT GGGGCTCACA GGGCAAGAAC
2501 GATACACTAT CATTATTCCA GAAAACCTCA AACCACAAAT GAAAGTCCAG
2551 GTCAAGCTGG ATACTGGCAA GACCTTCCAG GCTGTCATGA GGTTTGACAC
2601 TGATGTGGAG CTCACTTATT TCCTCAACGG GGGCATCCTC AACTACATGA
```

```
2651 TCCGCAAGAT GGCCAAAGCA TCCAACTTGA ACATTTCGAG AAAGCTTACC
2701 ATATCAACCC CATCATGCTC TTTCGAAAAT TCAAATAGCA CATCCATTCC
2751 TTCGCCCGCT TCCTCATCTC AAAGCCACAC TCCAATGAGA AACATGAGCT
2801 CACTCTCTGA TAACAGCGTT TTCAGCCGGA ATATGGAACA ATCATCACCA
2851 ATCACTCCAA GTATGTACCA ATTTGGTCAG CAGCAGTCAA ACAGTATATG
2901 TGGTAGCACC GTTAGTGTGA ATAGTCTGGT GAATACAAAT AACAAACAAA
2951 GGATCTACGA ACAAATCACG GGTCCTAACA GCAATAACGC AACCAATGAT
3001 TATATTGATT TGCTAAACCT AAATGAGTCT AACAAGGAAA ACCAAAATCC
3051 CGCAACGGCG CATTACCTCA ATGGGGGCCC ACCCAAGACA AGCTTCATTA
3101 ACCATGGAAT GTTCCCCTCG CCAACTGGGA CCATAAATAG CGGTAAATCT
3151 AGCAGTGCCT CATCTTTAAT TTCTTTTGGT ATGGGCAATA CCCAAGTAAT
3201 ATAGTGGCCA AGTAGGAGAC GTGCACTTGG TCGTGCGCCC AGGGAGGAAG
3251 CCGCACCACC AGCCAGCGCA GGCCCTGGTG GAGAGGCCTC CCTGGCTGCC
3301 TCTGGGAGGG GTGCTGCCTT GTAGATGGAG CAAGTGAGCA CTGAGGGTCT
3351 GGTGCCAATC CTGTAGGCAC AAAACCAGAA GTTTCTACAT TCTCTATTTT
3401 TGTTAATCAT CTTCTCTTTT TCCAGAATTT GGAAGCTAGA ATGGTGGGAA
3451 TGTCAGTAGT GCCAGAAAGA GAGAACCAAG
```

The fused PEST-rich C-terminal region of the Mateus and Avery Cln2 protein is highlighted in red. [216].

A.1.8 to demonstrate reading frame

```
   1 MSNPFAHLAE PLDPVQPGKK FFNLNKLEDS RYGRLPFSIR VLLEAAIRNC
  51 DEFLVKKQDI ENILHWNVTQ HKNIEVPFKP ARVILQDFTG VPAVVDFAAM
 101 RDAVKKLGGD PEKINPVCPA DLVIDHSIQV DFNRRADSLQ KNQDLEFERN
 151 RERFEFLKWG SQAFHNMRII PPGSGIIHQV NLEYLARVVF DQDGYYYPDS
 201 LVGTDSHTTM IDGLGILGWG VGGIEAEAVM LGQPISMVLP QVIGYRLMGK
 251 PHPLVTSTDI VLTITKHLRQ VGVVGKFVEF FGPGVAQLSI ADRATIANMC
 301 PEYGATAAFF PVDEVSITYL VQTGRDEEKL KYIKKYLQAV GMFRDFNDPS
 351 QDPDFTQVVE LDLKTVVPCC SGPKRPQDKV AVSDMKKDFE SCLGAKQGFK
 401 GFQVAPEHHN DHKTFIYDNT EFTLAHGSVV IAAITSCTNT SNPSVMLGAG
```

```
 451 LLAKKAVDAG LNVMPYIKTS LSPGSGVVTY YLQESGVMPY LSQLGFDVVG
 501 YGCMTCIGNS GPLPEPVVEA ITQGDLVAVG VLSGNRNFEG RVHPNTRANY
 551 LASPPLVIAY AIAGTIRIDF EKEPLGVNAK GQQVFLKDIW PTRDEIQAVE
 601 RQYVIPGMFK EVYQKIETVN ESWNALATPS DKLFFWNSKS TYIKSPPFFE
 651 NLTLDLQPPK SIVDAYVLLN LGDSVTTDHI SPAGNIARNS PAARYLTNRG
 701 LTPREFNSYG SRRGNDAVMA RGTFANIRLL NRFLNKQAPQ TIHLPSGEIL
 751 DVFDAAERYQ QAGLPLIVLA GKEYGAGSSR DWAAKGPFLL GIKAVLAESY
 801 ERIHRSNLVG MGVIPLEYLP GENADALGLT GQERYTIIIP ENLKPQMKVQ
 851 VKLDTGKTFQ AVMRFDTDVE LTYFLNGGIL NYMIRKMAKA SNLNISRKLT
 901 ISTPSCSFEN SNSTSIPSPA SSSQSHTPMR NMSSLSDNSV FSRNMEQSSP
 951 ITPSMYQFGQ QQSNSICGST VSVNSLVNTN NKQRIYEQIT GPNSNNATND
1001 YIDLLNLNES NKENQNPATA HYLNGGPPKT SFINHGMFPS PTGTINSGKS
1051 SSASSLISFG MGNTQVIWPS RRRALGRAPR EEAAPPASAG PGGEASLAAS
1101 GRGAALMEQV STEGLVPILA QNQKFLHSLF LLIIFSFSRI WKLEWWECQC
1151 QKERTK
```

The protein sequence of the IRP (white) with Mateus and Avery PEST-rich C-terminal region, demonstrating the in-frame insertion of the tag [216].

A.1.9 LexA Operator, *DCD1* promoter, and IRP Nucleotide Sequence

```
   1 CGAGTACTGT ATGTACATAC AGTACTCGAG TACTGTATGT ACATACAGTA
  51 CAAGCTTCTT CCTGCCTAAA CAGGAAGACA AAGCATGCGA GAGGCCCTGG
 101 GTTCAATTCC CAGCTCGCCC CATTATAATT TTTTCACTTT TTTGTTCTTT
 151 GCAAGAACGC GCGGTACGCA GTTATGAGAT GATGTAGGCA ATCTCGAGAA
 201 TTGAAACTTC TGCATTACCA TAGAATTCAA CATCTTTTTT TGGCACATTA
 251 AAGGTGTGAA TGGATCCAAT TATCTACTTA AGAACACAAA ACTCGAGAAC
 301 ATATGAGCAA CCCATTCGCA CACCTTGCTG AGCCATTGGA TCCTGTACAA
 351 CCAGGAAAGA AATTCTTCAA TTTGAATAAA TTGGAGGATT CAAGATATGG
 401 GC
```

The nucleotide sequence of the LexA operator region (yellow), the *DCD1* promoter (green) and the first 100 nucleotides of the IRP gene (red). The intervening white section between the *DCD1* promoter and the IRP is a multi-cloning site.

A.1.10 LexA Operator, *TEF1* promoter, and IRP Nucleotide Sequence

```
  1 CGAGTACTGT ATGTACATAC AGTACTCGAG TACTGTATGT ACATACAGTA
 51 CAAGCTTTGA TTACGCCTCC CAGTCACGAC GTTGTAAAAC GACGGCCAGT
101 GCTACAATCG CGGCCGCATA GGTATCGATC GTCACCCTGC AGGGTGACGG
151 GATCGATCCG TCACCCGCAT ATTACATATA ATACATATCA CATAGGAAGC
201 AACAGGCGCG TTGGACTTTT AATTTTCGAG GACCGCGAAT CCTTACATCA
251 CACCCAATCC CCCACAAGTG ATCCCCCACA CACCATAGCT TCAAAATGTT
301 TCTACTCCTT TTTTACTCTT CCAGATTTTC TCGGACTCCG CGCATCGCCG
351 TACCACTTCA AAACACCCAA GCACAGCATA CTAAATTTCC CCTCTTTCTT
401 CCTCTAGGGT GTCGTTAATT ACCCGTACTA AAGGTTTGGA AAAGAAAAAA
451 GAGACCGCCT CGTTTCTTTT TCTTCGTCGA AAAAGGCAAT AAAAATTTTT
501 ATCACGTTTC TTTTTCTTGA AAATTTTTTT TTTTAGTTTT TTTCTCTTTC
551 GATGACCTCC CATTGATATT TAAGTTAATA AACGGTCTTC AATTTCTCAA
601 GTTTCAGTTT CATTTTTCTT GTTCTATTAC AACTTTTTTT ACTTCTTGTT
651 CATTAGAAAG GATCCGTCGA CTCGAGAACA TATGAGCAAC CCATTCGCAC
701 ACCTTGCTGA GCCATTGGAT CCTGTACAAC CAGGAAAGAA ATTCTTCAAT
751 TTGAATAAAT TGGAGGATTC AAGATATGGG CG
```

The nucleotide sequence of the LexA operator region (yellow), the *TEF1* promoter (green) and the first 100 nucleotides of the IRP gene (red).

APPENDIX

B

APPENDIX

B.1 Python script for processing Copasi stochastic data

```
#!/usr/bin/python
#To execute the script call:
#calculate_statisitcs.py INPUTFILE MEAN_FILE SD_FILE NUM_STEPS NUM_REPEATS

import sys
import string
import math

if(len(sys.argv)!=6):
    print "Wrong number of arguments."
    sys.exit(1)

INPUTFILE=sys.argv[1]

MEAN_OUTFILE=sys.argv[2]

SD_OUTFILE=sys.argv[3]

NUM_STEPS=int(sys.argv[4])

NUM_REPEATS=int(sys.argv[5])

INPUT=file(INPUTFILE,"r").readlines()
```

```
if(len(INPUT) != (NUM_REPEATS * (NUM_STEPS + 2) + 1)):
    print "Wrong number of data points."
    sys.exit(1)

DATA=[]
LINENUMBER=1
NUMCOLUMNS=len(string.split(INPUT[1]))
MEAN=[]

HEADER=INPUT[0]

for X in range(0,NUM_STEPS+1):
    MEAN.append([])
    for Y in range(0,NUMCOLUMNS - 1):
        MEAN[X].append(0)

for X in range(0,NUM_REPEATS):
   SET=[]
   for Y in range(0, NUM_STEPS+1):
     LINE=INPUT[LINENUMBER]
     COLUMNS=string.split(LINE)
     if(len(COLUMNS)!=NUMCOLUMNS):
        print "Wrong number of elements on  line %d"%(LINENUMBER)
        sys.exit(1)
     ROW=[]
     for Z in range(0,len(COLUMNS)-1):
        v=float(COLUMNS[Z+1])
        ROW.append(v)
        MEAN[Y][Z]=MEAN[Y][Z]+v
     SET.append(ROW)
     LINENUMBER=LINENUMBER+1
   DATA.append(SET)
   LINENUMBER=LINENUMBER+1  # skip the empty line

MEANOUT=file(MEAN_OUTFILE,"w")
MEANOUT.write(HEADER)
for X in range(0,NUM_STEPS+1):
    line=string.join([str(X)],",")
    for Y in range(0,NUMCOLUMNS-1):
        MEAN[X][Y]=MEAN[X][Y]/NUM_REPEATS
        line=string.join([line,str(MEAN[X][Y])],",")
    line=line+"\n"
    MEANOUT.write(line)
MEANOUT.close()

SD=[]
for X in range(0,NUM_STEPS+1):
```

```
    SD.append([])
    for Y in range(0,NUMCOLUMNS-1):
        SD[X].append(0)


for X in range(0,NUM_REPEATS):
    for Y in range(0,NUM_STEPS+1):
        for Z in range(0,NUMCOLUMNS-1):
            SD[Y][Z]=SD[Y][Z]+math.pow((DATA[X][Y][Z]-MEAN[Y][Z]),2)


SDOUT=file(SD_OUTFILE,"w")
SDOUT.write(HEADER)
for X in range(0,NUM_STEPS+1):
    line=string.join([str(X)],",")
    for Y in range(0,NUMCOLUMNS-1):
        SD[X][Y]=math.sqrt(SD[X][Y]/NUM_REPEATS)
        line=string.join([line,str(SD[X][Y])],",")
    line=line+"\n"
    SDOUT.write(line)
SDOUT.close()
```