

PREDICTION OF COMBINATION EFFICACY
IN CANCER THERAPY

A THESIS

SUBMITTED TO THE UNIVERSITY OF MANCHESTER

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY (PHD)

IN THE FACULTY OF ENGINEERING AND PHYSICAL SCIENCES

JIE YANG

SCHOOL OF CHEMICAL ENGINEERING AND ANALYTICAL
SCIENCE 2012

CONTENTS

Abstract	8
Declaration	9
Copyright	10
Acknowledgements	11
1 Introduction	12
1.1 A Systems Biology Perspective Of Cancer.....	12
1.2 Cancer And The Eukaryotic Cell.....	17
1.3 The Cell Cycle.....	20
1.4 Regulation Of The Cell Cycle.....	24
1.5 Cell Cycle Metabolism And Growth.....	27
1.6 Normal vs. Cancer Cell Metabolism.....	34
2 Models Of The Cell Cycle	38
2.1 Modelling The Cell Cycle.....	38
2.2 From Yeast To Mammalian Cell Cycle Models.....	49
2.3 A Mammalian Cell Cycle Model.....	58
3 A Model Of Growth And Division	66
3.1 A Mammalian Cell Cycle And Metabolism Model.....	67
3.2 Model Implementation.....	72

3.3	Results And Discussions.....	78
4	A Model Of Cancer Growth And Division	85
4.1	A Mammalian Cell Cycle And Metabolism Model For Cancer Cells.....	85
4.2	Model Implementation.....	88
4.3	Results And Discussions.....	95
5	Drug Combination Efficacy In Cancer Therapy	101
5.1	Introduction.....	101
5.2	Methods.....	103
5.3	Results And Discussions.....	111
6	Summary And Future Directions	117
	Bibliography	123
	Appendix A	140
	Appendix B	148
	Appendix C	159

Final word count: 27,667

LIST OF FIGURES

1.1	The phases of the mammalian cell cycle.....	20
1.2	The two self-maintaining stable steady states of the cell cycle.....	23
1.3	The biosynthetic pathway of glutamate, glutamine and proline.....	32
1.4	The biosynthetic pathway of serine, glycine and cysteine.....	33
1.5	Diagram illustrating the differences between the metabolic pathways for normal and cancer cells.....	35
2.1	Schematic workflow of a cell cycle modelling process.....	45
3.1	Schematic diagram for the cell mass from the Conradie <i>et al.</i> mammalian cell cycle model.....	69
3.2	Schematic diagram for the extended model reactions involved in the cell cycle and metabolism model.....	70
3.3	Time course plots for cell mass when the initial glucose concentration is set to 1mM (yellow line) and 10mM (blue line).....	79
3.4	Time course plots for non-essential amino acids when the initial glucose concentration is set to 1mM (green line) and 10mM (red line).....	80
3.5	Time course plot for pyruvate when the initial glucose concentration is set to 10mM.....	81
3.6	Time course plot for lactate when the initial glucose concentration is set to 10mM.....	81
3.7	Time course plot for PEP when the initial glucose concentration is set	

	to 10mM.....	82
3.8	Time course plot for P2G when the initial glucose concentration is set to 10mM.....	82
3.9	Maximal metabolic rate constant for reaction 33 vs. Glucose Concentration.....	83
3.10	Cell Cycle Period (after 100 hours) as a function of Glucose concentration for normal cells.....	84
4.1	Methodology for creating the cancer cell cycle and metabolism model based on gene expression data.....	93
4.2	Schematic diagram showing the mean values of the cancer to normal gene expression level ratios involved in the model reactions found in the glycolysis and non-essential amino acid pathways.....	93
4.3	Comparison of cell mass time courses for normal and cancer cells.....	98
4.4	Comparison of cycle period vs. glucose concentration for normal and cancer cells.....	99
5.1	Comparison of the best value obtained for the cycle period when all the maximum metabolic rate constants of the cancer model are subjected to different optimisation algorithms in COPASI.....	107

LIST OF TABLES

1.1	Effect of cycloheximide on the cell cycle during different treatment Times.....	26
2.1	Examples of the notation utilised in the molecular interaction maps by Kohn.....	52
3.1	Abbreviations used in the cell cycle and metabolism model schematic diagram for model reactions as shown in Figure 3.2.....	70
4.1	Table showing the genes involved in the glycolysis pathway and their corresponding gene expression intensity values and expression ratio values given by Cancer (C) divided by Normal (N).....	89
4.2	A comparison of scaled sensitivity values generated from performing sensitivity analysis on the metabolic rate constants with regards to the cell cycle period for normal and cancer cells.....	96
5.1	Examples of algorithms implemented in COPASI to perform Optimisation.....	102
5.2	Details about the type of optimisation performed on different combinations of metabolic rate constants and their corresponding best cycle period values obtained through Hooke and Jeeves, and particle swarm.....	112
A.1	The reaction steps and their corresponding rate equations in the Conradie <i>et al.</i> mammalian cell cycle model.....	140
B.1	The kinetics involved in the mammalian cell cycle and metabolism	

model.....148

C.1 Gene expression data for the non-essential amino acids.....159

ABSTRACT

The cell cycle is an essential process in all living organisms that must be carefully regulated to ensure successful cell growth and division. Disregulation of the cell cycle is a key contributing factor towards the formation of cancerous cells. Understanding events at a cellular level is the first step towards comprehending how cancer manifests at an organismal level. Mathematical modelling can be used as a means of formalising and predicting the behaviour of the biological systems involved in cancer. In response, cell cycle models have been constructed to simulate and predict what happens to the mammalian cell over a time course in response to variable parameters.

Current cell cycle models rarely account for certain precursors of cell growth such as energy usage and the need for non-essential amino acids as fundamental building blocks of macromolecules. Normal and cancer cell metabolism differ in the way they derive energy from glucose. In addition, normal and cancer cells also demonstrate different levels of gene expression.

Two versions of a mammalian cell cycle and metabolism model, based on ordinary differential equations (ODEs) that respond to fluctuations in glucose concentration levels, have been developed here for the normal and cancer cell scenarios. Sensitivity analysis is performed for both normal and cancer cells using these cell cycle and metabolism models to investigate which kinetic reaction steps have a greater effect over the cell cycle period. Detailed analysis of the models and quantitatively assessing metabolite levels at various stages of the cell cycle may offer novel insights into how the glycolytic rate varies during the cell cycle for both normal and cancer cells.

The results of the sensitivity analysis are used to identify potential drug targets in cancer therapy. Combinations of these individual targets are also investigated to compare the different effects of single and multiple drug compounds on the time it takes to complete a cell division cycle.

DECLARATION

The University of Manchester
Traditional format PhD Candidate Declaration

Candidate Name: Jie Yang

Faculty: Engineering and Physical Sciences

Thesis Title: Prediction Of Combination Efficacy In Cancer Therapy

Declaration to be completed by the candidate:

I declare that no portion of this work referred to in this 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.

Signed:

Date: September 28th, 2012

COPYRIGHT

The author of this thesis (including any appendices and/or schedules to this thesis) owns any copyright in it (the “Copyright”)¹ and s/he has given The University of Manchester the right to use such Copyright for any administrative, promotional, educational and/or teaching purposes.

Copies of this thesis, either in full or in extracts, may be made only in accordance with the regulations of the John Rylands University Library of Manchester. Details of these regulations may be obtained from the Librarian. This page must form part of any such copies made.

The ownership of any patents, designs, trade marks and any and all other intellectual property rights except for the Copyright (the “Intellectual Property Rights”) and any reproductions of copyright works, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property Rights and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property Rights and/or Reproductions.

Further information on the conditions under which disclosure, publication and exploitation of this thesis, the Copyright and any Intellectual Property Rights and/or Reproductions described in it may take place is available from the Head of School of Chemical Engineering and Analytical Science (or the Vice-President) and the Dean of the Faculty of Engineering and Physical Sciences, for Faculty of Engineering and Physical Sciences candidates.

¹This excludes material already printed in academic journals, for which the copyright belong to said journal and publisher. Pages for which the author does not own the copyright are numbered differently from the rest of the thesis.

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Professor Pedro Mendes, Professor Hans Westerhoff, Professor Jacky Snoep from the University of Manchester and Fernando Ortega from AstraZeneca for their advice, guidance and encouragement.

I would also like to thank the EPSRC, BBSRC and the Doctoral Training Centre for the financial support they have provided during this PhD research project.

Finally, I am eternally grateful for the love, encouragement and support my parents, Xiaogang Yang and Xiaobing Huang (and the very adorable family cat, Alec), have given me through the PhD years. I love you all dearly and wish for many more happy memories to be made together.

INTRODUCTION

1.1 A Systems Biology Perspective of Cancer

“One day, we imagine that cancer biology and treatment at present, a patchwork quilt of cell biology, genetics, histopathology, biochemistry, immunology, and pharmacology will become a science with a conceptual structure and logical coherence that rivals that of chemistry of physics”.

(Hanahan and Weinberg, 2000)

Cancer cells arise when the conventional restraints on cell division are broken and dividing cells are no longer properly controlled by the regulatory mechanisms of the cell cycle. The effects of this disease are particularly detrimental due to the fact that cancerous cells and their progeny can invade and take control of territories that belong to cells with specific and vital responsibilities (Alberts *et al.*, 2002). In fact, cancer is one of the leading causes of death worldwide, with fatalities occurring predominately through cancerous growth of the lung, stomach, liver, colon and breast. In 2007, the

World Health Organisation established that cancer was responsible for 7.9 million deaths, which accounts for approximately 13% of all mortalities globally. This figure is expected to increase to 12 million deaths per year by 2030 (<http://www.who.int/cancer/en>).

Conventionally, the treatment of cancer has been through the exploitation of radiotherapy, chemotherapy and surgery (DeVita *et al.*, 2001). In recent years, the application of network-based drug design has promised cancer therapy to employ a more rational approach (Hornberg *et al.*, 2006). This has demanded a thorough and solid understanding of the multiple genomic events that lead to the origin of the disease. However, an intuitive perspective alone would certainly not be enough to comprehend the highly complex processes and feedback loops involved. In order to gain novel insights into the functional differences between a normal cell and a cancer cell, it is extremely important to investigate the extensive cross-talk between the overall network of pathways instead of merely observing elements of the individual pathways associated with encoding the affected genes (Laubenbacher *et al.*, 2009).

A more effective and tractable approach would be to observe and analyse events at a molecular level for the sake of ascertaining how cancer manifests at the organismal level. Creating a quantitative model of the network of interest could provide a beneficial tool for understanding existing biological data, aiding the generation of hypothesis and also helping to guide future laboratory experiments. This integration of biological and mathematical techniques into one coherent unit, for the purpose of breaking through the current constraints of reductionism, is one of the primary goals of

systems biology (Kitano, 2002). Around the 1960s, the research community began to support the notion that systemic methods were necessary to comprehend biological systems (Bertalanffy, 1968; Rosen, 1958a, 1958b). For instance, Henrik Kacser proposed the idea of using chemical kinetics to understand biological organization (Kacser, 1957, 1960). Systems biology, a novel discipline that has only become relatively established around the beginning of the 21st century, has already become an influential paradigm for biological investigations. The discipline was developed to advocate a systems-oriented perspective in order to fundamentally complement the well established reductionist approach to research. Furthermore, it accounts for systems where component-component interactions play an essential role in governing the dynamics that are a consequence of the components themselves (Ahn *et al.*, 2006). Therefore, systems biology may provide a good framework for studying the complex nature of cancer.

In the scientific community, the definition of exactly what systems biology entails can be rather varied in the details from one fellow academic to another. For instance, some view the discipline as the integration of experimental and computational research (Kitano, 2002), whilst others assert that systems biology involves the formulation of mathematical models that exemplify the structure of a biological system in response to individual perturbations (Ideker *et al.*, 2001). Nevertheless, the consensus appears to be that systems biology is a discipline that aims to clarify the elusive connection between molecules and physiology by delving deep into the complete set of dynamic interactions amongst system entities (Kitano, 2002; Westerhoff, 2005; Bruggeman and Westerhoff, 2007).

Since the advent of molecular biology in the latter half of the previous century, significant advancements have been made in understanding exactly how the individual molecule operates. The venture into the molecular biosciences has culminated in the successful milestone of the human genome project circa 2001, which has enabled the multiple variations of thousands of human genes to be entirely mapped and sequenced (Lander *et al.*, 2001; Venter, 2001).

Molecular biology has helped to drive a powerful reductionist approach to biological research, which has aimed to elucidate complex systemic events by examining the individual components themselves and defining their functional properties. In spite of this approach producing valuable information for modern medicine, reductionism has a key limiting factor in that it does not appreciate the system as a whole and only observes the effects when components are working in isolation. Emergent properties, that are imperceptible when only the parts are studied, can arise when the biological system is viewed in its entirety. These functional properties include metabolic steady states and the cell cycle, which must be studied carefully to impart the knowledge necessary for manipulating the progression of multifactorial diseases such as cancer (Bruggeman and Westerhoff, 2007).

To put it succinctly, the whole is indeed greater than the sum of the parts. Thus, a systems-oriented view is required to provide a satisfactory explanation for how function arises from the individual system components, which can be achieved by employing aspects of systems biology.

In order to put into perspective the high complexity of cancer as a disease, consider the following factors that have to be taken into account when choosing a potential drug target.

- Multiple sequential events such as protein binding and allosteric inhibition give rise to signalling pathways. These pathways form a large and intricate communication network that generates signals essential to cellular activity. This non-linear interaction between the signalling pathways often produces unpredictability with regards to signal behaviour (Weng, 1999).
- Feedback loops play an important role in regulatory pathways in that they regulate components upstream in the pathway, enabling the coordination of cellular signals for division and growth. In particular cases, such as the p53 pathway which is involved in cell proliferation and apoptosis, many feedback loops are present in order to act as a backup system to reduce the phenotype of mutations (Harris and Levine, 2005). The existence of multiple feedback loops can cause the dynamic complexity of the system to quickly overwhelm a biologist's ability to perform the relevant analysis at an intuitive level.
- The varying features between different cancer cells, which include their shape or size and the current phase of the cell division cycle they are undergoing, all contribute to the diverse range of responses possible from just one specific signal (Hornberg *et al.*, 2006).

The examples above highlight the fact that the occurrence of interactivity between biological mechanisms will frequently result in unexpected dynamical behaviour. Since cancer itself transpires when biological system properties fail to exert themselves properly, studying

cancer with a systems biology mindset may significantly benefit the current framework of oncological research.

Generating quantitative models that accurately reflect the dynamics of individual pathways and their interactions will enable the cell to be represented *in silico*. Consequently, the model can then be employed to examine areas of interest, for instance, identifying which reaction in a network possesses the highest control over the system flux. A fundamental step towards the creation of a mathematical model would be to acquire a firm understanding of the eukaryotic cell cycle process and the regulatory mechanisms involved in cell proliferation. This chapter will review the relevant background information being applied to the mammalian cell cycle and metabolism model being developed, as detailed in chapters 3 and 4.

1.2 Cancer And The Eukaryotic Cell

Eukaryotic cells are composed of several distinct structures including the plasma membrane, extracellular polymeric components known as the glycocalyx, the cytoplasm, the cytoskeleton and subcellular organelles surrounded by a membrane themselves. Some of these organelles possess particular functions for the maintenance of cells and the production of biological components. For example, the nucleus, containing chromosomes and the nucleolus, is where RNA is synthesised. RNA is used to encode the information of genes in the DNA in order for proteins to be synthesised in the cytoplasm. RNA and proteins are also the building blocks for another organelle called the ribosome, which is responsible for the synthesis of

proteins. The mitochondrion is another highly important organelle and is involved in cellular metabolism. It is responsible for generating most of the cell's energy supply through the production of adenosine triphosphate (ATP).

The onset and progression of cancers cells arise from a significant change in the biomechanical and biophysical properties of cells and their subcellular structures. For instance, genetic mutations may lead to the abnormal proliferation of tissues. If these particular cells reside at their initial location, then the tumour is referred to as benign. However, a cell becomes malignant through a multi-stage process where several crucial traits are acquired via the occurrence of numerous different types of mutations (Laubenbacher *et al.*, 2009). This results in a change of dynamics for the various molecular pathways that have a connection with the mutated gene. Consequently, the cell develops the capability to perform certain actions as follows.

- The cell can now completely evade apoptosis, which refers to programmed cell death. Apoptosis is essential for maintaining a healthy physiology by eliminating old cells or unwanted cells containing damaged DNA. Failure to carry out the process of apoptosis could lead to uncontrolled cell growth and the formation of tumours.
- Become insensitive to anti-growth signals that usually prevent cell proliferation.
- Achieve self-sufficiency with regards to growth signals. For example, cancer cells have the ability to interact with neighbouring normal cells and cause them to yield a very high level of signalling related to growth stimulation factors (Skobe and Fusenig, 1998).

- Maintain a prolonged process of angiogenesis where the cancerous cells are able to spread to other tissue areas through the development of these new blood vessels.
- Acquire the potential to carry out unlimited cellular replications.
- Gain the ability to penetrate and infiltrate both adjacent and non-adjacent tissues through the process of metastasis (Hanahan and Weinberg, 2000).

These effects can all be traced back to the occurrence of mutations in certain genes. A gene that is quiescent most of the time but is capable of giving rise to cancer if activated is identified as a proto-oncogene. Oncogenes, the carcinogenic form of a proto-oncogene, may arise in a variety of ways. For instance, exposure of the cell to ultraviolet radiation or chemical carcinogens and the invasion of DNA tumour viruses can all cause genetic damage leading to point mutations, which is the alteration of a single nucleotide base pair, and the eventual production of a modified protein that can not be regulated properly. Another important gene called a tumour suppressor encodes for proteins that inhibit cell division. In the event a tumour suppressor becomes damaged, the downstream signal for halting cell division is lost and cells begin to divide uncontrollably leading to the formation of tumours.

A cancer gene census was compiled in 2004 (Futreal *et al.*, 2004) and of the 300 genes currently documented, it was found that most genes were implicated in the progression of the cell cycle (Vogelstein and Kinzler, 2004). As a result, an effective method of restricting the growth of cancer cells selectively would be to carefully manipulate aspects of the cell division cycle. For example, cancer cells divide at a significantly faster rate than cancer cells. Therefore, a drug compound could be designed to increase the

duration it takes to complete one full cycle of cell division in cancer cells until it is equal to or greater than that of normal cells.

1.3 The Cell Cycle

The ultimate aim of the cell cycle is to divide a single parent cell into two daughter cells, complete with their own set of nucleus, cytoplasm and genetic components identical to the parent cell. This cycle is divided into several well defined stages and is commonly allocated four phases called G1, S, G2 and finally M. The cell cycle process is shown in Figure 1.1.

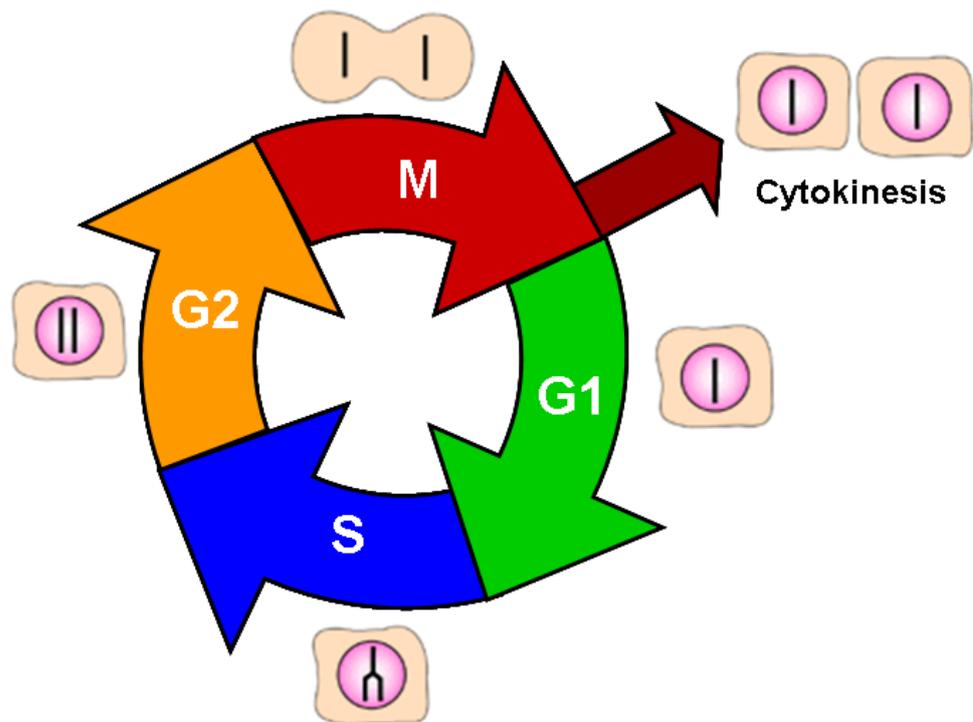


Figure 1.1. The phases of the mammalian cell cycle.

The first phase of the cell cycle, G₁, allows for the growth of chromosomes and prepares them for the replication of DNA. In the absence of sufficient growth factors or nutrients, the cell usually enters into a phase called G₀, which corresponds to a state of quiescence, although the cell itself may still be carrying out certain functions such as defending against pathogens. Here, the cell will remain until a certain threshold of internal and external growth conditions has been reached, stimulating the cell to re-enter and proceed with the cell cycle process once again. The difference between normal and cancer cells is that the latter does not enter G₀ and repeats the cell division cycle for an indefinite period of time.

Once the first growth phase has been completed, the cell then enters into a stage known as S where various cellular material including DNA and histones, which aid in the packaging of new DNA, are synthesised. In addition, each of the 46 chromosomes are replicated in two sister chromatids bound together by cohesin.

The third phase of the cell cycle, known as G₂, involves another period of growth and helps the cell prepare for division of the nucleus. Here, the chromosomes are subjected to rigorous checks in order to ensure that there have been no division errors with repairs being carried out if necessary.

The final stage of the cell division cycle is called M or mitosis. M itself is divided into several phases. The first of these phases is called prophase and it takes up over half the period required to complete mitosis. During prophase, the mitotic spindle is formed, chromosomes are condensed into compact structures and the nuclear envelope breaks down. The next two

stages are prometaphase and metaphase, where the chromosomes migrate and align at the spindle equatorial plane respectively. Anaphase follows metaphase and occupies the least portion of mitosis. The spindle fibres pull and separate the sister chromatids, which then move to the opposite ends of the cell. The spindle fibres dissolve and the nuclear envelope makes a re-appearance around the newly formed daughter chromosomes during telophase, the final stage of mitosis. Cytokinesis marks the end of the cell cycle when the nuclear membrane has properly formed around two identical daughter cells, each complete with cytoplasm and nucleus (Alberts *et al.*, 2002).

An alternative way of viewing the cell cycle is through a mathematical perspective as shown in Figure 1.2. The cell cycle in general may be thought of as an alternation between two self-maintaining stable steady states: G1 and S-G2-M (Nasmyth, 1996; Tyson *et al.*, 2001). In this scenario, the transition from G1 to S-G2-M is defined as “Start” and the move from M back to G1 as “Finish”. This view of the cell cycle implies that once the cell commits to the phase of DNA replication, the division process becomes irreversible. These two stable steady states are sustained primarily by the fundamental antagonism between two families of proteins, the cyclins and the cyclin-dependent kinases (CDK). The irreversible transition from G1 to S is due to positive feedback loops in the CDK control system. CDKs are enzymes that attach phosphate groups to specific target proteins and only become active when combined with cyclins, which act as regulatory subunits for CDKs, to form enzymatically active heterodimers.

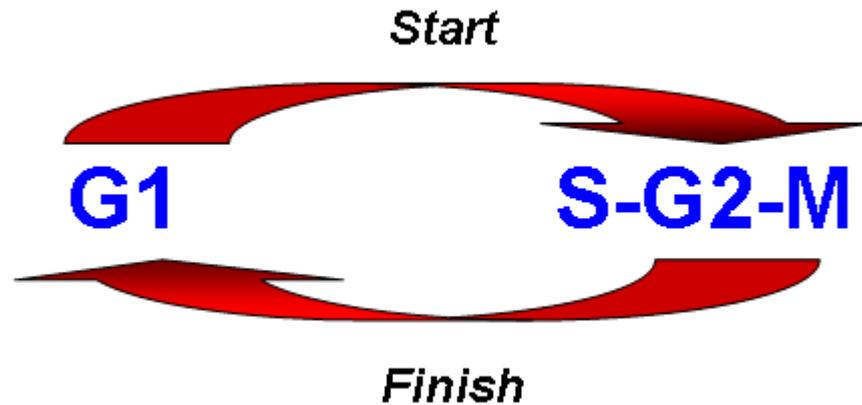


Figure 1.2. The two self-maintaining stable steady states of the cell cycle.

Cell cycle progression can be controlled by ordered activations of cyclin/CDKs in response to specific stresses. During G1, cyclin production is inhibited leading to their rapid degradation. As a consequence, CDK activity becomes low due a lack of available cyclin partners. Conversely, the Start transition is facilitated by an increase in CDK activity as a result of cyclin synthesis being promoted. The high level of CDK activity is maintained throughout the S to M phases of the cell cycle to cater for the requirements of DNA replication until Finish is reached. Finish is accomplished by the activation of the anaphase promoting complex (APC), a piece of proteolytic machinery, and the aid of two helper molecules (or auxiliary proteins), Cdc20 and Cdh1. Towards the end of the cell cycle, the combined efforts of APC and Cdc20/Cdh1 label cyclins in order to target them for degradation. Consequently, APC and CDK themselves are mutually antagonistic proteins. The APC inhibits CDK activity by destroying its cyclin partners, whereas the cyclin/CDK complex phosphorylates Cdh1 to limit the activity of the APC. An over-abundance of cyclin/CDK can also be resolved by the stoichiometric binding with a cyclin-dependent kinase inhibitor (CKI). In particular, p27

plays an important role in controlling cell cycle progression at G1 by binding to and preventing the activation of cyclin E/CDK2 or cyclin D/CDK4 complexes. The expression of this inhibitory subunit is tightly coupled to cell growth and it is often referred to as a cell cycle inhibitor protein because its major function is to slow or halt the cell division cycle (Alarcón *et al.*, 2004).

1.4 Regulation Of The Cell Cycle

The time taken to complete a single cell cycle varies between different cell types. For instance, embryonic cells have the ability to finish a whole cycle in minutes while some liver cells can take up to a year to complete the division process. Even cells that are genetically identical may exhibit heterogeneity with regards to the timing of their cell cycle progression (Smith, 2005). In general, the length of the cell cycle for fast-dividing mammalian cells is approximately 24 hours in accordance with circadian rhythms (Alberts *et al.*, 1997). Regardless of duration, it is imperative that cell cycle events transpire in a specific order for successful cell division to occur. To ensure this, the cell has a checkpoint system in place. Checkpoints in G1 and G2 verify whether or not damage has occurred to DNA. Damage to DNA before the cell enters S leads to CDK2 activity being inhibited, which halts the cell cycle until the problem is rectified or triggers entry into apoptosis if damage is too severe. Damage to DNA after S impedes CDK1 action and prevents the cell from entering M. Another type of checkpoint available are the spindle checkpoints, which detect if the spindle itself has been improperly aligned and blocks entry into cytokinesis.

In mammalian cells, the G1 checkpoint is known as the restriction point (RP) and is vital in controlling the progression from G1 to S. The RP is defined as a point in the cell cycle that separates two functionally different parts of G1 in continuously cycling cells, the postmitotic interval (G1-pm) and pre-S phase interval (G1-ps) of G1. Under suboptimal growth conditions the cell will exit the cell cycle at the RP and remain in G0 until a critical level of mitogenic stimulants is restored (Pardee, 1974). Thus, the RP acts a mechanism for survival under nutritional and growth factor deprivation.

Cell size, cell-cell interaction and the presence of nutrients all factor in the manipulation of signals that coordinate the shift between a proliferative and quiescent state. Experiments performed by Zetterberg and Larsson confirmed the existence and precise location of the RP (Zetterberg and Larsson, 1995). Cycloheximide (CHX), a biological compound that inhibits protein biosynthesis by blocking peptide chain elongation on eukaryotic ribosomes, was used in these experiments to ascertain the RP position. Table 1.1 demonstrates how the position of the restriction point was defined by treating the cell with CHX at three major treatment points during the cell cycle.

Table 1.1. Effect of cycloheximide on the cell cycle during different treatment times.

Effect	Treatment Time
Long delay in cell cycle division	Before RP
No effect on current or subsequent cell cycles	Directly after RP
Current cycle unaffected, significant delays in subsequent cycles	Late in cell cycle

A breakdown in the regulatory network that controls the RP may eventually lead to the malignant transformation of cells *i.e.* those that should be quiescent now actively proliferate. Failure of regulatory mechanisms can be instigated by a variety of factors such as a loss of CDK activity, overexpression of cyclins and the mutation or deletion of the p53 gene, a tumour suppressor. Restoring proper RP control to cancerous cells could enable them to return to a quiescent state. However, an alternative strategy for the treatment of cancer may be to use cytotoxic compounds that specifically target malignant cells by exploiting their uncontrolled proliferative nature in order to facilitate apoptosis. Employing a mathematical model that effectively predicts the biological behaviour of the cell cycle

should aid the development of existing treatment strategies and help researchers develop novel ones as well. A discussion of cell cycle models is given in Chapter 2.

1.5 Cell Metabolism And Growth

Successful growth and division of cells requires certain precursors. For instance, energy is necessary to drive the whole proliferation process and non-essential amino acids are needed to act as biological building blocks for macromolecules. In order to supply the cell with energy, cells must undergo multiple enzyme catalysed chemical reactions, collectively referred to as cellular metabolism.

Metabolic pathways contain two fundamental types of interdependent reactions. Catabolic reactions result in the release of energy when a larger complex organic compound is degraded into smaller constituents. Anabolic reactions are those that involve the assemblage of smaller simpler compounds into larger macromolecules. This type of constructive metabolism, also called biosynthesis, essentially depends on an input of energy in order to function properly. In general, the products obtained after catabolism will be required during anabolism.

Adenosine Triphosphate:

One of the primary sources of energy for cells comes in the form of the molecule specialised in storing energy after its release during catabolic

reactions and is known as adenosine triphosphate or ATP. Adenosine triphosphate is a nucleotide that is synthesised in the mitochondria. ATP is formed from a nitrogenous base, adenine and ribose sugar, which is bound to a chain of three phosphate groups. This highly important macromolecule acts as the energy currency for cells and is responsible for driving a myriad of processes related to transport, mechanical and chemical reactions. Through the process of phosphorylation, ATP transfers a phosphate group to another biomolecule to yield energy needed for cellular activities. The loss of the phosphate group renders the molecule to adenosine diphosphate (ADP), which contains only two phosphate groups, unlike ATP that has three. ADP is recharged into ATP through coupling to catabolic reactions, mostly in glycolysis (in the cytoplasm) or oxidative phosphorylation (in the mitochondrion).

Glycolysis:

The catabolic process of glycolysis is critical to mammalian cells for the release of energy. Glycolysis is ubiquitous to virtually all living cells. In eukaryotic cells, it is carried out in the cytoplasm. The aim of glycolysis is to convert a single molecule of glucose into two molecules of pyruvic acid, two energy-rich electron carrying molecules called nicotinamide adenine dinucleotide (NADH) and two molecules of ATP.

The metabolic pathway of glycolysis contains a series of enzymatic reaction steps. The first of these steps involves an enzyme called hexokinase catalysing the conversion of glucose into glucose-6-phosphate. ATP

phosphorylates glucose resulting in a sugar phosphate. Due to the negative charge of the phosphate, the sugar phosphate is unable to move through the plasma membrane. This effectively forces the glucose molecule to remain inside the cell. Phosphoglucose isomerase catalyses the second step of glycolysis where glucose 6-phosphate is converted into fructose 6-phosphate.

Phosphofructokinase is the enzyme involved in the third step of glycolysis and aids in the transformation of fructose 6-phosphate into fructose 1,6-biphosphate. The resulting product is a double phosphorylated six-carbon sugar that is separated during step four of glycolysis by the efforts of fructose biphosphate aldolase into two three-carbon molecules known as dihydroxyacetone phosphate and glyceraldehydes 3-phosphate. Only the latter product is allowed to automatically progress through glycolysis, whereas the former compound has to undergo a further step where it is transformed into its isomeric form, glyceraldehyde 3-phosphate, by the actions of triose phosphate isomerase. Glyceraldehyde 3-phosphate dehydrogenase catalyses the next reaction step that forms 1,3-biphosphoglycerate.

The first of the ATP molecules generated during glycolysis appears as a side-product in the reaction that has 1,3-biphosphoglycerate being converted into 3-phosphoglycerate by the enzyme, phosphoglycerate kinase. Phosphoglycerate mutase then helps transform 3-phosphoglycerate into 2-phosphoglycerate. In the penultimate step of glycolysis, water is removed from 2-phosphoglycerate in combination with the enzymatic actions of enolase to form phosphoenolpyruvate. The second ATP molecule is obtained during the final stage of glycolysis that has pyruvate kinase catalysing the

conversion of phosphoenolpyruvate into pyruvate (Alberts *et al.*, 2002). This results in the net gain of two ATP molecules from a single molecule of glucose.

Non-Essential Amino Acids:

In addition to a requirement for energy, cell growth also requires amino acids to serve as the basic units of proteins. Amino acids are classified as either nutritionally essential or non-essential. Essential amino acids are also referred to as indispensable because they are not synthesised by human cells and therefore must be introduced by dietary means. The group of essential amino acids consists of isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, valine, histidine and arginine. It should be noted that arginine is considered nutritionally essential only in young and growing mammals and not in adults. In contrast, non-essential or dispensable amino acids can be produced in the eukaryotic body from different carbon and ammonia sources (Lehninger, 1982). The mammalian cell cycle and metabolism model described in chapter 3 focuses on non-essential amino acids in particular.

Non-essential amino acids play important roles in cell signaling, regulating the expression of genes and can also function as a major source of metabolic fuel. Glutamate, glutamine and proline are three prominent examples of non-essential amino acids.

Glutamate is a negatively charged amino acid that is often involved in active or protein-binding sites as their negatively charged nature enables

them to interact with positively charged amino acids, creating stabilising hydrogen bonds that are critical to protein stability. It is actually the ionic form of glutamic acid after the removal of the proton. The final step in the formation of glutamate is carried out by the enzyme, glutamate dehydrogenase, from ammonia and 2-oxoglutarate, also referred to as alpha-ketoglutarate.

Glutamine is the most abundantly occurring amino acid in the mammalian body and can be found in all the organs and muscles. It has various functions including building protein in muscles, restoring glycogen levels and maintaining the immune system. Glutamine synthetase, also called glutamate-ammonia ligase, is the allosteric enzyme that transforms glutamate into glutamine.

Proline is also an eventual product of glutamate. It plays an important role in intracellular signaling. Glutamate is first converted into glutamate 5-semialdehyde, which is then reduced to lead to the production of proline. The relationship between glutamate, glutamine and proline is shown in the biosynthetic pathway diagram in Figure 1.3.

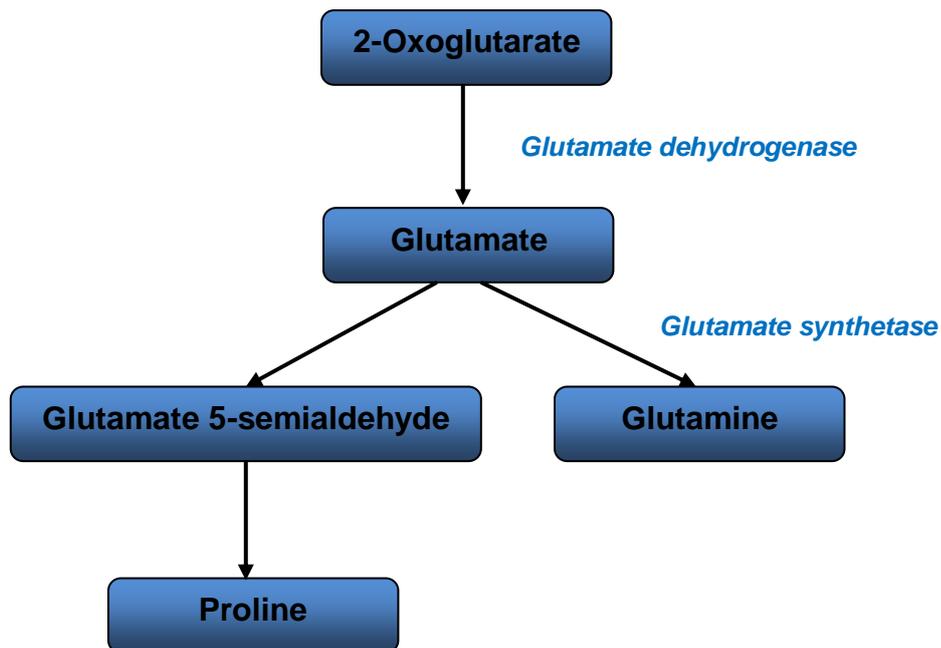


Figure 1.3. The biosynthetic pathway of glutamate, glutamine and proline.

Alanine and aspartate are two other non-essential amino acid synthesised in mammals. They are both the result of transaminations from glutamate. A transamination reaction is where an amino acid and keto acid exchange an amine for a keto group. Consequently, the amino acid now becomes a keto acid and vice versa. Alanine is a an amino acid with a hydrophobic side-chain that is generated by the reductive amination of pyruvate. Aspartate is converted from oxaloacetate by a transaminase enzyme. In mammalian cells, another non-essential amino acid called asparagine is derived from aspartate, catalysed by the ATP-dependent enzyme known as asparagines synthetase.

Serine, glycine and cysteine are non-essential amino acids that share a common biosynthetic pathway as shown in Figure 1.4. Serine phosphorylation facilitates the signal transduction process, while glycine and cysteine act as biosynthetic intermediates to various macromolecules. The formation of serine starts with the conversion of 3-phosphoglycerate, an intermediate of glycolysis, into 3-phosphohydroxypyruvate aided by phosphoglycerate dehydrogenase. Phosphoserine aminotransferase then catalyses the reaction involved in transforming this particular substrate to yield 3-phosphoserine. Phosphoserine phosphatase carries out hydrolysis to convert 3-phosphoserine into serine. From this point, serine acts as a precursor to glycine and cysteine and requires the enzymatic actions of serine hydroxymethyltransferase and cystathionine beta/gamma-lyases respectively.

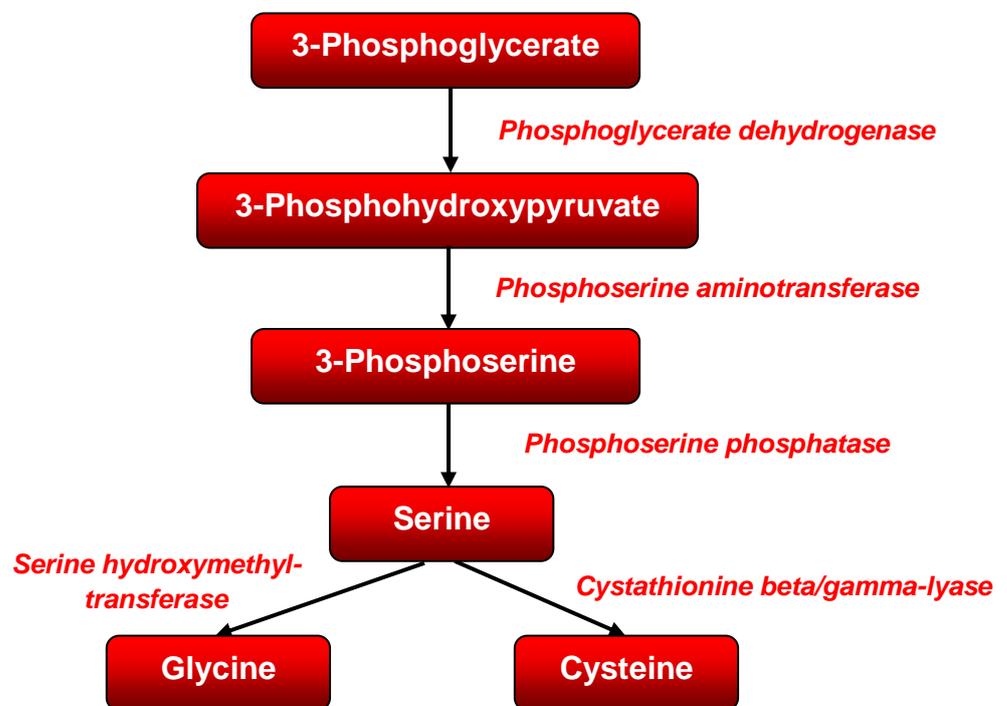


Figure 1.4. The biosynthetic pathway of serine, glycine and cysteine.

The final non-essential amino acid that has not been mentioned so far is tyrosine. It plays a structural role in almost every protein found in mammalian bodies. Although tyrosine itself is non-essential, it is actually formed from the essential amino acid, phenylalanine, with phenylalanineoxygenase catalysing the reaction. Therefore, tyrosine has not been included in the group of non-essential amino acids employed in the model of mammalian cell cycle and metabolism.

1.6 Normal vs. Cancer Cell Metabolism

An important distinction between normal and cancer cells is the dynamics involved in their metabolic pathway. Normal cells and cancer cells differ in the method they obtain the energy required for cellular processes from glucose. Most differentiated cells derive their energy through a combination of glycolysis and mitochondrial oxidative phosphorylation, whereas tumour cells predominantly utilise glycolysis with significantly reduced oxidative phosphorylation for the metabolism of glucose.

The majority of differentiated cells convert glucose to carbon dioxide by oxidising pyruvate in the tricarboxylic acid cycle (TCA), commonly referred to as the citric acid cycle, which then produces the reduced form of NADH necessary to drive the oxidative phosphorylation process (Alberts *et al.*, 2002). Lactate production is minimised for normal cells when there is ample oxygen available. Only a high level of lactate is produced when the differentiated cell is subjected to an anaerobic environment. For tumour cells on the other hand, there is a stark contrast to the dynamics of normal cell

metabolism in that large amounts of lactate are synthesised regardless of the presence or absence of oxygen. This preference of cancer cells to rely heavily on glycolysis even in the presence of high levels of oxygen has been shown to be a near-universal characteristic of cancer cells (Gatenby & Gillies, 2010). The phenomenon of the upregulation of glycolysis in tumour cells is known as the Warburg effect and is also referred to as aerobic glycolysis. A diagram illustrating the main differences between normal and cancer cell metabolism is shown in Figure 1.5.

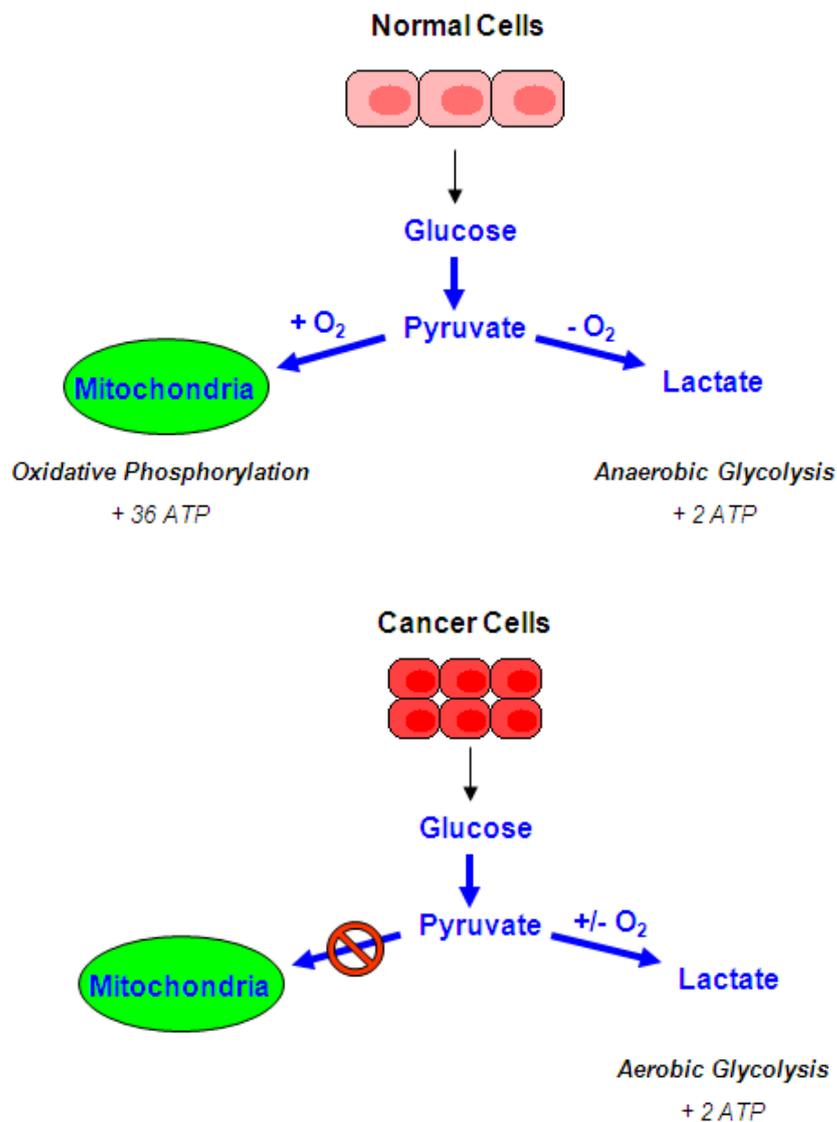


Figure 1.5. Diagram illustrating the differences between the metabolic pathways for normal and cancer cells.

For both normal and cancer cells, ATP is necessary for the transport of chemical energy within cells. During normal cell metabolism, oxidative phosphorylation generates 36 molecules of ATP per molecule of glucose, whilst 2 ATP molecules are produced from glycolysis. This means that for every glucose molecule, a total of 38 ATP molecules are generated. For tumour cells on the other hand, most of the glucose substrate is converted into lactate regardless of the level of oxygen. This process of aerobic glycolysis generates ATP in an inefficient manner with only 2 ATP molecules produced per molecule of glucose (Heiden *et al.*, 2009). However, aerobic glycolysis is a much faster process of producing ATP than is oxidative phosphorylation. This high ATP turnover rate may then be advantageously utilised by cancer cells to quickly fuel the conversion of important nutrients into the cellular biomass required for cell growth and proliferation.

There have been several explanations proposed as to why cancer cells prefer to heavily rely on glycolysis even in the presence of oxygen. Otto Warburg believed that the increase in glycolytic rates were due to irreversible damages to oxidative phosphorylation, which could have been caused by mitochondrial mutations. However, recent studies showed that inhibiting glycolysis in cancer cells may upregulate oxidative phosphorylation activity, which suggests that these malfunctions may not actually be irreversible. Another explanation that has been proposed is that cancer cells persist with glycolysis because this results in environmental acidosis, which is toxic to

normal cells whereas cancer cells remain unaffected, providing them with a significant proliferation advantage (López-Lázaro, 2008).

MODELS OF THE CELL CYCLE

2.1 Modelling The Cell Cycle

Understanding the molecular mechanisms involved in the cell division cycle and its regulation from a purely intuitive perspective is difficult due to the complexity of the regulatory system. However, a mathematical and computational approach may be applied towards comprehending the regulatory networks of the cell cycle in order to gain qualitative and quantitative insights into their method of functioning.

In general, models of the cell cycle and the mechanisms involved with its regulatory networks usually entail three stages. The initial stage is a phase of model design where the biological problem to be investigated is first identified and formulated into a well-defined mathematical model. This is the most crucial phase in the modelling process and can often be difficult and demanding. It consists of determining which pieces of information are relevant to the construction of the model and which ones are not important

and therefore should be omitted. The middle stage involves implementing the model using computational means to observe the effects on the biological system in response to various conditions. The information and data gathered should then be analysed to draw conclusions about the dynamic behavior of the cell cycle process. The final stage is model verification or validation, where the simulation results from the mathematical model are compared against the results from existing theoretical and experimental studies related to the current investigation but which had not been used to construct the model (Fuß *et al.*, 2005).

Model Design:

The first detailed step in the modelling process would be to create a mathematical model that provides a relatively accurate reflection of what is actually occurring within the cell cycle. This theoretical model should be able to reproduce certain aspects of existing experimental observations to the best of its ability. For example, a few key ideas that should be implemented in a basic cell cycle model include the following points.

- The model should be able to accurately reflect the response to checkpoint controls throughout the cell division cycle. For example, a lack mitogenic signaling would cause the cell to enter into a quiescent state until certain conditions are satisfied once again.
- The S phase of the cell cycle must always precede mitosis.
- Once the cell cycle passes “Start”, it must commit to the DNA replication process until complete as the transition from G1 to S is irreversible.

- A commitment must be made to remain in the mitosis phase until completion once the cell transitions from G2 to M.
- The cell should exit from mitosis at “Finish”
- For normal cells, simulations of the cell cycle time course should demonstrate balanced growth and division. For instance, the completion of a full cell cycle should result in the cell mass being halved to account for the new daughter cells.

Model formulation should always begin with a statement of the objective of the investigation by identifying the problem and determining any potential causes if possible.

During construction of a mathematical model, the system boundaries and components must be well-defined and any theoretical assumptions that are made to simplify the model should be within reason. Parameters and initial conditions should be specified during the model building process.

Cell cycle data for the mathematical model can be acquired from various sources. This includes reviewing existing literature relevant to the subject or gathering detailed information, such as the name of genes being expressed in a particular pathway, from online biological databases, such as the Kyoto Encyclopedia of Genes and Genomes (KEGG), and molecular interaction maps (Kanehisa, 1996; Kohn, 1999; Le Novère *et al.*, 2009).

Model Simulation And Analysis:

Computational tools can be employed to run simulations of the theoretical model in order to visualise the predictions made by the

mathematical model. Simulation can be a very powerful tool for studying and analysing complex systems. A simulation model imitates the operation of a real biological system as it evolves over time and is based on the form of a set of assumptions about the operation of the system. These assumptions are expressed as logical or mathematical relations between the system objects being investigated. One benefit of utilising a simulation model is the ability to calculate extreme values in addition to expected values. Hence, simulation can be used to gain a better understanding of the expected performance of the biological system and to determine the effectiveness of the design of the system.

A dynamic simulation, representing a biological system as it evolves over time, can be classified into two distinct types, deterministic or stochastic. A deterministic model contains no random variables, whereas a stochastic simulation model contains one or more random variables and can capture the effects of extrinsic and intrinsic noise. In most cases involving cell cycle research, the deterministic model is the preferred choice as it is usually sufficient enough to represent the general behaviour of the system, while a stochastic model is much more computationally intense.

A common problem with the formulation of a model is the lack of available information on the parameters needed to generate quantitative predictions of the behavior of the biological system. This ambiguity may be partially overcome by the use of parameter estimation techniques where information contained in accessible experimental data is used to extrapolate model parameters. However, a thorough and systematic review of existing

information is necessary to build a comprehensive and accurate model of the biological system.

In certain cases, researchers might not actually be aiming to investigate the behavior of the system based on specific parameters. They may be more interested in studying how the system is affected qualitatively when perturbations are applied to the parameters of interest. This particular investigation into how perturbed parameters influence and change the qualitative dynamical behaviour of a biological system is known as bifurcation analysis (Angeli *et al.*, 2004; Blüthgen *et al.*, 2001; Feller, 1994; Swat *et al.*, 2004; Tyson *et al.*, 2001).

Sensitivity analysis of model parameters has become another widely used technique in computational systems biology (Perumal and Gunawan, 2011). This technique has been employed in various fields of biological study, such as chemical kinetics, environmental modelling and the analysis of biological models (Horberger and Spear, 1981; Rabitz *et al.*, 1983; Rabitz, 1989; Saltelli *et al.*, 2005; Zheng and Rundell, 2006; Zi *et al.*, 2008).

Sensitivity analysis involves the use of sensitivity coefficients or values to observe the dependence of the mathematical model on the parameters involved. The sensitivity coefficients measure the ratio between the magnitudes of the resulting change in the model and the perturbations applied to the parameters that bring about this change. If the parameters have been subjected to infinitesimal perturbations, then the sensitivity analysis is referred to as local.

Global sensitivity analysis refers to a finite perturbation being applied to the model parameters and also allows all parameters to change

simultaneously (Marino *et al.*, 2008). The magnitude of the sensitivity values provides information on whether or not specific parameters heavily influence the behaviour of the biological system. A sensitivity value with a high magnitude would imply that the parameter has a great level of control over the cellular processes or pathway involved. Sensitivity analysis may provide potentially useful insights into the robustness of the mathematical outputs with respect to the varying model inputs and help identify the key factors affecting the model result. It can also be used to help determine the type of parameters estimation to be performed and to guide the design of future experimental work (Bentele *et al.*, 2004; Ingalls, 2008; Raue *et al.*, 2011, Rodriguez-Fernandez *et al.*, 2006).

Model Verification And Validation:

Once the analytical model is complete, the *in silico* results performed by the chosen simulation tool can be verified and validated by comparing them with existing information and data obtained from *in vivo* or *in vitro* experiments. This can help determine whether the mathematical model has worked as intended and if it is a credible representation of an actual biological system. After performing the comparison, the mathematical model can then be iteratively revised and adapted until a validated model of the dynamical system is accomplished.

After model verification and validation has been carried out, further analysis of the simulation results may help to identify potential new areas of theoretical study or suggest further experimental work to be carried out. The

model may be used to perform further simulations to explain non-intuitive problems about the biological system in question. In addition, predictions can be made to help guide future experiments by determining what areas need to be examined in more detail and what type of work should be carried out in the laboratory. A schematic diagram of the workflow involved in the cell cycle modelling process is shown in Figure 2.1.

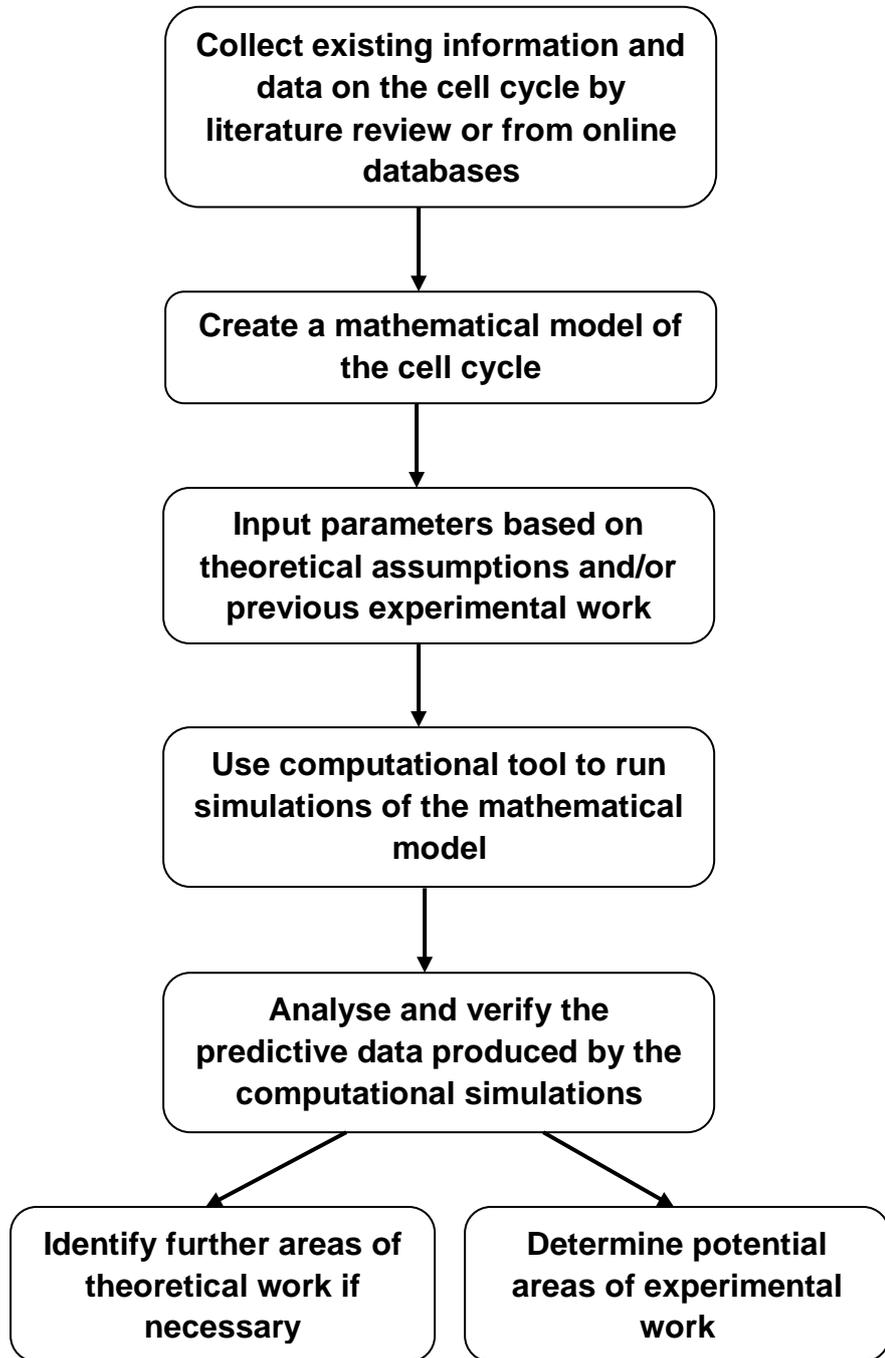


Figure 2.1. Schematic workflow of a cell cycle modelling process.

One major advantage of utilising mathematical modelling and computational techniques to observe the behavior of biological systems *in silico* is cost-effectiveness, compared to potentially expensive experimental work. In cancer therapy, a computational model can be used to predict the effects of drug compounds in order to determine which anti-cancer drug has the most influence over its intended target (Robert *et al.*, 2004). The model may also be used to rule out drug compounds before focusing on those that will target the pathway of interest. This is a much more effective and economical method than having to test the response of every drug compound in various *in vivo* trials. However, it is important to note that a mathematical model is only a representation of the biological system and that a certain degree of realism is usually lost during the construction process. Another benefit is the fact that mathematical models can be used to gain a better understanding of the system. Without a quantitative model, it would be impossible to comprehend the inner workings of some non-linear systems.

The simulation results of a mathematical model enables researchers to view exactly what happens over a period of time for the concentrations of the molecular species being investigated. One mathematical technique that can be employed to formulate the dynamical behavior of the cell cycle is the use of ordinary differential equations (ODEs), which are used in chemical kinetic theory to express the interactions between species related to the cell division cycle (Tyson *et al.*, 1996).

Ordinary Differential Equations:

Ordinary differential equations (ODEs) can be used to represent the deterministic relationship between the concentration of the species involved in the cell cycle and its regulatory networks, and their rate of change as time passes. If we define x to be the concentration of a certain species and t as time, then $x(t)$ can be used to denote the concentration of species x at time t . Now let F to be a given function of x , t and the derivatives of x . If the function F depends on more than one integration variable, the differential equation is then referred to as a partial differential equation (PDE). PDEs arise when we consider spatial location as well as time. If the location is ignored, then it means that the model assumes concentrations to be uniformly distributed in all of the space. PDEs require detailed information of the geometry of the spaces considered and it involves orders of magnitude with more degrees of freedom than ODEs and hence require greater computational power. Therefore, for this research project, as in most published cell cycle models, the attention will be restricted to ODE models.

The general form of the ODE describing what happens to a certain concentration over time can be expressed as equation (2.1).

$$x^{(n)}(t) = F\{t, x(t), x'(t), x''(t), \dots, x^{(n-1)}(t)\} \quad (2.1)$$

t = *The independent variable denoting time*

x = *The dependent variable denoting the concentration of a species*

n = *The order of the ordinary differential equation*

Here, $x'(t)$ is the derivative of x with respect to time and $x''(t)$ is the second derivative of x with respect to time. In general, $x^{(n)}(t)$ represents the derivative of order n for x with respect to time.

For a number of coupled ODEs that involve first order derivatives and represent multiple species with a role in the cell cycle, the system of equations can be expressed in column vector form by equation (2.2).

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_S'(t) \end{bmatrix} = \begin{bmatrix} F_1\{t, x(t), x_1(t), x_2(t), \dots, x_S(t)\} \\ F_2\{t, x(t), x_1(t), x_2(t), \dots, x_S(t)\} \\ \vdots \\ F_S\{t, x(t), x_1(t), x_2(t), \dots, x_S(t)\} \end{bmatrix} \quad (2.2)$$

S = *The number of species being modelled*

The system of ODEs can be solved using integration to obtain equations that describe the value of the dependent variable, x , as a function of the independent variable, t , given the value of x for a defined value of t . The resulting constants of integration are determined by taking into account the initial conditions defined in the mathematical model of the cell cycle.

Ordinary differential equations are classified as either linear or non-linear. A differential equation is linear when the dependent variable occurs only to the first power, as does its derivative. The solutions to linear type ODEs generate a vector space. Non-linear ODEs can be far more complicated to solve because they involve terms that include a dependent variable or the derivative of that particular variable to a power greater than one. The solutions obtained from non-linear ODEs can range from simple, complicated to chaotic.

In the event that it is not possible for an exact solution of an ODE can be obtained, numerical computation may be used to approximate the solution. Examples of numerical integration techniques include the Euler method for first order ODEs and Runge-Kutta for ODEs with a higher order (Atkinson, 1989; Butcher, 2003).

2.2 From Yeast To Mammalian Cell Cycle Models

At present, more is known in detail about the molecular machinery of cell cycle control for the unicellular organism, budding yeast or *Saccharomyces cerevisiae*, than any other eukaryotic organism. This experimental data-rich field makes budding yeast an ideal model organism for laying the foundations of mathematical modelling work for mammalian cells. In addition, studying the cell division cycle in yeast cells could potentially offer insights into the mutations that are the cause of cancer in humans due to similarities in the yeast and human cell cycles (Hartwell, 2002; Pray, 2008).

Kohn's Molecular Interaction Maps:

Kurt W. Kohn stated that “depicting the molecular networks involved in signaling pathway that regulate the cell function has proven challenging, due to the enormous amount of information that needs to be conveyed for each participant in the network and the cross-connection between pathways. This challenge must nevertheless be addressed in order to understand the underlying design of such networks, and to utilise the findings of modern biology most effectively to combat diseases...Another difficulty is that bioregulatory networks are replete with interconnections that make intuition about network function unreliable ” (Kohn and Aladjem, 2006a).

Kohn's statement refers to the fact that diagrams representing bioregulatory networks are often incomplete and ambiguous. For instance, simply having an arrow between two components to highlight their relationship may not detail the underlying mechanism involved, such as one component affecting the molecular state or quantity of the other. A complete and comprehensive diagram of complex bioregulatory networks must take into account three key ideas as follows.

- The binding of regulatory macromolecules such as proteins leads to new multi-compounds that may exhibit distinctive interaction behaviours.
- Enzymes used to catalyse specific chemical reactions are frequently found to be the substrates to other enzymes found in the network, irrevocably linking an interdependent system of enzymes together.

- Network molecules often undergo some form of modification such as being phosphorylated. This significantly affects their ability to bind to each other or to catalyse certain reactions.

In order to address the issue of comprehending the interactions between various elements of a complex biological system, Kohn developed a convention for molecular interactions maps that was capable of representing complex bioregulatory networks (Kohn, 1999). These molecular interactions maps employ specially designed and standardised notation. Each single molecular species in the diagram is represented at a unique location. Two types of interactions, the reactions and the contingencies, are denoted by well-defined arrowed or end-barred lines that connect the two species that are interacting. Reactions include processes such as the binding or cleaving of molecules. Events such as the stimulation and inhibition of enzymatic activity are referred to as the contingencies. Monomolecular species that bind together to form multi-complexes are represented by nodes, which are connected to each other by reaction or contingency lines. These resemble a large dot placed in the middle of the interaction line. A few examples of the notation used by Kohn to denote the interactions between species are given in Table 2.1.

Symbol	Representation
	Monomolecular species A
	Non-covalent interaction between species A and B (node in the middle represents the

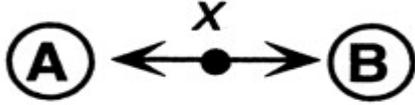
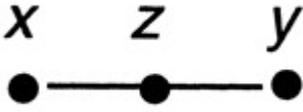
	A:B complex)
	A:B is denoted by x
	Species A contributes peptide that binds to receptor site on species B
	States x and y combine to form state z
	Degradation products
	Stimulation
	Inhibition
	Transcriptional activation
	Transcriptional inhibition

Table 2.1. Examples of the notation utilised in the molecular interaction maps by Kohn.

The molecular interactions maps allow the extensive protein-protein interactions and protein modifications to be expressed as well-defined canonical representations. Although this type of representation does not allow for explicit displays of temporal processes, Kohn's intricate biological circuit

enables proteins with functionally similar roles in both yeast and mammalian cells to be identified.

Since the publication of the widely cited 1999 paper, updates have been made to the notation used in the molecular interaction maps and published in a subsequent paper by Kohn and other researchers (Kohn et al., 2006b). Recently, Kohn's notation has become obsolete and the current accepted standardised notation for representing information about biochemical pathways is now the Systems Biology Graphical Notation (SBGN), for which Kohn is also a signatory. SBGN uses graphical representations to describe process descriptions, entity relationships and activity flows for the interactions of a biological network (Le Novère *et al.*, 2009). The objectives of SBGN are to be as simplistic and unambiguous as possible when representing the network of interest; be computationally tractable; make it convenient for researchers to draw the diagram by hand; be reproducible in black and white; allow for incomplete information and to follow accepted conventions whenever possible. These graphical representations act as a visual aid to help researchers understand what interactions are occurring and enable the information contained within the biological network to be summarised concisely.

The Chen *et al.* Yeast Cell Cycle Models:

The model presented by Chen *et al.* in 2000 elucidates the cell cycle of budding yeast by applying previously developed methods, such as ordinary differential equations, and drawing on the extensive data available

on this particular organism (Chen *et al.*, 2000). This biochemical kinetic model is used to describe how controlling protein concentrations vary during a proliferation cycle.

A set of 10 non-linear ODEs based on chemical kinetic rate equations govern the time-evolution for cell mass, the cyclins and their associated proteins. Three additional algebraic functions describe transcription factors. Events are also implemented to determine exactly when DNA synthesis, budding and spindle assembly should be triggered. The effect of the asymmetric division of the yeast cell mass once a full cell cycle has been completed is included in the model.

This model of kinetic equations comprises of around 50 parameters that include rate constants, binding constants, thresholds and relative efficiencies. Many of the parameters were derived from existing experimental data on kinetics, while others were manually fitted to training data reflecting genetic perturbations to the length of G1 and the cell mass at different times in the cycle. The aim of the budding yeast model was to have the equations and parameter set provide a reasonable and explicit explanation for the various properties of cell cycle control.

This model was able to reproduce a great deal of the observed physiology of *in vivo* cells, especially with regards to cell cycle regulation. Chen *et al.* predicted that the cell cycle exhibits hysteresis and bistability. In a more recent paper, Cross *et al.* showed this hypothesis to be true (Cross *et al.*, 2002). In particular, similar to the views of Nasmyth (1996) and Tyson *et al.* (2001), Chen *et al.* proposed that the G1 and S-G2-M phases were two alternative self-maintaining steady states. Cyclin-dependent kinases (CDKs)

drive the transition between the two states based on whether the concentration of cyclins available is high or low. This is caused by the mutual antagonistic relationship amongst the cyclin B-dependent kinases and their adversarial proteins, Sic1 and Cdh1.

Hysteresis was predicted in the transition between G1 and S-G2-M and vice versa. This means that the current state is dependent on the history of the previous state. Events that drive the transition between the two states, such as the interaction between the cyclins and APC/Cdh1, must exert themselves properly before the irreversible progression from one state to the next occurs.

The model developed by Chen *et al.* in 2000 provides an adequate representation of the transition from G1 to S. However, since its publication, additional information about the transition from M to G1, also known as the exit from mitosis, has arisen. These new pieces of information were added to the model by Chen *et al.* in 2000 and formulated into an improved version of the original model in 2004. This more recent model is able to quantitatively depict what is happening during all the stages related to the chromosome replication-segregation cycle in eukaryotic cells (Chen *et al.*, 2004).

The version of the budding yeast cell cycle model published in 2004 incorporates 36 ordinary differential equations containing 148 constants, in addition to algebraic equations. This particular model describes cell mass, DNA content, protein concentrations, the actions of the mitotic spindle, and details about the state of the emerging bud.

Chen *et al.* found that after subjecting some of the model parameters to manual fitting, simulations of the dynamical system behaviour were rather

adept at expressing properties of cell growth and division for wild-type cells. Another promising find was that the model was relatively effective in representing the phenotypes of more than 100 mutant strains with regards to certain experimental observations, such as the growth rate of the cell. However, in certain cases where specific parameters are modified to concur with existing experimental data on mutant cells, the effects on other mutants could be unexpected and unpredictable.

It is important to note here that in general, both the selection of a model for a particular biological system and parameter estimation of the chosen model must be dealt with carefully. For instance, the objective function, which is the sum of squared deviations between the predictive results and the experimental observations, may have multiple local optima in the parameter space. A specific set of parameters may offer a good model fit, but if this set contains parameter values that are actually unrealistic representations of their real-world counterparts, then the model might generate erroneous results when tested against data with different conditions.

The Novák and Tyson Mammalian Cell Cycle Model:

Novák and Tyson created a model for the restriction point control of the mammalian cell cycle by extending upon the Chen *et al.* budding yeast cell cycle model and exploiting the deep similarities between the CDK regulatory systems of yeast and mammalian cells as identified from Kohn's molecular interaction maps (Novák and Tyson, 2004).

This mammalian cell cycle model, consisting of 18 ODEs, is capable of recreating several cellular physiological responses such as the restriction point and the transient halting of cell growth through the inhibition of protein synthesis by the effects of cycloheximide. In addition, it takes into account the antagonism between p27 and the cyclin E/CDK 2 and cyclin A/CDK2 complexes.

Following on from the experimental work carried by Zetterberg and Larsson (1995), a term representing the effects of cycloheximide was included in the model in order to locate the position of the restriction point.

Recently, Yao *et al.* proposed that the RP is regulated by a bistable switch resulting from the interactions between the retinoblastoma protein (Rb) and E2F transcription factor (Yao *et al.*, 2008). Yao *et al.* of researchers discovered that if this particular bistable switch is on, a cell will continue to divide regardless of whether the cell is damaged or the growth signal has disappeared. In the Novák and Tyson mammalian cell cycle model, the Rb-E2F interactions are expressed as steady state algebraic equations. The model also satisfies the requirement of Cyclin E/Cdk2 oscillating in phase with the inactivated and phosphorylated form of Rb (pRb).

Several assumptions are made for the mammalian cell cycle model. Firstly, the ODEs are written exclusively for proteins while mRNAs are ignored because the mRNAs are always in steady state due to the postulated rapid message turnover. Secondly, the level of cyclin-dependent kinases (Cdk) are always assumed to be significantly higher than their cyclin partners so that they are not the limiting factor in the configuration of cyclin/Cdk complexes. Therefore, every cyclin/Cdk complex is defined by its cyclin

subunit and Cdk terms do not appear in the ODEs. Another assumption made is the existence of an intermediary enzyme (IE) that is responsible for causing a delay in the time it takes to accumulate Cyclin B before Cdc20 is activated. The phosphorylated version of IE, IEP, becomes dephosphorylated again by the actions of a phosphatase that is directly dependent on the efficiency of translation. For example, if the translation efficiency is working at half its potential full capacity, the activity of the phosphatase targeting the IE is also halved.

2.3 A Mammalian Cell Cycle Model

The mammalian cell cycle and metabolism model discussed in chapters 3 and 4 extends and builds upon the foundations of the work carried out by Conradie *et al.* (2010). Therefore, the focus of this particular section will be on detailing what is involved with their model of the mammalian cell cycle.

The mammalian cell cycle model developed by Conradie *et al.* is an extension of the mathematical model developed by Novák and Tyson. Certain techniques are used to formulate a method of using control analysis in order to identify which reaction steps were the most influential over the positioning of the restriction point (Conradie *et al.*, 2010). This model consists of a system of 23 coupled non-linear ODEs, each describing the rate of change of an important species involved in the mammalian cell cycle. The

concentrations of cyclins, p27, early response genes, delayed response genes and cell mass are amongst the variables being modelled to observe how they evolve over a period of time.

The increase in the number of ODEs found in this model compared to the 18 ODEs presented by Tyson and Novák is due to the fact that the Rb and E2F transcription factor interactions are now being expressed directly as ODEs instead of just steady-state algebraic equations. Rb is a tumour suppressor protein that binds to and inhibits E2F, which positively regulates many of the genes required for entry into the S phase of the cell cycle. The sequestering of E2F is carried out by cyclin D/CDK4 that phosphorylates Rb until it becomes inactive. E2F can then be released to enable the activation of genes required for initiation into the phase of DNA synthesis. Cyclin D production itself is controlled by a signal transduction pathway involving growth factors (GF), early response genes (ERG) and delayed response genes. This particular GF-ERG-DRG pathway is carefully regulated through the stimulation of the growth factors.

A combination of GF stimulation and the transcriptional induction of ERGs and DRGs results in a change of state for the cells found in the G0 phase. This allows the previously quiescent cells to re-enter the cell cycle to continue onwards with cellular growth and proliferation. This high level of feedback regulation required for proper cell growth and an ODE concerning the mass of the cell has been incorporated into the Conradie *et al.* model.

One similar aspect of the Conradie *et al.* model to the Tyson and Novák model is the involvement of 52 reaction steps contained within these ODEs. The series of equations utilise mass action kinetics to elucidate a

functional unit within the cell cycle network. Around 70 parameters are employed in the model to characterize rate constants, relative efficiencies and binding constants.

All the terms related to the rate of synthesis for proteins have a factor, ϵ , which denotes the translation efficiency of the ribosomes. This particular term is affected by translation inhibitors and growth factors. It is used to represent ribosomal activity in response to cycloheximide exposure. The value of ϵ ranges from zero to one. A very low value close to zero denotes low translation efficiency, while a value of one implies that the ribosome is working at full translation efficiency.

Conradie *et al.* Model Assumptions:

Several key model assumptions, inherited from the model developed by Novák and Tyson (Novák and Tyson, 2004), are made when deriving equations of the model.

- Binding and release reactions are relatively fast compared to the phosphorylation and dephosphorylation reactions.
- Cdc20 synthesis is assumed to be dependent on the level of cyclin B in order to model the effect of Cdc20 accumulating during the S-G2-M phases and receding during G1.
- The cyclin/CDK complexes involving cyclins A, B, D, and E all contribute towards phosphorylating Rb.

- The total levels of E2F and Rb (which includes both the hypophosphorylated and phosphorylated forms) remain constant during the cell cycle process.
- Messenger RNA (mRNA) is not allocated an ODE as rapid message turnover is assumed i.e. mRNAs are considered to be in steady state constantly.
- The levels of CDKs are assumed to be always higher than their cyclin counterparts. This abundance of CDKs compared to cyclins means that the formation of cyclin/CDK complexes is not dependent on the rate of CDK production.
- The value of the cell mass is exactly halved when the concentration of Cdh1 becomes greater than a critical value.

Conradie *et al.* Model Reactions Steps:

The 52 reaction steps and their corresponding rate equations used by Conradie *et al.* to construct their model of the mammalian cell cycle and its regulatory mechanisms are given in Table A.1 found in Appendix A, with V_r describing the rate equation for reaction step r .

In order to reproduce certain time course results from the paper by Novák and Tyson, Conradie *et al.* actually used a time re-scaling factor of 10, meaning that all ODEs of the system should be multiplied by ten. This factor is not displayed in the rate equations as shown in Table A.1, but it is important to note that the eventual time course plots generated by Conradie *et*

al. through simulation of the model record the units of the species concentration and time in arbitrary units and hours respectively.

Species such as Cyclin A, B, D and E are abbreviated to CYCA, CYCB, CYCE, CYCD respectively. The complexes formed by the cyclin and Kip1, which is the cyclin-dependent kinase inhibitor in G1, are denoted by CA, CE and CD with regards to the respective cyclins A, E and D. The concentration of a phosphatase is represented as PPX.

Conradie *et al.* Model ODEs:

ODEs used in this particular mammalian cell cycle model are shown in Appendix A. The NDSolve function and the Event-Locator method were employed in Mathematica, version 6.0 (Wolfram Research, Champaign, IL, USA, <http://www.wolfram.com>) to numerically solve the ODEs. Refer to Table A.1. for the relevant rate equations.

Conradie *et al.* Model Parameter Values:

The parameter values found in the Conradie *et al.* (2010) model are inherited from the mammalian cell cycle model by Novák and Tyson (2004). These particular parameters themselves were either taken from earlier models of yeast and frog cell cycles developed by Novák and Tyson or chosen to give an approximation of the *in vivo* mammalian cell cycle activity (Novák

and Tyson, 1993; Novák *et al.*, 1998, 1999). The values of the parameters are given in Appendix A.

An ODE For The Cell Mass:

The ODE that describes the cell mass over time, as given by equation (2.3), implies that the cells are accumulating cytoplasmic mass at an exponential rate before each subsequent cell division. There is an event implemented in the model that once the concentration of Cdh1 passes a certain threshold value, in this case 0.2, the value of the cell mass is to be divided by a factor of 2. This is to reflect the exact point in time during the cell cycle at which a mother cell divides into two identical daughter cells.

$$\frac{d[\text{Mass}]}{dt} = \mu * \varepsilon * [\text{GM}] \quad (2.3)$$

μ = *Growth rate*

ε = *Translational efficiency*

GM = *General machinery for protein synthesis*

Metabolic Control Analysis:

Metabolic control analysis (MCA) was developed in 1973 by Kacser and Burns to act as a rigorous framework for the purpose of assessing the behaviour of reaction steps in the context of a pathway (Kacser and Burns, 1973).

In mathematical terms, MCA is a 1st order sensitivity analysis in the areas around a stable, and structurally stable, fixed point. A control coefficient, as defined by equation (2.4), measures the relative steady state change in a system variable, such as the flux or species concentration, in response to a relative change in a parameter e.g. enzyme activity.

$$C_{\text{Parameter}}^{\text{Variable}} = \frac{\partial \log_e(\text{Variable})}{\partial \log_e(\text{Parameter})} \quad (2.4)$$

Conradie *et al.* (2010) implemented an extension to MCA by using a perturbation method to quantify the control of certain reactions steps involved in the cell cycle pathway on the time the restriction point occurs. With this method of control analysis in place, they concluded the restriction point is most sensitive to perturbations made to the cyclin E/CDK2:p27 (CE) complex. It was suggested that identifying the reactions steps with the greatest effect on the CE complex and perturbing these reactions by affecting them with relevant compounds could possibly shift the restriction point back to its normal position.

Control And Sensitivity Analysis:

In the Conradie *et al.* (2010) paper, the control coefficient of reaction r on the position of the restriction point (RP), C_r^{RP} , was defined as the percentage change in the fraction of the cell cycle length at which RP occurs for 1% change in the activity of (reaction step) r .

A positive value of C_r^{RP} implies the time of RP occurrence has been delayed in the cell cycle, whereas a value less than zero suggests the position of the RP has been advanced. A value of zero means there was no effect on the RP position.

In order to perform this control analysis, 52 constants, α_1 to α_{52} , were multiplied to the 52 existing rate equations involved in the mammalian cell cycle model. The explicit definition of this process is given in equation (2.5).

$$\{M_r\} = \{\alpha_r\} \cdot \{V_r\} \quad , \quad r = 1, 2, \dots, 52 \quad (2.5)$$

$\{M_r\}$ = *The set of r modified rate equations*

$\{V_r\}$ = *The set of r original rate equations*

$\{\alpha_r\}$ = *The set of r perturbation constants*

Each of these 52 modified rate equations represented by $\{M_i\}$ was perturbed individually by perturbation of 0.0000001 up and down. The model was then simulated with varying times at which the cells are exposed to cycloheximide. This allows the RP occurrence time to be identified by observing the first time point for which the cell mass begins to show stationary behaviour. Finally, control coefficients can be calculated using the RP occurrence times for the different perturbation values.

Different cell types exhibit a diverse range of times taken to complete one full division cycle. One advantage of this type of control analysis is that control coefficients are expressed as a fraction of the cell division cycle. Hence, the time scale is dimension-free when determining the value of C_r^{RP} .

Chapter
THREE

A MODEL OF GROWTH AND DIVISION

A vital aspect of cell growth and proliferation involves the usage of energy derived from nutrients. This includes the metabolism of glucose, which is the most utilised carbohydrate in mammalian cells as it provides a major source of energy necessary to fuel the process of cellular division. Previous research has also demonstrated that glucose takes on an important role in cancer cell metabolism (Li *et al.*, 2010).

The restriction of glucose metabolism has been considered as a powerful method of preventing the development of cancer. This is due to the fact that cancer cells appear to carry out the metabolism of glucose at an elevated rate compared to normal cells. Limiting the availability of glucose acts as a metabolism stress factor that activates several signal transduction pathways (Hammerman *et al.*, 2004). In addition, limiting the metabolism of glucose leads to significant changes in the expression of multiple genes including those associated with cell growth and survival (Gupta, *et al.*, 1997; Lee *et al.*, 1998).

In order to ensure the successful completion of a cell growth and division cycle, certain precursors such as non-essential amino acids, nucleotides and mRNA are also required. These molecules are mostly synthesised from glucose, either from intermediates of the glycolytic pathway or from intermediates of the downstream TCA cycle.

The model developed here focuses on combining an existing cell cycle model with the pathways involved in glycolysis and the biosynthesis of non-essential amino acids.

3.1 A Mammalian Cell Cycle And Metabolism Model

The processes of cell growth and proliferation requires a substrate such as glucose as a source of energy and biochemical precursors to synthesise new materials. However, current cell cycle models rarely account for the consumption of a substrate required to drive the cell division process (Gauthier and Pohl, 2011). The Conradie *et al.* mammalian cell cycle model does not take into account the requirements of glucose for the cell growth process. Therefore, an extended version of the Conradie *et al.* model has been developed that also incorporates elements of both glycolysis and the non-essential amino acids pathway, crucial for acting as the building blocks of proteins, which themselves are necessary for ensuring successful cell growth and division.

The aim of this research project is to construct a novel mammalian cell cycle and metabolism model that will respond to different concentrations

of glucose with different division times. The focus in this particular chapter is placed on modeling the division cycle for normal cells. Chapter 4 will discuss how a version of the model for cancer cells was created.

Amino acids are the foundation for the biosynthesis of proteins and are also required for the expression of genetic data. Non-essential amino acids, also known as dispensable amino acids, do not need to be obtained by dietary means because they are synthesised directly from ammonia and carbon sources already present. This particular group of amino acids include glutamate, glutamine, proline, aspartate, asparagine, alanine, glycine, serine, cysteine and tyrosine. Although tyrosine is a non-essential amino acid, it can only be synthesised by mammalian cells from the essential amino acid phenylalanine through the process of hydroxylation by phenylalanine oxygenase (Lehninger, 1982). Therefore, tyrosine is not incorporated into the non-essential amino acids pathways of the cell cycle and metabolism model discussed in this chapter.

The cell cycle and metabolism model builds upon the foundations laid by the Conradie *et al.* mammalian cell cycle model and extends it to incorporate certain properties of cellular metabolism, such as the pathways involved in glycolysis and synthesising non-essential amino acids.

The cell cycle and metabolism model consists of 36 species, containing around 125 parameters, that convey the concentrations of the mammalian cell cycle and related regulatory network species, the metabolites involved in glycolysis, the total level of non-essential amino acids and finally, the cell mass. The cell mass is a particularly important entry for the mathematical model as it represents the mass of the cell over time and in the

new model its production rate is dependent on the concentrations of glucose.

In the original Conradie *et al.* mammalian cell cycle model, as shown in Table A.1, reaction 33 described the accumulation of cell mass at an exponential rate of growth before the value of the mass halved. Figure 3.1 shows the schematic diagram for the cell mass in the original mammalian cell cycle model.



Figure 3.1. Schematic diagram for the cell mass from the Conradie *et al.* mammalian cell cycle model (original reaction number in blue)

The single reaction shown in Figure 3.1 has been modified and extended to incorporate the reactions involved in the glycolytic pathway and the synthesis of non-essential amino acids, as shown in Figure 3.2. The arrows displayed in Figure 3.2 show the direction of conversion from substrate to product, while the numbers highlighted in red indicate the specified reaction step number. The full names of the abbreviations used in Figure 3.2 are specified in Table 3.1.

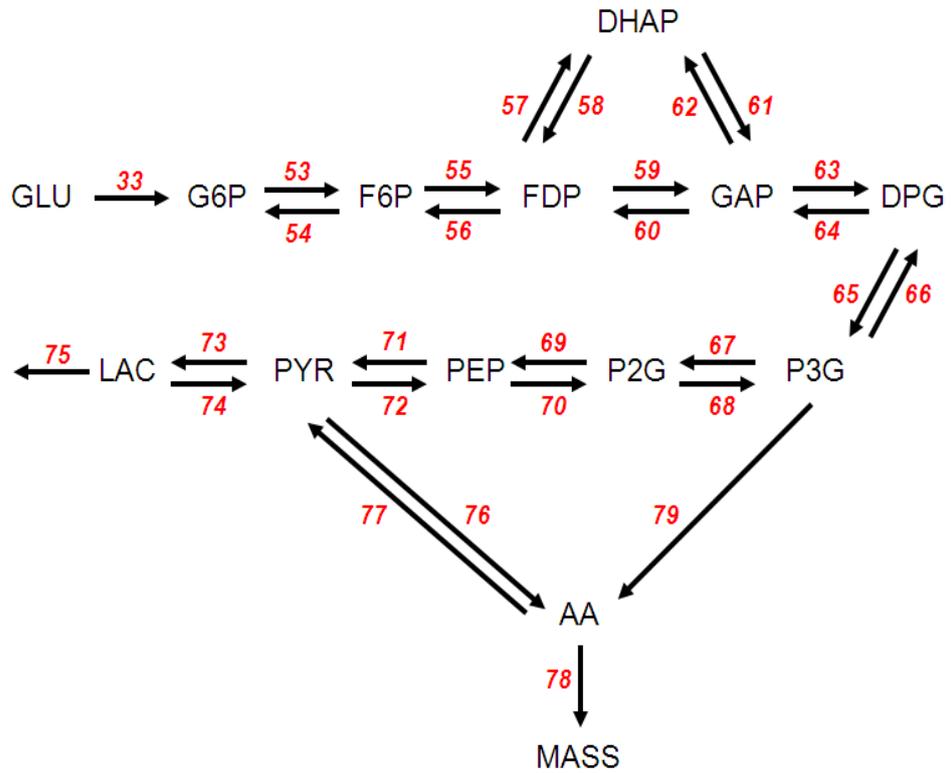


Figure 3.2. Schematic diagram for the extended model reactions involved in the cell cycle and metabolism model (reaction numbers in red)

Table 3.1. Abbreviations used in the cell cycle and metabolism model schematic diagram for model reactions as shown in Figure 3.2.

Abbreviation Used	Full Term
GLU	Glucose
G6P	Glucose 6-phosphate
F6P	Fructose 6-phosphate
FDP	Fructose 1,6-biphosphate
DHAP	Dihydroxyacetone phosphate
GAP	Glyceraldehyde 3-phosphate

DPG	1,3-bisphosphoglycerate
P3G	3-phosphoglycerate
P2G	2-phosphoglycerate
PEP	Phosphoenolpyruvate
PYR	Pyruvate
LAC	Lactate
AA	Non-Essential Amino Acids
MASS	Cell Mass

As shown by Figure 3.2, the species that are modelled for the pathway of glycolysis include glucose 6-phosphate, fructose 6-phosphate, fructose 1,6-biphosphate, dihydroxyacetone phosphate, glyceraldehyde-3-phosphate, 3-phosphoglycerate, 2-phosphoglycerate, phosphoenolpyruvate, pyruvate and lactate. A reaction has also been included to describe the degradation of lactate and its removal from the system.

The non-essential amino acids being modelled comprises of the total concentration of glutamate, glutamine, proline, asparatate, asparagines, alanine, glycine, serine and cysteine.

3.2 Model Implementation

The Cell Cycle And Metabolism Model ODEs:

The 36 ordinary differential equations that model the species involved in the cell cycle and its related metabolic pathways consists of 79 reaction steps in total. Their corresponding reaction rate laws are shown in Table B.1 found in Appendix B. Note that the time-rescaling factor of 10 used by Conradie *et al.* (2010) is displayed explicitly in the following rate equations. The discrete event that causes the cell mass to be halved when the Cdh1 concentration value becomes greater than 0.2 is kept for this particular model.

The Cell Cycle And Metabolism Model Parameter Values:

The parameter values pertaining to the division cycle and its regulatory mechanisms in the new model are inherited from the model by Conradie *et al.* (2010). New parameters are introduced to the metabolism side of the model and include the maximum metabolic rate constants for certain reactions and K_m values that represent the binding affinity between an enzyme and its substrate.

Formulation Of The Model Rate Law Equations:

As shown in Figure 3.2, the reaction rate law describing the conversion of glucose to pyruvate in the mammalian cell cycle and metabolism, V_{33} , is governed by a hyperbolic rate law akin to the Monod equation of growth, which is itself similar to the Michaelis-Menten enzyme kinetics.

The Monod equation can be used to express the growth rate of a population primarily as a function of the limiting nutrient concentration (Monod, 1949). The general expression of the Monod equation is given by equation (3.1).

$$\mu = \frac{\mu' \cdot [S]}{K_s + [S]} \quad (3.1)$$

μ = *The specific growth rate coefficient of the population of interest*

μ' = *denotes the maximum population growth rate*

K_s = *Half saturation coefficient*

S = *The concentration of the limiting nutrient*

The Monod equation can be applied to the cell cycle and metabolism model by defining our limiting nutrient specifically as glucose. V_{33} is therefore defined as follows by equation (3.2).

$$V_{33} = \frac{K_{33} \cdot [GLUCOSE]}{KM_{GLU} + [GLUCOSE]} \quad (3.2)$$

K_{33} = Maximum metabolic rate constant for reaction 33

KM_{GLU} = The affinity constant for glucose.

Here, KM is an inverse measure of the strength of binding or affinity between the catalysing enzyme and its substrate. The lower the value of KM , the higher the affinity.

Similarly, the rate equations describing the other species found in Figure 3.2 are expressed in the form of equation (3.3).

$$V_i = \frac{K_i \cdot [C_i]}{KM_{C_i} + [C_i]} \quad , i = 53,54 \dots, 76,77,79 \quad (3.3)$$

K_i = Maximum metabolic rate constant

C_i = Substrate concentration

KM_{C_i} = Affinity constant for the denoted substrate

Finally, the reaction rate law for describing the conversion of non-essential amino acids (AA) to cell mass, V_{78} , is given by equation (3.4)

$$V_{78} = \frac{K_{78} \cdot [AA] \cdot \varepsilon \cdot [GM]}{KM_{AA} + [AA]} \quad (3.4)$$

K_{78} = Maximum metabolic rate constant for reaction 78

KM_{AA} = Affinity constant for AA

AA = The total concentration of non-essential amino acids

Model Implementation in COPASI:

The simulation platform chosen to solve and analyse the ODEs of the mammalian cell cycle and metabolism model was COPASI, which is the abbreviated name for Complex Pathway Simulator (Hoops *et al.*, 2006).

The modelling and simulation tools available in COPASI are derived from the previously published GEPASI simulation platform (Mendes, 1993). COPASI is an open source software application that is available to download free online (<http://www.COPASI.org>). It was created through the collaborative efforts of computational biologists from the University of Manchester, UK, the Virginia Bioinformatics Institute, USA, and the University of Heidelberg, Germany.

COPASI is a software that is able to support models in the Systems Biology Markup Language (SBML). Therefore, mathematical models can be imported or exported into COPASI provided they are coded in the SBML format. SBML is a machine readable language, based on the Extensible

Markup Language (XML), and is generally used to decompose biological models into explicitly labelled constituent elements. COPASI is also able to write mathematical models into files that are compatible with several other computational formats such as Berkeley Madonna, XPPAUT and the C programming language.

COPASI can be used to construct a model of complex biochemical networks and their dynamics in such a way that the chemical reactions are easily translated into their corresponding rate equation forms through features such as events, compartments, global parameters and reaction rate laws. Certain tasks are also available in COPASI as a tool for analysing the results generated by the mathematical model, such as parameter scans, optimisation, metabolic control analysis and parameter estimation.

A comprehensive set of numerical algorithms are available in COPASI to simulate what happens to a certain model species over time in a quick and efficient manner. The types of time course solvers available can be classified into three main categories, a deterministic time course simulator, a stochastic simulator, such as the Gibson and Bruck solver (Gibson and Bruck, 2000), and a hybrid of the deterministic and stochastic methods that attempts to combine the best of both worlds. The choice of time course simulator should depend on the type of biological system being modelled. Stochastic simulators account for the introduction of noise in a system but can be very time-consuming for large and complex network models, so in certain cases it might be more prudent to choose a deterministic-type solver.

LSODA:

LSODA is deterministic-type solver based on the algorithm developed by Petzold (Petzold, 1983). The biggest advantage of LSODA is its ability to switch between non-stiff and stiff methods in response to the system dynamics. A non-stiff method is initially applied. If the problem then becomes stiff *i.e.* the step size of the numerical solution becomes significantly more limited by the stability of the numerical method rather than by the accuracy of the method, a dense or banded Jacobian is used instead to solve the system of equations. It is the default time course solver in COPASI.

The options included for LSODA are as follows.

- Integrate Reduced Model: A value of 1 notifies COPASI to use all the mass conservation laws and a value of 0 instructs COPASI to determine all the variables through the specified ODEs only.
- Relative Tolerance: A value determining the relative tolerance to be attained. The smaller the value, the more accuracy is achieved. The lowest possible numerical value that can be currently entered is around 2.22×10^{-16} . The default value is set at 1×10^{-6} .
- Absolute Tolerance: A value determining the absolute tolerance to be attained. The smaller the value, the more accuracy is achieved. The default value is set at 1×10^{-12} .

- Max Internal Steps: A value designating the maximum number of steps the solver is allowed prior to the next reported time point. The default value is set at 10000.

During this research project, LSODA, which automatically selects between non-stiff and stiff methods, was the chosen solver for providing deterministic solutions and to simulate the time course of the various model species. The default options for LSODA were selected for the time course simulation.

A global quantity was included during the construction of the mammalian cell cycle and metabolism model to measure the time it takes for the cell to complete one cycle of the division process. This global quantity takes the time point at which the cell mass is halved and the previous time point at which cell mass halving also occurs in order to calculate the time difference that defines a single cell cycle period. The first cell cycle period is measured only after 100 hours (h) have passed during the time course in order to allow sufficient time for the cell mass to oscillate and settle to their stable stationary period.

3.3 Results And Discussion

In order to investigate whether the response of the mammalian cell cycle and metabolism model to varying concentrations of glucoses provided a reasonable reflection of real-life behaviour for normal cells, separate simulations were carried out to observe how the cell mass evolves over time

with respect to a certain initial glucose concentration. The time course results for the cell mass when 1mM and 10mM of glucose are applied to the model is displayed in Figure 3.3.

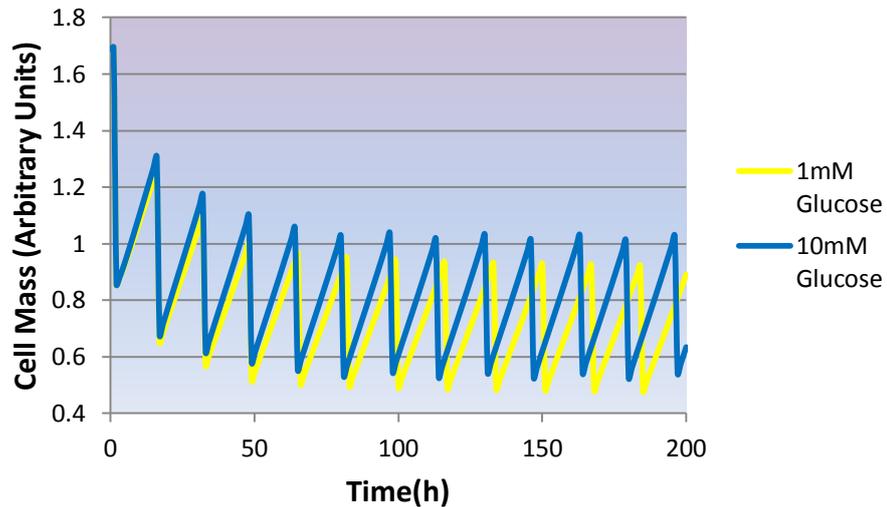


Figure 3.3. Time course plots for cell mass when the initial glucose concentration is set to 1mM (yellow line) and 10mM (blue line)

In Figure 3.3, the time course plot for when 1mM of glucose is initially applied shows the overall concentration of the cell mass decreasing until it eventually settles to a maximum of approximately 0.95 arbitrary units (au) and a minimum of 0.5au. When the initial concentration of glucose is set to 10mM, the overall cell mass also decreases until it settles to a slightly higher maximum of 1.05au and a minimum of 0.55au. This biologically reflects the fact that although normal cells may differ in the times they take to complete one full cycle of division, they should be maintaining a similar cell size per cycle, although the level of glucose present may affect the cell size to some degree by increasing the cell size when more glucose is available. In

eukaryotic cells, these critical cell size thresholds are imposed during the G1-S and G2-M transitions (Jorgensen and Tysers, 2004).

Similar time course plots, displayed in Figure 3.4, were produced for the non-essential amino acids (AA) when 1mM and 10mM of glucose are the initially enforced on the model.

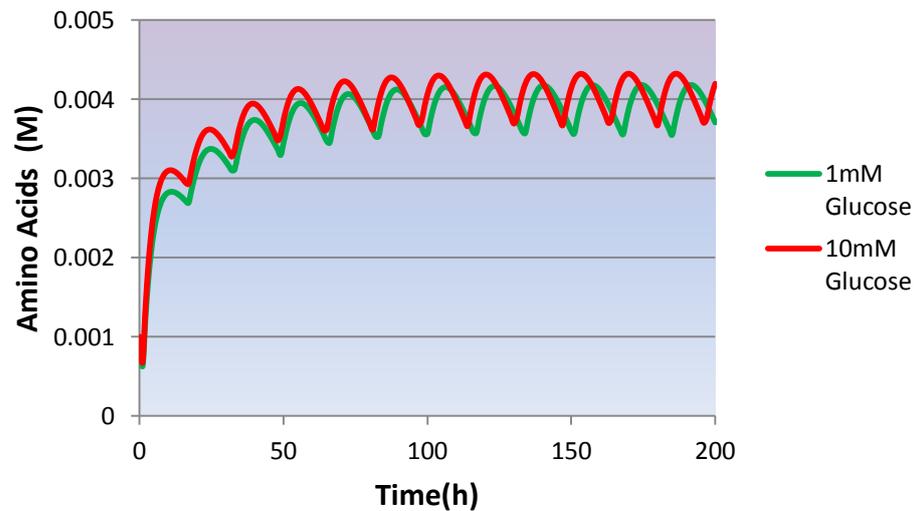


Figure 3.4. Time course plots for non-essential amino acids when the initial glucose concentration is set to 1mM (green line), and 10mM (red line)

As shown in Figure 3.4, the time course plot for when the initial concentration of glucose applied is 1mM shows the concentration of non-essential amino acids increasing before eventually settling to a stable cycle with values between 0.0035M and 0.0042M. This cyclical behaviour is clearly driven by the events of the cell division process. The time course for when 10mM of glucose is initially applied exhibits a similar trend but settles to a stable cycle with very slightly higher values of between 0.0036M to

0.0043M. This suggests that a higher initial concentration of glucose will slightly increase the overall concentration of non-essential amino acids.

Other time course plots were produced for the model species found towards the end of the metabolic pathway, such as pyruvate, lactate, PEP and P2G, before incorporation into cell mass. These plots are shown in Figures 3.5 to 3.8.

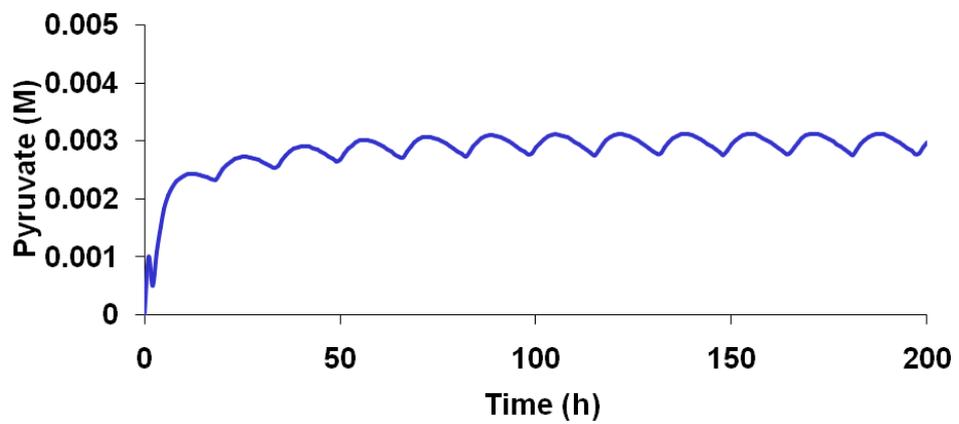


Figure 3.5. Time course plot for pyruvate when the initial glucose concentration is set to 10mM.

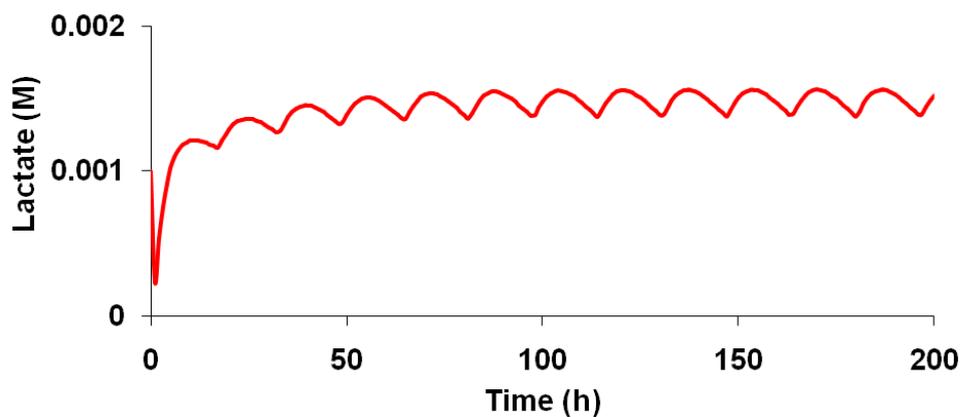


Figure 3.6. Time course plot for lactate when the initial glucose concentration is set to 10mM.

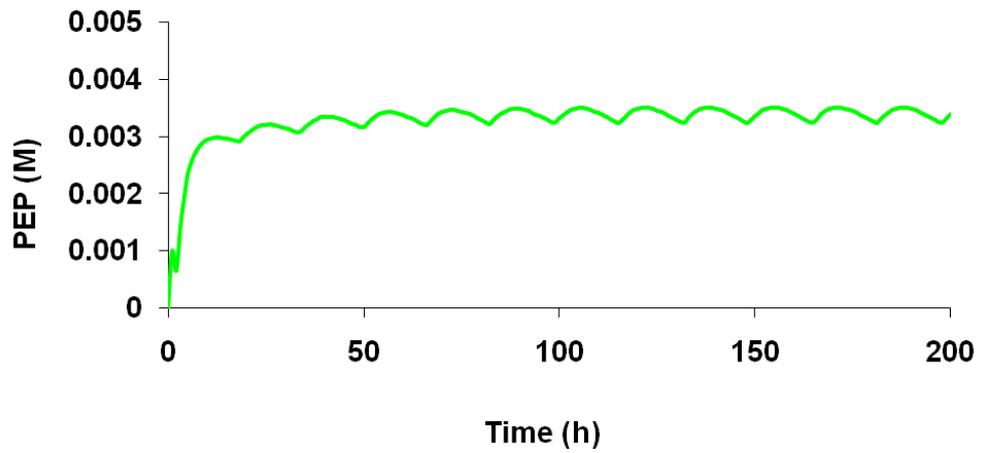


Figure 3.7. Time course plot for PEP when the initial glucose concentration is set to 10mM.

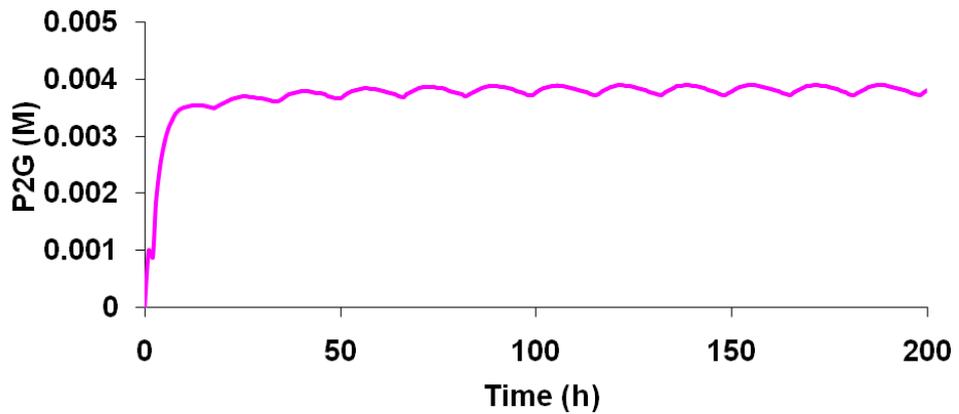


Figure 3.8. Time course plot for P2G when the initial glucose concentration is set to 10mM.

Figures 3.5 to 3.8 all show the model species of interest initially increasing before oscillating between two certain concentrations, although it appears that as a backtrack is made through the glycolysis pathway, the amplitude of oscillations becomes smaller in size. For instance, P2G comes before PEP in

the metabolic pathway and the eventual amplitude of the species concentration oscillations is higher for PEP than for P2G.

The parameter scan function was performed in COPASI on the maximal metabolic rate constant involved in the reaction step describing the conversion from Glucose to G6P in response to different initial concentrations of glucose. The resulting plot obtained from the parameter scan is shown in Figure 3.9.

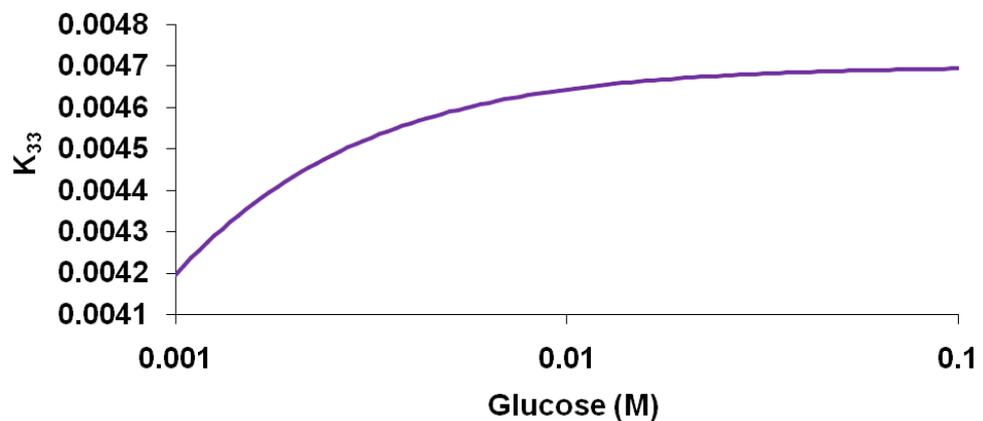


Figure 3.9. Maximal metabolic rate constant for reaction 33 vs. Glucose concentration

As displayed in Figure 3.9, the value of the maximal metabolic rate constant for reaction 33 increases to a steady state asymptotically as the concentration of glucose is increased. This type of plateauing behaviour is to be expected if equation (3.2) is viewed from an analytical perspective. As the concentration of glucose becomes significantly larger than KM_{GLU} , the ratio between the glucose concentration, and the sum of KM_{GLU} and glucose concentration becomes 1 to give a new rate law expression of $V_{33} = K_{33}$.

A plot of the first cell cycle period taken after 100 hours of division time for normal cells is shown for different concentrations of glucose in Figure 3.10.

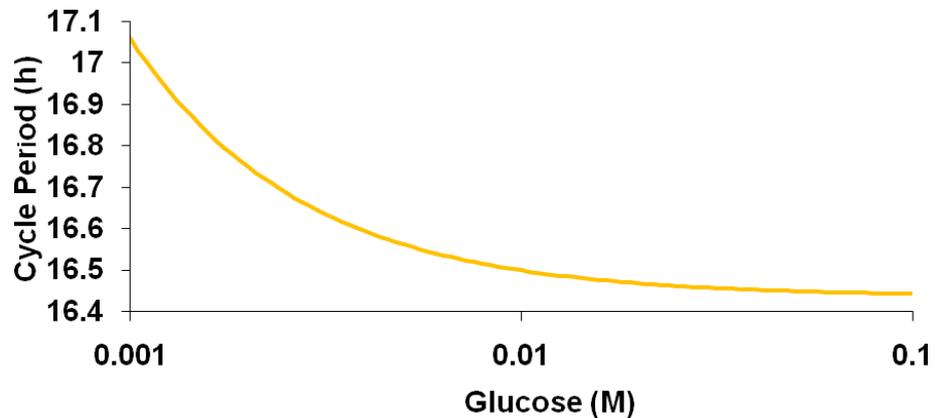


Figure 3.10. Cell Cycle Period (after 100 hours) as a function of Glucose concentration for normal cells.

Figure 3.10 shows that as the concentration of glucose increases, the time taken to complete one cell cycle gradually decreases asymptotically to 16.5 hours. This suggests that as the concentration of glucose becomes larger, the cells begin to divide at a slightly faster rate *i.e.* display shorter cell cycle period times up to a certain threshold of glucose concentration.

At present, a novel mathematical model that incorporates aspects of both the mammalian cell cycle and metabolic pathway has been developed here. Time course simulation results demonstrate the model is able to respond to different concentrations of glucose in a way that is compatible with real-life observations (Higuchi *et al.*, 1997).

A MODEL OF CANCER GROWTH AND DIVISION

4.1 Mammalian Cell Cycle And Metabolism Model For Cancer Cells

This chapter will discuss how gene expression data was utilised to create a version of the mammalian cell cycle and metabolism model, as described in chapter 3, but now focusing on cancer cells.

Gaglio *et al.* Discuss The Upregulation Of Cancer Gene Expression:

Gaglio *et al.* carried out transcriptional profiling experiments to analyse metabolic alterations that are dependent on the oncogene K-Ras in order to gain insights into how cancer cell metabolism is regulated (Gaglio *et al.*, 2011). The flux of ^{13}C -labelled glucose and glutamine was measured for downstream metabolites in normal and cancer cells obtained from the same

cell lines. The generated microarray transcription profiles for the normal and cancer cells were then compared with each other.

Gaglio *et al.* found that several genes associated with glycolysis, glutamine metabolism and nucleotide biosynthesis upon transformation with K-Ras oncogenes demonstrated elevated expression levels compared to the case for normal cells i.e. cancer cells have demonstrated higher gene expression levels compared to normal cells.

Microarray Data Collection:

The collection of microarray data by Gaglio *et al.* based on human raw expression data for U133A arrays was extracted from the publicly accessible biological databases, GEO (<http://www.ncbi.nlm.nih.gov/geo>) and CellMiner (<http://discover.nci.nih.gov/cellminer/home.do>).

The Gene Expression Omnibus (GEO) is a database containing information related to functional genomics. A tool for queries and browsing is available to enable researchers to find and download specific gene expression profiles (Barrett *et al.*, 2011; Edgar *et al.*, 2002).

CellMiner is a online application created by the Genomics and Bioinformatics Group that allows researchers to obtain and analyse molecular and pharmacological datasets for NC1-60 lines, which consist of a panel of 60 diverse human cancer cell lines that have been heavily utilised by the Developmental Therapeutics Program of the U.S. National Cancer Institute

since the 1990s to screen chemical compounds (Liu *et al.*, 2010; Reinhold *et al.*, 2012; Shankavarum *et al.*, 2009). The web application includes genomic data and drug data query tools and allows cell line metadata to be downloaded as well.

Gaglio *et al.* obtained their transcription profiles data of the human normal breast tissue from GEO, based on G2M44683 tissue, and breast cancer cell line MDA-MB-231 from CellMiner. The transcriptional data profiles for the normal and cancer cells, encoded in Affymetrix™ CEL files, were imported into the GeneSpring software platform (Agilent Technologies Inc.), where they were then normalised and summarised as probe-level measurements using the robust multi-array average (RMA) method (Gaglio *et al.*, 2004). The RMA methodology involves taking microarray data and subjecting the raw intensity values contained within them to background corrections, log₂ transformations and then quantile normalisation. The normalised data is then fitted against a linear model to generate an expression measure for each probe set on each array (Bolstad *et al.*, 2003; Irizarry *et al.*, 2003).

Gene Expression Ratios:

It is proposed for this research project that this upregulation effect found in cancer gene expression may be modelled by identifying the ratio between the cancer and normal gene expression levels. For instance, if the

maximum metabolic rate constant found in the mammalian cell cycle and metabolism model is originally 1 and the level of gene expression is detected to be twice as high in cancer cells, then that particular rate constant can be given the value of 2 in the new model for the mammalian cell cycle and metabolism model for cancer cells. The main advantage of calculating the ratio value between the cancer and normal gene expression levels is that it is dimension-free. This allows for different gene expression datasets to be analysed but still share a standard measure for exactly how much the gene expression level has been upregulated from normal to cancer cells.

The gene expression ratios were calculated for genes involved in the pathways of glycolysis and non-essential amino acids. The information provided by the ratios were then entered into the cell cycle and metabolism model for normal cells by replacing the relevant parameter values with the multiple of the ratio and the original parameter value to form a version of the model for cancer cells.

4.2 Model Implementation

During this project, gene expression data recorded in the paper by Gaglio *et al.* was found to be readily available for the glycolysis pathway. The genes encoding enzymes that play a role in the model reactions involved with the glycolysis pathway and their corresponding expression intensities for both normal breast tissue (Normal) and MDA-MB-231 cancer cells (Cancer) were recorded as shown in Table 4.1.

Table 4.1. Table showing the genes involved in the glycolysis pathway and their corresponding gene expression intensity values and expression ratio values given by Cancer (C) divided by Normal (N).

Gene Title	Gene Symbol	Normal Breast (N)	MDA-MB-231 (C)	C/N	C/N Mean Value
		Expression Intensity			
hexokinase 1	HK1	0.78	1.20	1.53	1.39
hexokinase 2	HK2	0.76	0.95	1.24	
<i>glucose phosphate isomerase 1 (data obtained from Mouse cell line as missing from Human Dataset)</i>	<i>Gpi1</i>	<i>0.77</i>	<i>1.42</i>	<i>1.85</i>	<i>1.85</i>
phosphofructokinase, liver	PFKL	0.96	1.00	1.03	3.41
phosphofructokinase, muscle	PFKM	0.63	0.62	0.99	
phosphofructokinase, platelet	PFKP	0.16	1.29	8.21	
aldolase A, fructose-bisphosphate	ALDOA	0.42	0.69	1.66	1.66
triosephosphate isomerase 1	TPI1	0.20	0.76	3.84	3.84
glyceraldehyde-3-phosphate dehydrogenase	Gadph	0.31	0.86	2.79	2.79
phosphoglycerate kinase 1	PGK1	0.19	1.35	7.02	7.02
2,3-bisphosphoglycerate mutase	BPGM	0.87	1.57	1.81	1.81
enolase 1, (alpha)	ENO1	0.23	1.31	5.58	4.35
enolase 2 (gamma,	ENO2	0.29	1.74	6.02	
enolase 3 (beta, muscle)	ENO3	0.68	0.99	1.44	
pyruvate kinase, muscle	PKM2	0.11	1.00	9.26	9.26
lactate dehydrogenase A	LDHA	0.25	1.32	5.18	3.71
lactate dehydrogenase B	LDHB	0.30	0.68	2.23	

In Table 4.1, the ratio between the value for Cancer and Normal is first calculated for each individual gene expressing an enzyme of the

glycolysis pathway. In the event where there is more than one possible gene involved for each reaction step modelled in the pathway for glycolysis, the mean value of the expression ratios is also calculated. The arithmetic mean is taken here for an unbiased approximation of the ratios involved. For example, the reaction step describing the conversion from glucose to G6P involves the genes, hexokinase 1 and hexokinase 2, with corresponding expression ratios of 1.53 and 1.24 respectively. The mean value of these two ratios is therefore 1.39 and this is the value that will determine the parameter value entered into the mammalian cell cycle and metabolism for cancer cells.

Gene expression data for the non-essential amino acids pathway was found to be lacking from the paper by Gaglio *et al.* and its supplementary information tables. Therefore, data about the genes involved in expressing the necessary enzymes required for synthesising non-essential amino acids had to be extracted through an alternative method.

The first step comprises of accessing the KEGG database (Kanehisa and Goto, 2000; Kanehisa *et al.*, 2012) to identify all the genes involved in expressing the enzymes that catalyse the reactions of the non-essential amino acids pathways.

The names of these particular genes were then checked against GeneAnnot (<http://genecards.weizmann.ac.il/geneannot/index.shtml>), which is a public database containing annotations of microarray probe-sets from human HG-U133. If these genes were found to be exclusively in the Human Genome U133A microarrays, then their probe-set IDs were recorded.

CEL files containing gene expression intensity values for human normal breast tissue (NCBI GEO accession GSM44683) and human breast

cancer cell line MDA-MB-231 (CellMiner accession 13409hg133a21) were downloaded from the online databases, GEO and CellMiner respectively.

The next step involved loading BioConductor software in a session of the statistical programming language R (which can be downloaded from <http://www.r-project.org>). Bioconductor (www.bioconductor.org) is an open source and open development software project that is based on R and enables high-throughput genomic data to be analysed by the user (Gentleman *et al.*, 2004). In a session of R, the appropriate BioConductor software was used to extract the robust multi-array (or multichip) average (RMA) of the gene expression intensity values from the CEL files.

The cancer to normal gene expression ratios related to non-essential amino acids could then be calculated. The mean expression ratio values for the non-essential amino acids, originally synthesised from pyruvate and P3G respectively, were used to update the relevant parameters in the model by multiplying the ratio value with the original parameter value. This results in a new version of the mammalian cell cycle and metabolism model for cancer cells. The methodology involved in this process is summarised in Figure 4.1. A schematic diagram displaying the mean values of the cancer to normal gene expression level ratios involved in the model reactions related to the glycolysis and non-essential amino acid pathways is shown in Figure 4.2.

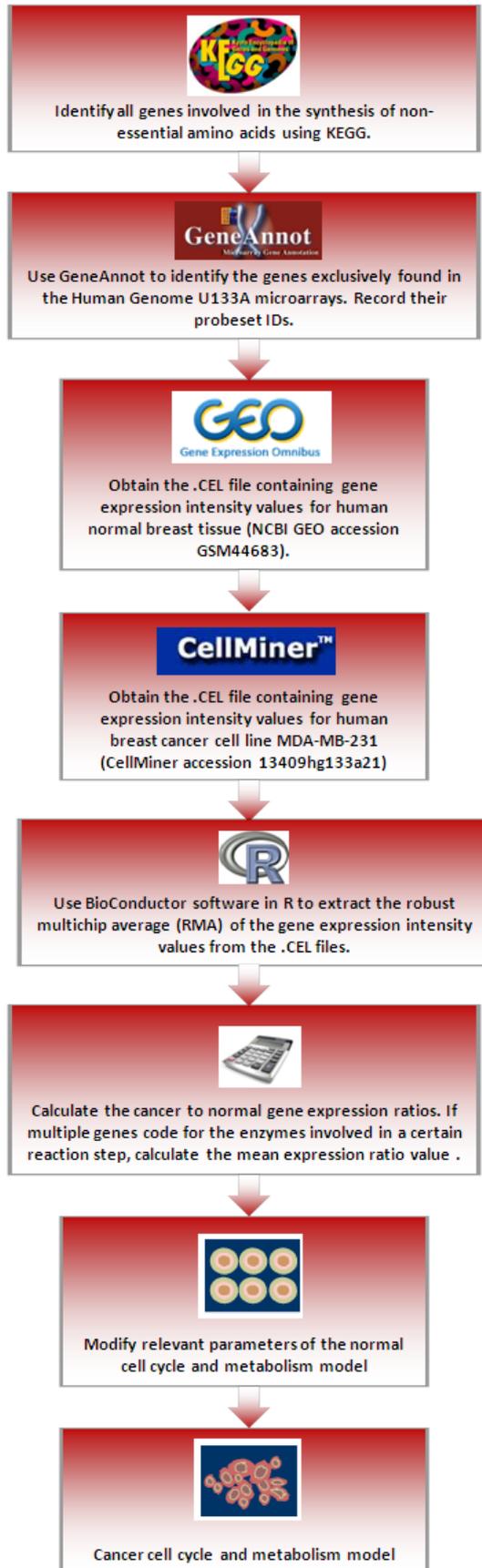


Figure 4.1. Methodology for creating the cancer cell cycle and metabolism model based on gene expression data.

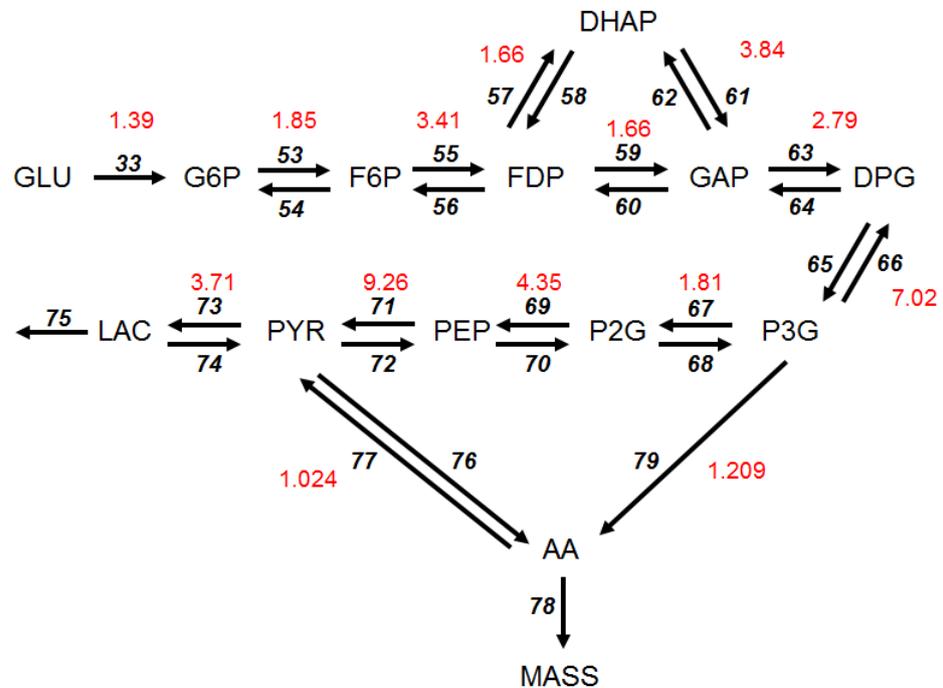


Figure 4.2. Schematic diagram showing the mean values of the cancer to normal gene expression level ratios (highlighted in red) involved in the model reactions (reaction number in black) found in the glycolysis and non-essential amino acid pathways.

Gene expression data involved with the non-essential amino acids are given in Table C.1, which can be found in Appendix C. The parameter values for the cancer version of the mammalian cell cycle and metabolism can also be found in Appendix C.

Sensitivity Analysis:

The generalised sensitivity values of a mathematical model can be calculated using COPASI, which uses finite differences to perform numerical differentiation. The differentiation is carried out based on a list of functions with respect to one or two lists of variables specified by the modeler, usually leading to the generation of a two-dimensional matrix. The result becomes a three-dimensional array of second derivatives when all the first derivatives are differentiated again with respect to the all the variables provided in an additional list of variables for second derivatives. COPASI allows the option of viewing both unscaled and scaled result matrices for the sensitivities obtained. The scaled values are obtained from the unscaled values by scaling them with the relevant steady state concentrations.

COPASI offers two options for calculating sensitivity values.

- **Delta Factor:** This value instructs COPASI on what it should set the delta value to be for the finite difference numerical differentiation calculations. The delta is obtained by multiplying the value of the delta factor with the current absolute value of the variable of interest. The default value for the delta factor is set to 1×10^{-6} .
- **Delta Minimum:** This value is set such that if the resulting value of the delta calculated is smaller than the value of delta minimum, then the delta minimum value is used instead of the delta value. The default value for the delta factor is set to 1×10^{-12} .

The default values for the two options of sensitivity analysis in COPASI were used for the purposes of analysing the mammalian cell cycle and metabolism model.

4.3 Results And Conclusions

Sensitivity analysis was first performed in COPASI on the maximum metabolic rate constants with regards to the cell cycle period (after the cells have been allowed to grow and divide for at least 100 hours) for normal cells. The scaled values that were generated were recorded. This process was then repeated for the case where the cancer version of the mammalian cell cycle and metabolism model is analysed in order to allow comparisons between the normal and cancer cells to be deduced. The results of the sensitivity analysis are shown in Table 4.2.

Table 4.2. A comparison of scaled sensitivity values generated from performing sensitivity analysis on the metabolic rate constants with regards to the cell cycle period for normal and cancer cells (including the percentage change from normal to cancer values)

Parameter Name	Sensitivity (Normal)	Sensitivity (Cancer)	Percentage Change (%)
K_{33}	-0.197876	-0.522774	-164.193
K_{53}	0.0117697	-0.178098	-1613.191
K_{54}	-0.0405423	-0.0705738	-74.074
K_{55}	0.129972	-0.0515921	-139.695
K_{56}	0.00363284	-0.151819	-4279.072
K_{57}	0.0983428	0.0136298	-86.141
K_{58}	0.0276506	-0.173826	-728.652
K_{59}	0.0449509	0.060646	34.916
K_{60}	0.0791869	-0.215546	-372.199
K_{61}	0.0700632	-0.00464293	-106.627
K_{62}	0.0885207	-0.0682989	-177.156
K_{63}	-0.107524	0.0454292	142.250
K_{64}	0.0250515	-0.228922	-1013.806
K_{65}	-0.118299	-0.0451652	61.821
K_{66}	0.159162	-0.111336	-169.951
K_{67}	0.247337	-0.154575	-162.496
K_{68}	-0.0112034	-0.188435	-1581.945
K_{69}	0.0825949	-0.0623434	-175.481
K_{70}	0.220275	-0.157342	-171.430
K_{71}	0.129999	0.0334539	-74.266
K_{72}	0.00570696	-0.0825292	-1546.115
K_{73}	-0.057377	-0.0496935	13.391
K_{74}	0.00170704	-0.165791	-9812.192
K_{75}	0.0355715	-0.204981	-676.251
K_{76}	-0.0414956	-0.158925	-282.992
K_{77}	0.0496609	0.193859	290.365
K_{78}	-0.00234197	-0.238771	-10095.306
K_{79}	0.0752917	-0.0857682	-213.915

In Table 4.2, a sensitivity value highlighted in blue denotes a positive sensitivity value. This implies that if the corresponding parameter value is increased, the cell cycle period will also increase accordingly. Conversely, a value highlighted in red means that it is a negative sensitivity value and should the corresponding parameter increase in value, the time taken to complete a full cycle of division will decrease.

Theoretically, the larger the magnitude of the sensitivity value is, the greater the influence the corresponding parameter value should hold over the cycle period when a perturbation is applied to that specific parameter.

As shown in Table 4.2, if you take the absolute value of K_{33} , it then demonstrates the largest sensitivity value in the case for cancer cells. This particular parameter represents the maximum metabolic rate constant involved in the conversion of glucose to G6P. If the original sensitivity value of K_{33} is taken, then it is observed to be a negative value in both the categories for normal and cancer cells. This suggests that if this parameter value is increased in value, the cell cycle period should decrease in magnitude. A consequence of this result is that this parameter and its associated reaction step may make for a potentially effective drug target since an aim of the drug compound would be to halt or slow down the cancerous cells from dividing as quickly as they are presently doing so.

Table 4.2 also shows that K_{78} exhibits the greatest percentage change from normal to cancer sensitivity values. This specific parameter describes the maximum metabolic rate constant involved in the reactions that sees the

non-essential amino acids incorporated in the cell mass. Again, this may be a potential drug target site for at least restoring the period of cell division for cancer cells to that found in normal cells.

The results of the sensitivity analysis shown in Table 4.2 can be used to guide the choice of drug combinations in cancer therapy as discussed in chapter 5.

The time course of cell mass for both normal and cancer cells were simulated in COPASI and a single plot comparing the two time courses are shown in Figure 4.3.

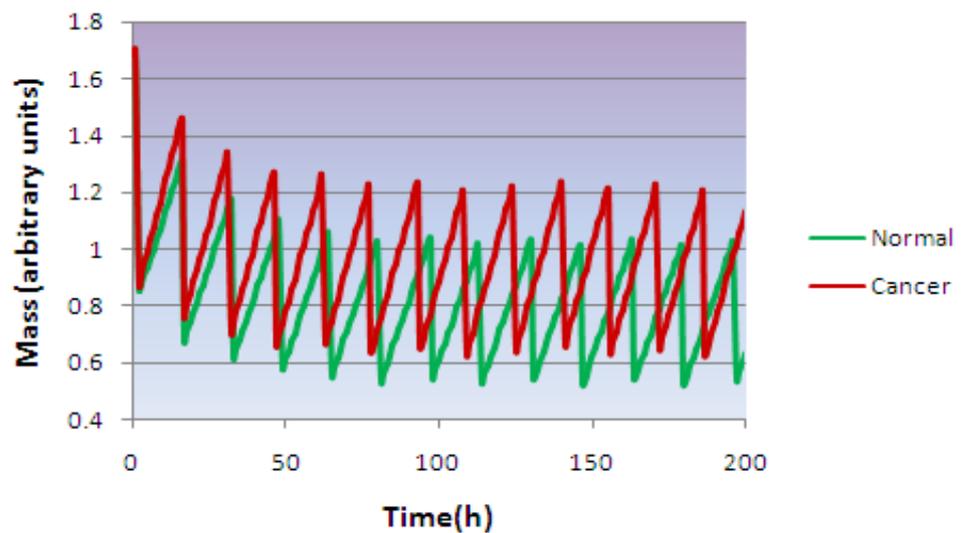


Figure 4.3. Comparison of cell mass time courses for normal and cancer cells

The time courses displayed in Figure 4.3 illustrates the fact that time taken for a normal cell to complete one full division cycle is greater than that of the cancer cell. In other words, the cancer cells are dividing at a faster rate than

normal cells. This reduction in cycle period time for cancer cells is due to a faster accumulation of biomass and accelerated level of glucose metabolism.

The cycle periods for different concentrations of glucose were computed in COPASI, using the parameter scan function, for both normal and cancer cells. A comparison between the two cells types are is shown in Figure 4.4.

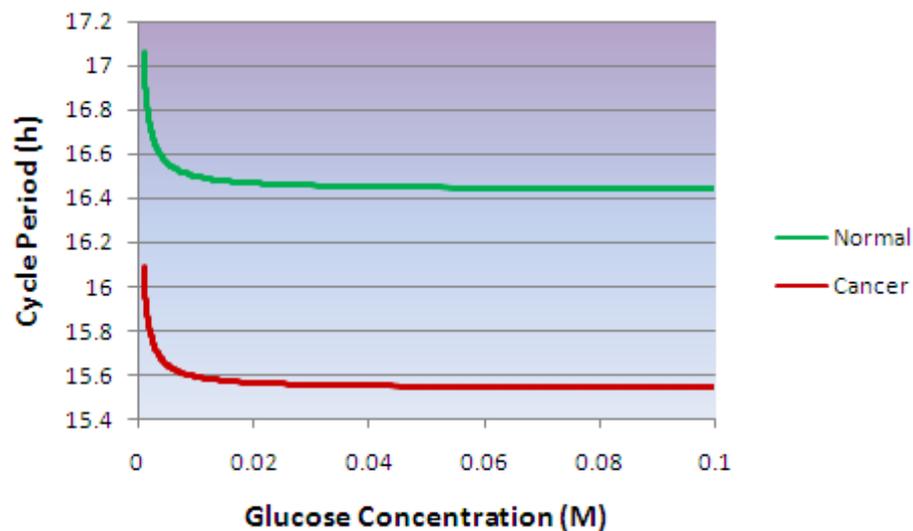


Figure 4.4. Comparison of cycle period vs. glucose concentration for normal and cancer cells.

Figure 4.4 shows that as the concentration of glucose is increased, the cycle period behaviour for both normal and cancer cells demonstrate a similar trend by showing the period decreasing gradually until it reaches an asymptotic state. This means that the cells will divide at a faster rate when the glucose concentration levels are higher, but only up to a certain threshold concentration of glucose. As implied in Figure 4.4, the cycle period for normal cells is always longer than for the cancer cell situation. In this

simulation, the cycle periods for normal and cancer cells settle to around 16.45 hours and 15.55 hours respectively in response to increasing concentrations of glucose.

DRUG COMBINATION EFFICACY IN CANCER THERAPY

5.1 Introduction

During the development of a drug compound to treat cancer cells, pharmacologists should be asking themselves certain questions as follows.

- Does the drug have high efficacy?
- Is the route of administration of the drug compound as non-invasive as possible to cancer patients?
- Does the drug show low toxicity and incur a minimal amount of side-effects?
- How specific is the drug to treating a particular type of cancer?
- Can the drug compound be manufactured at a reasonable price?

For a long time in the history of drug development, drug testing tried to rely on the ability of a cytotoxic compound to kill or slow the progression of tumour cells more effectively than the normal cells. This usually results in significant side-effects for cancer patients who take the drug, leading to the mindset that sometimes the cure may be far more detrimental than the actual disease itself.

Recently, the precision and effectiveness of novel drugs has been improved through the use of targeted monotherapies (Zahorowska *et al.*, 2012). These drugs work by focusing on antagonising, blocking, inhibiting or activating a single endogenous protein or enzyme, while attempting to exhibit minimal side-effects. However, as mentioned cancer is a complex disease that is driven by multiple genomic events of which the network dynamics involved are highly non-linear, and often targeted monotherapies may not be offering a full and realistic picture of what is actually happening with the network dynamics.

The pharmaceutical industry has now recognised the importance of pursuing a combinatorial approach to molecular targets in order to identify potential target combinations for testing as well as prioritising combinations of existing and experimental therapeutics. For instance, studies have demonstrated the benefits of combination chemotherapy over monotherapy in the first-line treatment of metastatic colorectal cancer (Arnold, 2008).

Another example is the proven effectiveness of pretreating pancreatic cells with apigenin in order to sensitise the cancer cells to gemcitabine-induced killing (Lee *et al.*, 2008).

Systems biology can be utilised to explore combination therapy by coalescing mathematical and biological techniques to discover emergent properties of the complex system involved behind cancer. In particular, mathematical models may help in determining optimum drug regimes for inducing maximum toxicity in cancer cells while minimising the damage done to normal cells.

5.2 Methods

In order to explore this combinatorial approach to cancer therapy, the model developed for the cell growth and division of cancer cells is subjected to optimisation algorithms in COPASI to search for combinations of maximum metabolic rate parameters that generate the highest cell cycle periods for the defined model. Here, the values of the parameters are used to represent the effect of drugs. Lower value of maximal rates mean that a drug should be used to inhibit that particular enzyme.

Many algorithms have been included in COPASI for minimising or maximising an objective function. Here, the objective of this study is to increase the time taken to complete one division cycle. The maximum objective function is desired in this situation so that the period time is as high as possible.

Examples of optimisation algorithms found in COPASI are shown in Table 5.1 and whether they are able to locate the local or global maxima in the system being investigated.

Table 5.1. Examples of algorithms implemented in COPASI to perform optimisation

Local	Global
Gauss-Newton	Evolutionary Algorithm
Steepest Descent	Genetic Algorithm
Nelder-Mead	Simulated Annealing
Hooke-Jeeves	Particle Swarm
Levenberg-Marquardt	

The first of these local optimisation algorithms shown in Table 5.1 is the Gauss-Newton algorithm, which can be used to minimise a sum of squared function values through the process of iteration. The main advantage of this method is that it does not require second derivatives that may sometimes prove a challenge to compute.

The Steepest Descent method is another method for finding a local minimum (Fogel *et al.*, 1992). It does this by following the direction of the steepest descent on the hyper-surface of the objective function. The direction taken is defined by the negative of the gradient of the objective function.

The Nelder-Mead method (Nelder and Mead, 1965), also referred to as the downhill simplex method, begins by forming an initial simplex, which is a polytope of $N+1$ vertices in N -dimensions, from $N+1$ test points. The function value is then evaluated at the test points and the worst of these test

points are replaced by a point determined by reflecting the worst point through the centroid of the remaining N points. If this new point is better than the best current point, the algorithm will try to stretch out exponentially along this line. If not, then the algorithm will shrink the polytope towards the best point.

The Levenberg-Marquardt algorithm is also a gradient descent method for finding a local optimum (Levenberg, 1944; Marquardt, 1963). It is able to effectively switch between the method of steepest descent to the Newton optimisation method. In addition, the second derivatives do not require calculation as they can be estimated from the gradient of the residuals.

A method for ascertaining a global optimum is the Evolutionary Algorithm, which mimics the process of evolution (Fogel *et al.*, 1992). It consists of individuals, each representing a potential solution of the optimisation being performed, which reproduce and compete. At the end of each generation, the algorithm produces double the number of individuals. An individual is allocated wins if its competitors have a worse solution than itself. The individuals are ranked according to their number of wins, and the population size is reduced to the original number of individuals by removing those that have the worst solutions.

The genetic algorithm is similar to the evolutionary algorithm except each individual is represented as a gene (Bäck and Schwefel, 1997). After each generation of the algorithm, an individual is paired up with another at random. Two new individuals are then produced from a combination of the

genomes of their parents. The population size is reduced to the original number by eliminating the individuals that have performed the worst in accordance with the number of wins they have.

One other method for finding a global optimum is Simulated Annealing (Kirkpatrick *et al.*, 1983). It mimics the analogy of metals being slowly cooled to make them reach a state of low energy where they are very strong. The objective function is considered a measure of the energy of the system. Simulated annealing picks a random variable and value during each step. If the assignment of that particular value to the variable does not increase the number of conflicts, then that assignment is accepted. One benefit of this algorithm is that it is guaranteed to converge for an infinite number of iterations. However, a drawback of this is that the algorithm could potentially run for a very long time and be computationally intensive.

The choice of optimisation algorithm would depend on its ability to maximise the objective function and its speed of computation. First of all, in order to determine exactly which optimisation algorithms in COPASI are the most effective for determining the maximum time period, the collection of all the metabolic rate constants involved in the cancer model of cell growth and division underwent different optimisation techniques. A plot comparing the best cycle period values obtained for the various COPASI optimisation algorithms is shown in Figure 5.1.

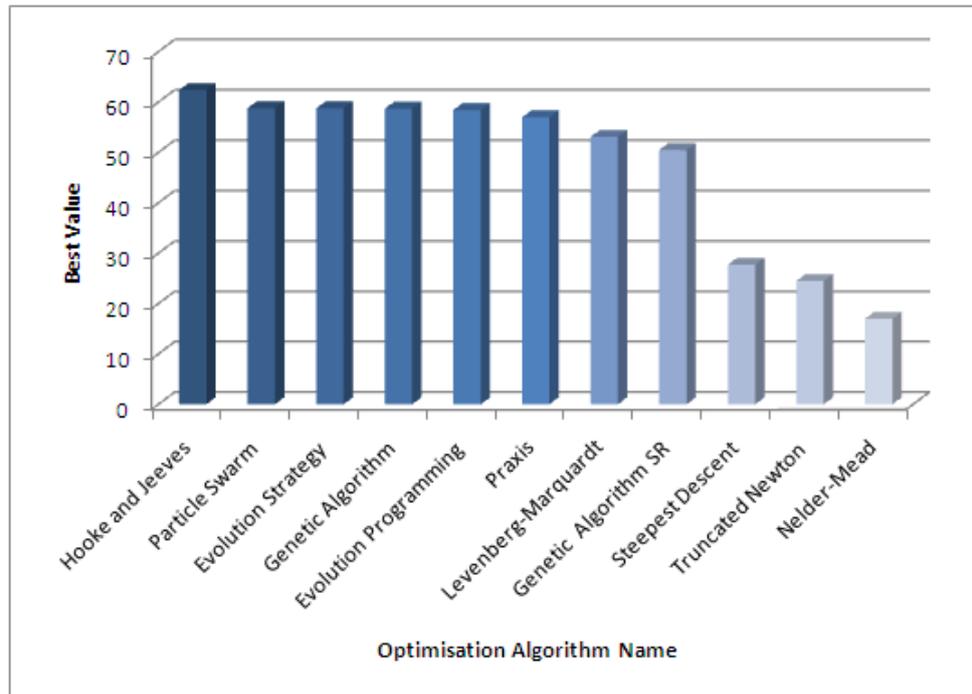


Figure 5.1. Comparison of the best value obtained for the cycle period when all the maximum metabolic rate constants of the cancer model are subjected to different optimisation algorithms in COPASI

Figure 5.1 demonstrates that the Hooke and Jeeves, and particle swarm algorithms were the most effective at generating the highest cycle periods, whereas other algorithms such as Nelder-Mead had not performed very well with regards to model optimisation.

Hooke and Jeeves, and particle swarm were therefore chosen as the optimisation algorithms to be used for computing the maximum cycle period value for different combinations of metabolic rate constants to account for one local optimisation method and one global optimisation method.

Hooke And Jeeves:

The algorithm developed by Hooke and Jeeves uses a direct search method of coordinate patterns (Hooke and Jeeves, 1961). The Hooke and Jeeves method involves two major stages called the exploratory search, also referred to as a coordinate search, and the pattern search. The exploratory search phase seeks out an improvement in the directions parallel to the coordinates axes, while the pattern search aims to accelerate the search process by moving to a new and improved position in the direction of the point obtained by the exploratory search. Hooke and Jeeves is able to search for the minimum or maximum of a non-linear function without needing information about the derivatives of the function.

The following options are included for Hooke and Jeeves in COPASI.

- **Iteration Limit:** This value must be an integer that is positive. It defines the maximum number of iterations the method should go through. The default value is set to 50.
- **Tolerance:** This value must be positive and defines the level of tolerance the solution obtained should fall within. The default value is set to 10^{-5} .
- **Rho:** This entry is a positive value between 0 and 1. It defines exactly how much the step size should be decreased between iterations. The default value is set to 0.2.

Particle Swarm:

The particle swarm optimisation algorithm was developed by Eberhart and Kennedy (Eberhart and Kennedy, 2005) and originated as a real life example of a simplified social system. It was initially intended to graphically simulate the choreography of a flock of birds, but it was found that this method could also be used as a solver. Particle swarm works by first initializing the system with a population of random solutions and searches for optima by updating generations without any crossovers. The particles fly through the problem space in a random direction and velocity, but also biased by the direction of other near-by particles. This provides a population-based search procedure for the optimisation process.

The following options are included for Particle Swarm in COPASI.

- **Iteration Limit:** The default value is set to 2000.
- **Swarm Size:** This value should be a positive integer that determines the number of particles in the swarm. The default value is set to 500.
- **Standard Deviation:** This entry provides another criteria for terminating the optimisation process. When the standard deviation of the values of the objective function for each particle becomes less than the specified value, the algorithm is terminated. The default value is 10^{-6} .

During the cycle period optimisation process, the default option values were used for both Hooke and Jeeves, and particle swarm. Some of the default values such as swarm size may be decreased to enable faster computations, but the higher these certain values are, the better the representation the optimisation result has over the actual model dynamics.

Condor High-Throughput Computing Platform:

For both Hooke and Jeeves, and particle swarm, the start value of the metabolic control rate constants was randomised to fall between the range of $\{-50\%, +50\% \} * \text{Original Parameter Value}$ for a number of optimisation runs, before the best value obtained for the cycle period was recorded. This was to observe whether changing the start values of the parameters had any effect over the eventual highest cycle period value obtained.

20 optimisation runs using both algorithms were performed for every set of metabolic control rate constant combinations being investigated. The highest value of the cycle period obtained out of the 20 optimisation runs is defined as the best value for the cell period.

In order to ease computational simulation speeds, especially as particle swarm can sometimes take up to days to complete a certain optimisation task, the Condor high-throughput computing platform was utilised (<http://research.cs.wisc.edu/condor/>).

Condor is an open source high-throughput computing software framework that enables users to submit computationally intensive jobs, which

are then delegated to and carried out by simultaneous processors. The main advantage of Condor is its efficiency in employing the use of idle computers. This allows potentially wasted processing power to be harvested. As the number of machines linked to the Condor grid system grows, the greater the computational resources will become available to individual researchers.

However, the use of Condor would require programming scripts, creating job specification files and altering COPASI files. All of these tasks are not hard, but would involve a tedious and time-consuming process of implementation. An alternative is the package Condor-COPASI (Kent *et al.*, 2012) which automates all of these tasks through a simple web user-interface. The Condor-COPASI package running in the Mendes research group accesses a large pool of *circa* 2500 CPUs in the Faculty of Engineering and Physical Sciences of the University of Manchester. This is the system that was used to run the simulations and optimizations described in this chapter.

5.3 Results And Discussion

The types of optimisation performed on certain sets of metabolic rate constants are shown in Table 5.2.

Table 5.2. Details about the type of optimisation performed on different combinations of metabolic rate constants and their corresponding best cycle period values obtained through Hooke and Jeeves, and particle swarm

Type Of Optimisation Performed	Cycle Period Best Value (Hooke and Jeeves)	Cycle Period Best Value (Particle Swarm)
The complete set of metabolic rate constants	62.3839	58.8396
{ K_{AA2} , K_{F6P2} , K_{G6P1} , K_{PEP1} , K_{33} }	25.3868	25.3868
Metabolic rate constants with the top 3 greatest (absolute value) sensitivities	23.7425	23.7425
K_{33}	19.9447	19.9447
Metabolic rate constants with the top 3 greatest sensitivity (absolute value) percentage change from the normal to cancer cell model sensitivities	18.828	18.828
{ K_{AA1} , K_{F6P1} , K_{PYR2} }	17.3064	17.3064
Metabolic rate constants with a positive sensitivity value	16.5742	16.5742
Metabolic rate constants involved in PEP ↔ PYR and DPG ↔ P3G (which had the two highest cancer to normal gene expression ratios in the metabolic chain)	16.0497	16.0497
Metabolic rate constants involved in PEP ↔ PYR (which had the highest cancer to normal gene expression ratio in the metabolic chain)	16.0474	16.0474

As shown in Table 5.2, nine different types of optimisation were performed on the metabolic rate constants.

The first of these involve the complete set of metabolic rate constants from the model of cell growth and division. In this scenario, the best cycle period value out of 20 optimisation runs with randomised parameter start values obtained with the Hooke and Jeeves algorithm was 62.3839 hours. This is slightly higher than the value obtained by particle swarm, which is

58.8396 hours. It was also found that in general, the Hooke and Jeeves algorithm ran significantly faster than the particle swarm for this particular problem. In this case, particle swarm was still applied in order to offer a comparison of the capabilities between global and local optimisation methods.

For 10mM of Glucose, cancer cells had a cycle period of 15.5924 hours. It is not surprising then that the cycle period has dramatically increased so much from this particular value when the complete set of metabolic rate constants are optimised to bring about the high cycle period value. Biologically speaking for a metabolic network, if a combination of drugs are applied and their effects manage to extend to all areas of the network, then an intuitive result would be that the species involved will all be affected in some manner. Of course, a drug, or even drug combinations, capable of achieving such a feat is unrealistic and so, smaller combinations of the maximum metabolic rate constants were also investigated to see how targeting their corresponding reactions with drug compounds influences the cycle period of cancer cells.

Sets of metabolic rate constants that were involved in the reactions possessing high cancer to normal gene expression ratios did not have much of an effect over the best cycle period value. For example, metabolic rate constants that are part of the reversible reactions between PEP and pyruvate, which had the highest gene expression ratio in the metabolic pathway being modelled,, exhibited a best period value of 16.0474 for the two optimisation algorithms. This value only slightly higher than the original cancer cycle period of 15.5924.

Results generated by the sensitivity analysis performed on the model of cell growth and division for normal and cancer cells, as discussed in chapter 4, were used as a guide for determining what types of metabolic control rate parameters sets could also potentially be investigated using with optimisation methods. For example, as shown in Table 5.2, optimisation with respect to cycle period were performed for the following scenarios related to the results of the sensitivity analysis.

- Maximum metabolic rate constants for cancer cells with a positive sensitivity value (see parameters K_{57} , K_{59} , K_{63} , K_{71} and K_{77} in Table 4.2): Here, the best cycle period value achieved by Hooke and Jeeves, and particle swarm was 16.5742 hours for both optimisation algorithms. Although the optimized period value is only slightly higher than the cycle period of 15.5924 hours for cancer cells prior to model optimisation, it is important to note that for 10mM of Glucose, normal cells demonstrated a cycle period (hrs) of 16.4994. This means that if the combination of reactions related to these particular metabolic rate constants were targeted by drug compounds, the cell division rate of present cancer cells can still be restored to the slower rate of division found in normal cells. The final parameters values for K_{57} , K_{59} and K_{63} were pushed towards the lower boundary value imposed for the optimisation task. These three metabolic parameters are involved in the reactions describing the conversion from FDP to DHAP, FDP to GAP and GAP to DPG respectively. This would suggest that in order to maximise the cell cycle period as much as possible, the drug compounds targeting these three reactions should aim to lower those specific metabolic rate constants as much as possible. The final parameter values for K_{71} and K_{79} occurred at the

upper boundary limit and these were involved in the reactions describing the conversions from PEP to PYR and AA to PYR respectively. Therefore, the drugs targeting these two reactions would be aiming to increase those particular metabolic rate constants.

- Metabolic rate constants with the top 3 greatest sensitivity (absolute value) percentage change from the normal to cancer cell model sensitivities (see parameters K_{56} , K_{74} and K_{78} in Table 4.2): The best cycle period value obtained for both optimisation algorithms was 18.828. The final value for the parameter K_{56} , which is involved in the reaction describing the conversion of FDP to F6P, occurred at the upper limit of the interval used during the optimisation task. Therefore, the drug compound targeting this particular reaction would be aiming to increase the metabolic rate as much as possible. The final values for the parameters K_{74} and K_{78} , which are involved in the reactions describing the conversion from LAC to PYR and AA to MASS, were pushed towards the lower boundary value and hence, the drug compounds targeting these two reactions should aim to lower the metabolic rates.

- Metabolic rate constants with the top 3 greatest (absolute value) sensitivities (see parameters K_{33} , K_{64} and K_{78} in Table 4.2): The best cycle period value achieved by both optimisation algorithms now becomes the significantly higher value of 23.7425. The final values for the parameters K_{33} and K_{78} , which are involved in the reactions describing the conversion from Glucose to G6P and AA to MASS, occurred at the lower boundary value of the interval. Therefore, in order to maximise the cell cycle period, the drug compounds targeting these two reactions should aim to lower the metabolic

rates as much as possible. The final value for the parameter K_{64} , which is involved in the reaction describing the conversion of DPG to GAP, occurred at the upper limit of the interval used during the optimisation task. Therefore, the drug compound targeting this specific reaction should aim to increase the metabolic rate.

From these results, it appears that the major factor towards achieving as high a cycle period value as possible is not whether the optimised parameters have a positive sensitivity value (if the value of those particular parameters are increased, the cycle period should increase as well), but is dependent on those parameters that exhibit the largest magnitudes in sensitivity value.

Another observation made from Table 5.2 was that when only the metabolic rate constant (K_{33}) involved in the conversion from glucose to G6P was subjected to optimisation, it showed quite a high impact factor over the cell cycle period on its own with a corresponding best value of 19.9447. As postulated in chapter 4, this reinforces the idea that K_{33} would make an effective potential target for drug compounds.

SUMMARY AND FUTURE DIRECTIONS

Summary:

Understanding the changes made to events at a cellular level could aid in the prevention or treatment of diseases including cancer at the organismal level.

Systems Biology looks at the specific interactions of components found in biological systems, which are composed of many different types of multifunctional elements, all interacting selectively and non-linearly with each other to form coherent system behaviour. This type of approach requires a combination of theoretical, computational and experimental work to be carried out in order to gain a comprehensive understanding of complex biological systems such as the networks involved in cell cycle and metabolic pathways.

Mammalian cell cycle models have arisen from previous mathematical models based on the budding yeast cycle. This is due to the fact

that yeast and mammalian cells share fundamentally similar system properties such as the cyclin-dependent kinase regulatory systems. These similarities can be identified from molecular interaction maps that use a standardized set of notation to describe the interactions between various entities.

Genome replication during S, or DNA synthesis stage of the cell cycle, and the halving of the cell mass to reflect the splitting of a parent cell into two identical daughter cells, are the two major events that must be included in the model of a cycle division cycle. The mathematical model should also reflect the fact it is compulsory for cell cycle events to be highly regulated in a temporal manner into to ensure successful cell growth and division.

One limitation of current cell cycle models is that they rarely account for certain precursors of cell growth such as energy usage and the need for non-essential amino acids as fundamental building blocks of proteins, which are required themselves to drive the cell division process.

At present, a mathematical model consisting of 36 ODEs that describe species found in both the mammalian cell cycle and metabolic pathways has been developed. This model is based upon a previous cycle model for mammalian cells by Conradie *et al.* (2010), which was described in chapter 2. This model of cell growth and division was discussed in detail in chapter 3 and includes species that play a vital role in glycolysis and the synthesis of non-essential amino acids. The model was designed in such a way as to allow it to respond to different concentrations of glucose in a reasonable depiction

of what is actually happening in real-life. Time course plots of the cell mass demonstrated that regardless of the initial concentration of glucose, the cell mass eventually settles into a stable cycle.

A version of the cell growth and division model was also created for cancer cells by taking into account the fact that the expression levels of genes associated with metabolism are often upregulated for transformed cells compared to normal cells. The genes involved in expressing the enzymes that catalyse the reactions of glycolysis and non-essential amino acid pathways were identified and the ratio of gene expression levels between cancer and normal cells were calculated. These gene expression ratios were incorporated into the mammalian cell cycle and metabolism model by multiplying the ratio values with the relevant parameters to produce a cancer growth and division model.

Sensitivity analysis was performed on the maximum metabolic rate constants for both versions of the cell growth and division model to determine which of the parameters had the greatest control over the time it takes to complete one full cell division cycle. The sign of the sensitivity values obtained determines whether a perturbation made to the parameter of interest will cause the cell cycle period to increase or decrease. It was shown that the rate constant involved in the reaction describing the conversion from glucose to G6P exhibited the highest magnitude out of all the other sensitivities. This makes that particular reaction, and the enzymes that catalyse it (hexokinase and glucokinase) potential drug targets.

Finally, different optimisation algorithms were used to mirror the effects of a combination of drug compounds by determining whether certain sets of metabolic rate constants and their optimised values affect the rate of cell division in cancer cells. The aim of single and combinations of drug compounds would be to increase the cell cycle period of cancer cells until it is at least equal to or greater than the cycle period for normal mammalian cells. It was discovered that the magnitude of the sensitivity value attached to a particular metabolic rate constant was highly important in determining whether drug combinations targeting their related pathway reactions showed a high level of impact on the cancer cell cycle period.

Future Directions For The Model:

In order to provide a more complete and comprehensive cell cycle and metabolism model, the current set of ODEs could be extended to include the behaviour of the pathway involved in the oxidative phosphorylation process. Throughout this building process, it should be ensured that the cell mass continues to respond realistically to differences in the concentration of glucose. In addition, the creation and analysis of ODEs describing the effects of ATP and oxygen may offer deeper insights into the Warburg effect. By perturbing specific concentrations throughout the metabolic pathway, the observed effects can be used to ascertain which metabolites have the highest influence over cellular metabolism. The quantitative values of the species concentrations at different stages of the cell cycle could be quantitatively

measured and analysed to investigate the differences in the glycolytic rate between normal and cancer cells.

Ultimately, the model could be used to investigate which metabolites or section of the metabolic pathway becomes dysregulated during cancer cell metabolism. Detailed knowledge of the key differences of a metabolic pathway between normal proliferating cells and tumour cells may enable the development of chemical compounds with superior efficacy and efficiency that exploit the weaknesses of cancer cells without detrimentally affecting normal cells.

Another potential extension that could be made to the ODEs describing the systems of the cell cycle and metabolic pathway is to introduce noise terms. Trajectories modelled by ODEs may sometimes not match up with what is happening in reality as large and complex regulatory networks often give rise to chaotic system behaviour. Therefore, by transforming the ODEs into stochastic differential equations (SDEs), an element of unpredictability could be factored in.

One important aspect of combination therapy that could be investigated further, using the mammalian cell cycle and cancer model, is the dosing schedule for combinations of drug compounds.

Administering a combination of drugs in a certain order and at a specific time can have a profound effect on the efficacy of the therapy being applied. This is illustrated by a recent piece of theoretical work carried out where a predictive mathematical model was used to explore the effects of administering a combination of SNS-314, an Aurora kinase inhibitor, and Docetaxel (Orrell and Fernandez, 2010). Orrell and Fernandez discovered

that the sequence of the drugs being given was very important as an administration of SNS-314 followed by Docetaxel was shown to be much more effective in reducing the size of the tumour being treated than administering the two drugs in the alternative order.

The model of cell growth and division could be employed to predict the effects of scheduling multiple drug treatments on system properties such as the cell cycle period in particular; the simulation results should then be tested and verified by related experiments.

BIBLIOGRAPHY

Alberts, B., Bray, D., Johnson A., Lewis, J., Raff, M., Roberts, K., Walter, P., (1997), "*Essential cell biology: an introduction to the molecular biology of the cell*", Garland Publishing Inc, Har/Cdr edition.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., (2002), "*Molecular Biology of the Cell*", New York and London: Garland Publishing.

Ahn, A.C., Tewari, M., Poon, C., Phillips, R.S., (2006), "*The clinical applications of a systems approach*", PLoS medicine, 3, e209.

Alarcón, T., Byrne, H.M., Maini, P.K. (2004), "*A mathematical model of the effects of hypoxia on the cell-cycle of normal and cancer cells*", Journal of theoretical biology, 229, 395-411.

Angeli, D., Ferrel, J.E., Sontag, E.D. (2004), "*Detection of multistability, bifurcations, and hysteresis in large class of biological positive-feedback systems*", PNAS, 7, 1822-1827.

Atkinson, K.A. (1989), “*An Introduction to Numerical Analysis*”, New York: John Wiley & Sons, 2nd edition.

Arnold, D. (2008), “*Assessing the role of combination therapy in mCRC*”, EJC Supplements, 6, 5-11.

Bäck, T., Schwefel, H.P. (1997), “An overview of evolutionary algorithms for parameter optimization”, Evolutionary Computation, 1, 1-23.

Barrett, T., Troup, D.B., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Muerter, R.N., Holko, M., Ayanbule, O., Yefanov, A. Soboleva, A. (2011), “*NCBI GEO: archive for functional genomics data sets – 10 years on*”, Nucleic Acids Res., 39 (Database Issue): D1005-1010.

Bentele, M., Larvik, I., Ulrich, M., Stöber, S., Heermann, D.W., Kalthoff, H., Krammer, P.H., Eils, R. (2004), “*Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis*”, Journal of Cell Biology, 166, 839-851.

Bertalanff, L. (1968), “*General System Theory: Foundations, Development, Applications*”, New York: George Braziller, revised edition 1976.

Blüthgen, N., Herzog, H. (2001), “*NAP-Kinase-Cascade: Switch, Amplifier or Feedback Controller*”, 2nd Workshop on Computation of Biochemical Pathways and Genetic Networks, Logos Verlag Berlin, 55-62.

Bolstad, B.M., Irizarry, R.A., Astrand, M., Speed, T.P. (2003), “*A Comparison of Normalization Methods for High Density Oligonucleotide Array Data Based on Bias and Variance*”, *Bioinformatics*, 19, 185-193.

Bruggeman, F.J., Westerhoff, H.V. (2007), “*The nature of systems biology*”, *Trends in microbiology*, 15, 45-50.

Butcher, J.C. (2003), “*Numerical Methods For Ordinary Differential Equations*”, New York: John Wiley & Sons.

Chen, K.C., Calzone, L., Csikász-Nagy, A., Cross, F.R., Novák, B., Tyson, J.J. (2004), “*Integrative Analysis of Cell Cycle Control in Budding Yeast*”, *Molecular Biology of the Cell*, 15, 3841-3862.

Chen, K.C., Csikász-Nagy, A., Gyorffy, B., Val, J., Novák, B., Tyson, J.J. (2000), “*Kinetic Analysis of a Molecular Model of the Budding Yeast Cell Cycle*”, *Molecular Biology of the Cell*, 11, 369-391.

Conradie, R., Bruggeman, F.J., Ciliberto, A., Csikász-Nagy, A., Novák, B., Westerhoff, H.V., Snoep, J.L. (2010), “*Restriction Point Control of the Mammalian Cell Cycle via the Cyclin E/Cdk2:p27 Complex*”, *FEBS Journal*.

Cross, F.R., Archambault, V., Miller, M., Klovstad, M. (2002), "*Testing a mathematical model of the yeast cell cycle*", Mol. Biol. Cell, 13, 52-70.

DeVita, V.T., Hellman, S., Rosenberg, S.A. (2001), "*Cancer: Principles and Practice of Oncology*", Lippincott Williams and Wilkins, 6th edition.

Edgar, R., Domrachev, M., Lash, A.E. (2002), "*Gene Expression Omnibus: NCBI gene expression and hybridization array data repository*", Nucleic Acids Res., 30, 207-210.

Faure, A., Thieffry, D. (2009), "*Logical modelling of cell cycle control in eukaryotes: a comparative study*", Molecular Biosystems, 5, 1567-1581.

Feller, A.D. (1994), "*Specifying epigenetic states with autoregulatory transcription factors*", Journal of Theoretical Biology, 170, 175-181.

Fogel, D.B., Fogel, L.J., Atmar, J.W. (1992), "*Meta-evolutionary programming*", 25th Asiloma Conference on Signals, Systems and Computers, IEEE Computer Society, Asilomar, 540-545.

Fuß, H., Dubitzky, D., Downes, C.S., Kurth, M.J. (2005), "*Mathematical model of cell cycle regulation*", Briefings in Bioinformatics, 6, 163-177.

Futreal, P.A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., Rahman, N., Stratton, M.R. (2004), “*A census of human cancer genes*”, *Nat. Rev. Cancer*, 4, 177-183.

Gaglio, D., Metallo, C.M., Gameiro, P.A., Hiller, K., Danna, L.S., Balestrieri, C., Alberghina, L., Stephanopoulos, G., Chiaradonna, F. (2011), “*Oncogenic K-Ras decouples glucose and glutamine metabolism to support cancer cell growth*”, *Molecular Systems Biology*, 7.

Gatenby, R., Gillies, R. (2010), “*Why Do Cancers Have High Aerobic Glycolysis*”, *Nature Reviews*, 4, 891-899.

Gauthier, J., Pohl, P. (2011), “*A general framework for modelling growth and division of mammalian cells*”, *BMC Systems Biology*, 5.

Gentleman *et al.* (2004), “*Bioconductor: open software development for computational biology and bioinformatics*”, *Genome Biology*, 5.

Gibson, M.A., Bruck, J. (2000), “*Efficient Exact Stochastic Simulation of Chemical Systems With Many Species and Many Channels*”, *J. Phys. Chem.*, A104, 1876-1889.

Gupta, A., Lee, Y., Galoforo, S., Berns, C., Martinez, A., Corry, P., Wu, X., Guan, K. (1997), “*Differential effect of glucose deprivation on MAPK*

activation in drug sensitive human breast carcinoma MCF-7 and multidrug resistant MCF-7/ADR cells”, Mol. Cell. Biochem. 170, 23–30.

Hammerman, P., Fox, C., Thompson, C. (2004), “*Beginnings of a signal-transduction pathway for bioenergetic control of cell survival*”, Trends Biochem. Sci., 29, 586–592.

Hanahan, D., Weinberg, R.A. (2000), “*The hallmarks of cancer*”, Cell, 100, 57-70.

Harris, S.L., Levine, A.J. (2005), “*The p53 pathway: positive and negative feedback loops*”, Oncogene, 24, 2899-2908.

Hartwell, L.H. (2002), “*Yeast and Cancer*,” The Nobel Prizes 2001, ed. Tore Frängsmyr (Nobel Foundation, Stockholm).

Heiden, M., Cantley, L. (2009), “*Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation*”, Science, 324, 1029-1033.

Higuchi, C., Sanaka, T., Sato, T., Omata, M., Watanabe, M., Mine, S., Inuzuka, N., Nihei, H. (1997), “*The effect of glucose on the proliferation of peritoneal fibroblasts*”, Adv. Perit. Dial., 13, 253-256.

Hooke, R., Jeeves, T.A. (1961), “*Direct search solution of numerical and statistical problems*”, *Journal of the Association for Computing Machinery*, 8, 212-229.

Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., Singhal, M., Xu, L., Mendes, P., Kummer, U. (2006), “*COPASI—a Complex Pathway Simulator*”, *Bioinformatics*, 22, 3067-3074.

Hornberg, J.J., Bruggeman, F.J., Westerhoff, H.V., Lankelma, J. (2006), “*Cancer: a Systems Biology disease*”, *Bio. Systems*, 83, 81-90.

Hornberger, G.M., Spear, R.C. (1981), “*An approach to the preliminary analysis of environmental systems*”, *J. Environ. Manage.*, 12, 7-18.

Ideker, T., Galitski, T., Hood, L. (2001), “*A new approach to decoding life: systems biology*”, *Essays Biochem.*, 45, 177-193.

Ingalls, B. (2008), “*Sensitivity analysis: from model parameters to system behaviour*”, *Annual review of genomics and human genetics*, 2, 343-372.

Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B., Speed, T.P. (2003), “*Summaries of Affymetrix GeneChip probe level data*”, *Nucleic Acids Research*, 31.

Jorgensen, P., Tyers, M. (2004), "*How Cells Coordinate Growth and Division*", *Current Biology*, 14, 1014-1027.

Kacser, H. (1957), "*Some physio-chemical aspects of biological organization*", Appendix to *The strategy of the genes* by C.H. Waddington, George Allen and Unwin, London, 191-249.

Kacser, H. (1960), "*Kinetic models of development and heredity*", *Symp. Soc. Exp. Biol.*, 13-27.

Kacser, H., Burns, J. (1973), "*The control of flux*", *Symp. Soc. Exp. Biol.*, 27, 65-104.

Kanehisa, M. (1996), "*Toward pathway engineering: a new database of genetic and molecular pathways*", *Science & Technology Japan*, 59, 34-38.

Kanehisa, M., Goto, S. (2000), "*KEGG: Kyoto Encyclopedia of Genes and Genomes*", *Nucleic Acids Res.*, 28, 37-30.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., Tanabe, M. (2012), "*KEGG for integration and interpretation of large-scale molecular datasets*", *Nucleic Acids Res.*, 40, D109-D114.

Kennedy, J., Eberhart, R. (1995), "*Particle Swarm Optimization*", Proceedings of the Fourth IEEE International Conference on Neural Networks, Perth, Australia, 1942-1948.

Kent, E., Hoops, S., Mendes, P. (2012), "*Condor-COPASI: High-Throughput Computing for Biochemical Networks*", BMC Systems Biology, 6(91).

Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P. (1983), "*Optimization by simulated annealing*", Science, 220, 671-680.

Kitano, H. (2002), "*Computational systems biology*", Nature, 420, 206-210.

Kohn, K.W. (1999), "*Molecular interaction map of the mammalian cell cycle control and DNA repair systems*", Molecular Biology of the Cell, 10, 2703-2734.

Kohn, K.W., Aladjem, M.I. (2006a), "*Circuit diagrams for biological networks*", Molecular Systems Biology, 2.

Kohn, K.W. (2006b), "*Molecular interaction map of the mammalian cell cycle control and DNA repair systems*", Molecular Biology of the Cell, 10, 2703-2734.

Lander, E.S. *et al.* (2001), "*Initial sequencing and analysis of the human genome*", Nature, 409, 860-921.

Laubenbacher, R., Hower, V., Jarrah, A., Torti, S.V., Shulaev, V., Mendes, P., Torti, F.M., Akman, S. (2009), “A *systems biology view of cancer*”, *Biochimica et biophysica acta*, 1796, 129-139.

Le Novère, N., Hucka, M., Mi, H., Moodie, S., Schreiber, F., Sorokin, A. *et al.* (2009), “*The Systems Biology Graphical Notation*”, *Nat. Biotechnol.*, 27, 735-741.

Lee, S.H., Ryu, J.K., Lee, K., Woo, S.M., Park, J.K., Yoo, J.W., Kim, Y., Yoon, Y.B. (2008), “*Enhanced anti-tumour effect of combination therapy with gemcitabine and apigenin in pancreatic cancer*”, *Cancer Letters*, 259, 39-49.

Lee, Y., Galoforo, S., Berns, C., Chen, J., Davis, B., Sim, J., Corry, P., Spitz, D. (1998), “*Glucose deprivation-induced cytotoxicity and alterations in mitogen-activated protein kinase activation are mediated by oxidative stress in multidrug-resistant human breast carcinoma cells*”, *J. Biol. Chem.* 273, 5294–5299.

Lehninger, A.L. (1982), “*Principles of Biochemistry*”, Worth Publishers, New York, Chapter 22.

Levenberg, K. (1944), “*A method for the solution of certain nonlinear*

problems in least squares”, *Quart. Appl. Math.*, 2, 164-168.

Li, Y., Liu, L., Tollefsbol, T.O. (2010), “*Glucose Restriction Can Extend Normal Cell Lifespan And Impair Precancerous Cell Growth Through Epigenetic Control Of hTERT and p16 expression*”, *The FASEB Journal*, 24(5), 1442-1453.

Liu, H., D’Andrade, P., Fulmer-Smentek, S., Lorenzi, P., Kohn, K.K., Weinstein, J.N., Pommier, Y., Reinhold, W.C. (2010), “*mRNA and microRNA Expression Profiles of the NCI-60 Integrated with Drug Activites*”, *Molecular Cancer Therapeutics*, 9.

López-Lázaro, M. (2008), “*The warburg effect: why and how do cancer cells activate glycolysis in the presence of oxygen*”, *Anticancer Agents Med Chem.*, 8(3), 305-312.

Marino, S., Hogue, I.B., Ray, C.J. (2008), “*A Methodology For Performing Global Uncertainty And Sensitivity Analysis In Systems Biology*”, *Journal of Theoretical Biology*, 254, 178-196.

Marquardt, D.W. (1963), “*An algorithm for least squares estimation of nonlinear parameters*”, *SIAM Journal*, 11, 431-441.

Mendes, P. (1993), “*GEPASI: A software package for modeling the dynamics, steady states and control of biochemical and other systems*”, *Computer applications in the biosciences*, 9, 563-571.

Monod, J. (1949), "*The Growth of Bacterial Cultures*", Annual Review of Microbiology, 3, 371.

Nasmyth, K. (1996), "*At the heart of the budding yeast cell cycle*", Trends in genetics, 12, 405-412.

Nelder, J.A., Mead, R. (1965), "*A simplex method for function minimization*", Computer Journal, 7, 308-313.

Novák, B., Tyson, J.J. (1993), "*Numerical analysis of a comprehensive model of M-phase control in Xenopus oocyte extracts and intact embryos*", J. Cell Sci., 106, 1153-1168.

Novák, B., Tyson, J.J. (1997), "*Modeling the control of DNA replication in fission yeast*", Proc. Natl. Acad. Sci. (USA), 94, 9147-9152.

Novák, B., Tyson, J.J. (2004), "*A model for restriction point control of the mammalian cell cycle*", Journal of Theoretical Biology, 230, 563-579.

Novák, B., Csikasz-Nagy, A., Gyorffy, B., Chen, K., Tyson, J.J. (1998), "*Mathematical model of the fission yeast cell cycle with checkpoint controls at the G1/S, G2/M and metaphase/anaphase transitions*", Biophys. Chem., 72, 185-200.

Novák, B., Toth, A., Csikasz-Nagy, A., Gyorffy, B., Tyson, J.J., Nasmyth, K. (1999), "*Finishing the cell cycle*", Journal of Theoretical Biology, 199, 223-233.

Orrell, D., Fernandex, E. (2010), "*Using Predictive Mathematical Models to Optimise the Scheduling of Anti-Cancer Drugs* ", Innovations in Pharmaceutical Technology, Biopharma, 59-62.

Pardee, A. (1974), "*A restriction point for control of normal animal cell proliferation*", PNAS, 71, 1286-1290.

Perumal, T.M., Gunawan, R. (2011), "*Understanding dynamics using sensitivity analysis: caveat and solution*", BMC Systems Biology, 5(41).

Petzold, L. (1983), "*Automatic Selection of Methods for Solving Stiff and Nonstiff Systems of Ordinary Differential Equations*", SIAM J. Sci. Stat. Comput., 4, 136-148.

Pray, L. (2008), "*L.H. Hartwell's yeast: A model organism for studying somatic mutations and cancer*", Nature Education, 1(1).

Rabitz, H., Kramer, M., Dacol, D. (1983), "*Sensitivity analysis in chemical kinetics*", Annu. Rev. Phys. Chem., 34, 419-461.

Rabitz, H. (1989), “*Systems analysis at the molecular scale*”, Science, 246, 221-226.

Raue, A., Kreutz, C., Maiwald, T., Klingmuller, U., Timmer, J. (2011), “*Addressing parameter identifiability by model-based experimentation*”, IET Systems Biology, 5, 120.

Reinhold, W.C., Sunshine, M., Liu, H., Varma, S., Kohn, K.K., Morris, J., Doroshow, J., Pommier, Y. (2012), “*CellMiner: A Web-Based Suite of Genomic and Pharmacologic Tools to Explore Transcript and Drug Patterns in the NCI-60 Cell Line Set*”, Cancer Research, 72.

Robert, J., Vekris, A., Pourquier, P., Bonnet, J. (2004), “*Predicting drug response based on gene expression*”, Crit. Rev. Oncology/Hematology, 51, 205-227.

Rodriguez-Fernandez, M., Mendes, P., Banga, J.R. (2006), “*A hybrid approach for efficient and robust parameter estimation in biochemical pathways*”, Biosystems, 83, 248-265.

Rosen, R. (1958a), “*A Relational Theory Of Biological Systems*”, Bulletin of Mathematical Biophysics, 20, 245-260.

Rosen, R. (1958b), "*The Representation of Biological Systems from the Standpoint of the Theory of Categories*", *Bulletin of Mathematical Biophysics*, 20, 317-341.

Saltelli, A., Ratto, M, Tarantola, S., Campolongo, F. (2005), "*Sensitivity analysis for chemical models*", *Chem. Rev.*, 105, 2811-2827.

Shankavaram, U.T., Varma, S., Kane, D., Sunshine, M., Chary, K.K., Reinhold, W.C., Pommier, Y., Weinstein, J.N. (2009), "*CellMiner: a relational database and query tool for the NCI-60 cancer cell lines*", *BMC Genomics*, 10.

Skobe, M., Fusenig, N.E. (1998), "*Tumorigenic conversion of immortal human keratinocytes through stromal cell activation*", *Proc. Natl. Acad. Sci. USA*, 95, 1050-1055.

Smith, K.A. (2005), "*Determining to Divide: How Do Cells Decide?*", *Journal of Biological Physics*, 31, 261-272.

Swat, M., Kel, A., Herzel, H. (2004), "*Bifurcation analysis of the regulatory modules of the mammalian G1/S transition*", *Bioinformatics*, 20, 1506-1511.

Tyson, J.J., Novák, B, Odell, G.M., Chen, K., Thron, C.D. (1996), "*Chemical kinetic theory: understanding cell-cycle regulation*", *Trends Biochem. Sci.*, 21, 89-96.

Tyson, J.J., Chen, K., Novák, B. (2001), “*Network dynamics and cell physiology*”, *Nature reviews, Molecular cell biology*, 2, 908-916.

Venter, J.C. (2001), “*The sequence of the human genome*”, *Science*, New York, 291, 1304-1351.

Vogelstein, B., Kinzler, K.W. (2004), “*Cancer genes and the pathways they control*”, *Nature medicine*, 10, 789-99.

Warburg, O. (1956), “*On the Origin of Cancer Cells*”, *Science*, 123, 309-314.

Weng, G. (1999), “*Complexity in Biological Signaling Systems*”, *Science*, 284, 92-96.

Westerhoff, H.V. (2005), “*Systems biology...in action*”, *Current Opinion in Biotechnology*, 16, 326-328.

Yao, G, Lee, T.J., Mori, S., Nevins, J.R., You, L. (2008), “*A bistable Rb-E2F switch underlies the restriction point*”, *Nature Cell Biology*, 10, 476-482.

Zahorowska, B., Crowe, P.J., Yang, J.L. (2009), “*Combined therapies for cancer: a review of EGFR-targeted monotherapy and combination treatment with other drugs*”, *J. Cancer Res. Clin. Oncol.*, 135, 1137-48.

Zetterberg, A., Larsson, O. (1995), “*Cell cycle progression and growth in mammalian cells: kinetic aspects of transition events*”, Oxford University Press, Hutchison C and Glover DM edition.

Zheng, Y., Rundell, A. (2006), “*Comparative study of parameter sensitivity analyses of the TC-activated Erk-MAPK signaling pathway*”, 153, 201-211.

Zi, Z., Zheng, Y., Rundell, A.E., Klipp, E. (2008), “*SBML-SAT: a systems biology markup language (SBML) based sensitivity analysis tool*”, BMC Bioinf., 9, 342.

APPENDIX

A

MODELS OF THE CELL CYCLE – SUPPLEMENTARY INFORMATION

Reaction Rate Laws:

Table A.1. The reaction steps and their corresponding rate equations in the Conradi *et al.* mammalian cell cycle model.

Reaction Number	Rate Equations
1	$V_1 = k16 \cdot ERG(t)$
2	$V_2 = k18 \cdot DRG(t)$
3	$V_3 = K10 \cdot CD(t)$
4	$V_4 = K10 \cdot CYCD(t)$
5	$V_5 = K25 \cdot P27(t) \cdot CYCE(t)$
6	$V_6 = K25 \cdot P27(t) \cdot CYCA(t)$
7	$V_7 = k24 \cdot P27(t) \cdot CYCD(t)$
8	$V_8 = k24r \cdot CD(t)$
9	$V_9 = K30 \cdot CDc20(t) \cdot CYCA(t)$
10	$V_{10} = K30 \cdot CDc20(t) \cdot CA(t)$
11	$V_{11} = K25R \cdot CE(t)$

12	$V_{12} = 25R \cdot CA(t)$
13	$V_{13} = V8 \cdot CE(t)$
14	$V_{14} = V8 \cdot CYCE(t)$
15	$V_{15} = V6 \cdot P27(t)$
16	$V_{16} = V6 \cdot CE(t)$
17	$V_{17} = V6 \cdot CD(t)$
18	$V_{18} = V6 \cdot CA(t)$
19	$V_{19} = V2 \cdot CYCB(t)$
20	$V_{20} = \frac{(K3a + K3 \cdot CDc20(t)) \cdot (1 - CDh1(t))}{J3 - CDh1(t) + 1}$
21	$V_{21} = \frac{V4 \cdot CDh1(t)}{J4 + CDh1(t)}$
22	$V_{22} = K34 \cdot PPX(t)$
23	$V_{23} = \frac{K31 \cdot CYCB(t) \cdot (1 - IEP(t))}{J31 - IEP(t) + 1}$
24	$V_{24} = \frac{K32 \cdot PPX(t) \cdot IEP(t)}{J32 + IEP(t)}$
25	$V_{25} = K12 \cdot CDc20T(t)$
26	$V_{26} = \frac{K13 \cdot IEP(t) \cdot (CDc20T(t) - CDc20(t))}{J13 - CDc20(t) + CDc20T}$
27	$V_{27} = \frac{K14 \cdot CDc20(t)}{J14 + CDc20(t)}$
28	$V_{28} = K12 \cdot CDc20(t)$
29	$V_{29} = E2F_Rb(t) \cdot K20 \cdot (CYCDT(t) \cdot LD + CYCA(t) \cdot LA + CYCB(t) \cdot LB + CYCE(t) \cdot LE)$
30	$V_{30} = p_E2F_Rb(t) \cdot K20 \cdot (CYCDT(t) \cdot LD + CYCA(t) \cdot LA + CYCB(t) \cdot LB + CYCE(t) \cdot LE)$
31	$V_{31} = K27 \cdot MASS \cdot If \left[\frac{Rb(t) + E2F_Rb(t) + p_E2F_Rb(t)}{pp_Rb(t) + Rb(t) + E2F_Rb(t) + p_E2F_Rb(t)} > 0.8; 0, 1 \right]$

32	$V_{32} = K28 \cdot GM(t)$
33	$V_{33} = \varepsilon \cdot \mu \cdot GM(t)$
34	$V_{34} = \frac{\varepsilon \cdot k15}{1 + (DRG(t)/J15)^2}$
35	$V_{35} = \varepsilon \cdot (K11a + K11 \cdot CYCB(t))$
36	$V_{36} = \varepsilon \cdot K29 \cdot E2F(t) \cdot MASS(t)$
37	$V_{37} = \varepsilon \cdot K33$
38	$V_{38} = \varepsilon \cdot (K7a + K7 \cdot E2F(t))$
39	$V_{39} = \varepsilon \cdot K9 \cdot DRG(t)$
40	$V_{40} = \varepsilon \cdot K5$
41	$V_{41} = \varepsilon \cdot (k17a \cdot ERG + \frac{k17 \cdot (DRG(t)/J17)^2}{1 + (DRG(t)/J17)^2})$
42	$V_{42} = \varepsilon \cdot (K1a + \frac{K1 \cdot (CYCB(t)/J1)^2}{1 + (CYCB(t)/J1)^2})$
43	$V_{43} = Rb(t) \cdot K20 \cdot (CYCDT(t) \cdot LD + CYCA(t) \cdot LA + CYCB(t) \cdot LB + CYCE(t) \cdot LE)$
44	$V_{44} = pp_Rb(t) \cdot (K19a \cdot (PP1T - PP1A) + K19 \cdot PP1A)$
45	$V_{45} = E2F_Rb(t) \cdot K26R$
46	$V_{46} = E2F(t) \cdot (K23a + K23 \cdot (CYCA(t) + CYCB(t)))$
47	$V_{47} = p_E2F(t) \cdot K22$
48	$V_{48} = E2F(t) \cdot Rb(t) \cdot K26$
49	$V_{49} = p_E2F_Rb(t) \cdot K26R$
50	$V_{50} = Rb(t) \cdot p_E2F(t) \cdot K26$
51	$V_{51} = p_E2F_Rb(t) \cdot K22$
52	$V_{52} = E2F_Rb(t) \cdot (K23a + K23 \cdot (CYCA(t) + CYCB(t)))$

Conradie *et al.* Model ODEs:

$$CA'(t) = V_6 - V_{10} - V_{12} - V_{18}$$

$$CD'(t) = V_7 - V_8 - V_{17} - V_3$$

$$CDc20'(t) = V_{26} - V_{27} - V_{28}$$

$$CDc20T'(t) = V_{35} - V_{25}$$

$$CDh1'(t) = V_{20} - V_{21}$$

$$CE'(t) = V_5 - V_{11} - V_{13} - V_{16}$$

$$CYCA'(t) = V_{36} - V_9 - V_6 + V_{12} + V_{18}$$

$$CYCB'(t) = V_{42} - V_{19}$$

$$CYCD'(t) = V_{39} + V_{17} + V_8 - V_7 - V_4$$

$$CYCE'(t) = V_{38} - V_{14} - V_5 + V_{11} + V_{16}$$

$$DRG'(t) = V_{41} - V_2$$

$$E2F'(t) = V_{29} + V_{45} + V_{47} - V_{46} - V_{48}$$

$$E2F_Rb'(t) = V_{51} + V_{48} - V_{52} - V_{29} - V_{45}$$

$$ERG'(t) = V_{34} - V_1$$

$$GM'(t) = V_{31} - V_{32}$$

$$IEP'(t) = V_{23} - V_{24}$$

$$MASS'(t) = V_{33}$$

$$P27'(t) = V_{40} + V_3 + V_8 - V_{15} - V_5 - V_6 - V_7 + V_{11} + V_{12} + V_{13} + V_{10}$$

$$p_E2F'(t) = V_{30} + V_{49} + V_{46} - V_{47} - V_{50}$$

$$p_E2F_Rb'(t) = V_{52} + V_{50} - V_{51} - V_{30} - V_{49}$$

$$pp_Rb'(t) = V_{29} + V_{30} + V_{43} - V_{44}$$

$$PPX'(t) = V_{37} - V_{22}$$

$$Rb'(t) = V_{44} + V_{45} + V_{49} - V_{48} - V_{50} - V_{43}$$

Conradie *et al.* Model Definitions And Steady-State Relations:

$$PP1A = \frac{PP1T}{1 + K21 \cdot (FE \cdot (CYCA(t) + CYCE(t)) + FB \cdot CYCB(t))}$$

$$V2 = K2aa \cdot CDc20(t) + K2a \cdot (1 - CDh1(t)) + K2 \cdot CDh1(t)$$

$$V4 = K4 \cdot (GA \cdot CYCA(t) + GB \cdot CYCB(t) + GE \cdot CYCE(t))$$

$$V6 = K6a + K6 \cdot (HA \cdot CYCA(t) + HB \cdot CYCB(t) + HE \cdot CYCE(t))$$

$$V8 = \frac{(YE \cdot (CYCA(t) + CYCE(t)) + YB \cdot CYCB(t)) \cdot K8}{CYCET + J8} + K8a$$

$$CYCET = CE(t) + CYCE(t)$$

$$CYCDT = CD(t) + CYCD(t)$$

$$CYCAT = CA(t) + CYCA(t)$$

$$P27T = CA(t) + CD(t) + CE(t) + P27(t)$$

$$\varepsilon(t) = 1$$

Conradie *et al.* Model Parameters:

k15 = 0.025
k16 = 0.025
J15 = 0.1
k17a = 0.035
k17 = 1.
J17 = 0.3
k18 = 1.
K9 = 0.25
K10 = 0.5
k24 = 100.
k24r = 1.
K7a = 0.
K7 = 0.06
K8a = 0.01
K8 = 0.2
K25 = 100.
K25R = 1.
J8 = 0.1
YE = 1.
YB = 0.05
K29 = 0.005
K30 = 2.
K1a = 0.01
K1 = 0.06
J1 = 0.1
K2a = 0.005
K2 = 2.
K2aa = 0.1
K5 = 2.
K6a = 1.
K6 = 10.
HE = 0.5
HB = 1.
HA = 0.5
LD = 3.3
LE = 5.
LB = 5.
LA = 3.
K20 = 1.
K19a = 0.
K19 = 2.
K21 = 1.
PP1T = 1.
FE = 25.

FB = 2.
K3a = 0.75
K3 = 14.
J3 = 0.01
J4 = 0.01
K4 = 4.
GE = 0.
GB = 1.
GA = 0.3
K33 = 0.005
K34 = 0.005
K31 = 0.07
K32 = 0.18
J31 = 0.01
J32 = 0.01
K11a = 0.
K11 = 0.15
K12 = 0.15
K13 = 0.5
K14 = 0.25
J13 = 0.005
J14 = 0.005
K22 = 0.1
K23a = 0.0005
K23 = 0.1
K26 = 1000.
K26R = 20.
K27 = 0.02
K28 = 0.02
 $\mu = 0.0061$

Conradie *et al.* Model Initial Concentrations:

$$CA(0) = 0.0356927$$

$$CD(0) = 0.010976$$

$$Cdc20(0) = 0.00220177$$

$$Cdc20T(0) = 2.36733$$

$$CDh1(0) = 0.000653278$$

$$CE(0) = 0.000542587$$

$$CYCA(0) = 1.4094$$

CYCB(0) = 2.72898
CYCD(0) = 0.43929
CYCE(0) = 0.0229112
DRG(0) = 0.900533
ERG(0) = 0.0121809
GM(0) = 1.35565
IEP(0) = 0.154655
MASS(0) = 1.68776
P27(0) = 0.00922806
PPX(0) = 1
pp-RB(0) = 9.97574
E2F(0) = 0.989986
p-E2F(0) = 3.98594
Rb(0) = 0.000190871
E2F-Rb(0) = 0.00478911
p-E2F-Rb(0) = 0.0192822

APPENDIX

B

A MODEL OF GROWTH AND DIVISION – SUPPLEMENTARY INFORMATION

Reaction Rate Laws:

Table B.1. The kinetics involved in the mammalian cell cycle and metabolism model.

Reaction Number	Rate Equations
1	$V_1 = 10 \cdot k16 \cdot ERG(t)$
2	$V_2 = 10 \cdot k18 \cdot DRG(t)$
3	$V_3 = 10 \cdot K10 \cdot CD(t)$
4	$V_4 = 10 \cdot K10 \cdot CYCD(t)$
5	$V_5 = 10 \cdot K25 \cdot P27(t) \cdot CYCE(t)$
6	$V_6 = 10 \cdot K25 \cdot P27(t) \cdot CYCA(t)$
7	$V_7 = 10 \cdot k24 \cdot P27(t) \cdot CYCD(t)$
8	$V_8 = 10 \cdot k24r \cdot CD(t)$
9	$V_9 = 10 \cdot K30 \cdot CDc20(t) \cdot CYCA(t)$
10	$V_{10} = 10 \cdot K30 \cdot CDc20(t) \cdot CA(t)$
11	$V_{11} = 10 \cdot K25R \cdot CE(t)$

12	$V_{12} = 10 \cdot K25R \cdot CA(t)$
13	$V_{13} = 10 \cdot V8 \cdot CE(t)$
14	$V_{14} = 10 \cdot V8 \cdot CYCE(t)$
15	$V_{15} = 10 \cdot V6 \cdot P27(t)$
16	$V_{16} = 10 \cdot V6 \cdot CE(t)$
17	$V_{17} = 10 \cdot V6 \cdot CD(t)$
18	$V_{18} = 10 \cdot V6 \cdot CA(t)$
19	$V_{19} = 10 \cdot V2 \cdot CYCB(t)$
20	$V_{20} = 10 \cdot \frac{(K3a + K3 \cdot CDc20(t)) \cdot (1 - CDh1(t))}{J3 - CDh1(t) + 1}$
21	$V_{21} = 10 \cdot \frac{V4 \cdot CDh1(t)}{J4 + CDh1(t)}$
22	$V_{22} = 10 \cdot K34 \cdot PPX(t)$
23	$V_{23} = 10 \cdot \frac{K31 \cdot CYCB(t) \cdot (1 - IEP(t))}{J31 - IEP(t) + 1}$
24	$V_{24} = 10 \cdot \frac{K32 \cdot PPX(t) \cdot IEP(t)}{J32 + IEP(t)}$
25	$V_{25} = 10 \cdot K12 \cdot CDc20T(t)$
26	$V_{26} = 10 \cdot \frac{K13 \cdot IEP(t) \cdot (CDc20T(t) - CDc20(t))}{J13 - CDc20(t) + CDc20T}$
27	$V_{27} = 10 \cdot \frac{K14 \cdot CDc20(t)}{J14 + CDc20(t)}$
28	$V_{28} = 10 \cdot K12 \cdot CDc20(t)$
29	$V_{29} = 10 \cdot E2F_Rb(t) \cdot K20 \cdot (CYCDT(t) \cdot LD + CYCA(t) \cdot LA + CYCB(t) \cdot LB + CYCE(t) \cdot LE)$
30	$V_{30} = 10 \cdot p_E2F_Rb(t) \cdot K20 \cdot (CYCDT(t) \cdot LD + CYCA(t) \cdot LA + CYCB(t) \cdot LB + CYCE(t) \cdot LE)$
31	$V_{31} = 10 \cdot K27 \cdot MASS \cdot If \left[\frac{Rb(t) + E2F_Rb(t) + p_E2F_Rb(t)}{pp_Rb(t) + Rb(t) + E2F_Rb(t) + p_E2F_Rb(t)} > 0.8; 0, 1 \right]$

32	$V_{32} = 10 \cdot K28 \cdot GM(t)$
33	$V_{33} = 10 \cdot \frac{K_{33} \cdot GLUCOSE(t)}{K_m + GLUCOSE(t)}$
34	$V_{34} = 10 \cdot \frac{\varepsilon \cdot k15}{1 + (DRG(t)/J15)^2}$
35	$V_{35} = 10 \cdot \varepsilon \cdot (K11a + K11 \cdot CYCB(t))$
36	$V_{36} = 10 \cdot \varepsilon \cdot K29 \cdot E2F(t) \cdot MASS(t)$
37	$V_{37} = 10 \cdot \varepsilon \cdot K33$
38	$V_{38} = 10 \cdot \varepsilon \cdot (K7a + K7 \cdot E2F(t))$
39	$V_{39} = 10 \cdot \varepsilon \cdot K9 \cdot DRG(t)$
40	$V_{40} = 10 \cdot \varepsilon \cdot K5$
41	$V_{41} = 10 \cdot \varepsilon \cdot (k17a \cdot ERG + \frac{k17 \cdot (DRG(t)/J17)^2}{1 + (DRG(t)/J17)^2})$
42	$V_{42} = 10 \cdot \varepsilon \cdot (K1a + \frac{K1 \cdot (CYCB(t)/J1)^2}{1 + (CYCB(t)/J1)^2})$
43	$V_{43} = 10 \cdot Rb(t) \cdot K20 \cdot (CYCDT(t) \cdot LD + CYCA(t) \cdot LA + CYCB(t) \cdot LB + CYCE(t) \cdot LE)$
44	$V_{44} = 10 \cdot pp_Rb(t) \cdot (K19a \cdot (PP1T - PP1A) + K19 \cdot PP1A)$
45	$V_{45} = 10 \cdot E2F_Rb(t) \cdot K26R$
46	$V_{46} = 10 \cdot E2F(t) \cdot (K23a + K23 \cdot (CYCA(t) + CYCB(t)))$
47	$V_{47} = 10 \cdot p_E2F(t) \cdot K22$
48	$V_{48} = 10 \cdot E2F(t) \cdot Rb(t) \cdot K26$
49	$V_{49} = 10 \cdot p_E2F_Rb(t) \cdot K26R$
50	$V_{50} = 10 \cdot Rb(t) \cdot p_E2F(t) \cdot K26$
51	$V_{51} = 10 \cdot p_E2F_Rb(t) \cdot K22$
52	$V_{52} = 10 \cdot E2F_Rb(t) \cdot (K23a + K23 \cdot (CYCA(t) + CYCB(t)))$
53	$V_{53} = 10 \cdot \frac{K_{G6P1} \cdot G6P(t)}{KM_{G6P} + G6P(t)}$

54	$V_{54} = 10 \cdot \frac{K_{F6P1} \cdot F6P(t)}{KM_{F6P} + F6P(t)}$
55	$V_{55} = 10 \cdot \frac{K_{F6P2} \cdot F6P(t)}{KM_{F6P} + F6P(t)}$
56	$V_{56} = 10 \cdot \frac{K_{FDP1} \cdot FDP(t)}{KM_{FDP} + FDP(t)}$
57	$V_{57} = 10 \cdot \frac{K_{FDP2} \cdot FDP(t)}{KM_{FDP} + FDP(t)}$
58	$V_{58} = 10 \cdot \frac{K_{DHAP1} \cdot DHAP(t)}{KM_{DHAP} + DHAP(t)}$
59	$V_{59} = 10 \cdot \frac{K_{FDP3} \cdot FDP(t)}{KM_{FDP} + FDP(t)}$
60	$V_{60} = 10 \cdot \frac{K_{GAP1} \cdot GAP(t)}{KM_{GAP} + GAP(t)}$
61	$V_{61} = 10 \cdot \frac{K_{DHAP2} \cdot DHAP(t)}{KM_{DHAP} + DHAP(t)}$
62	$V_{62} = 10 \cdot \frac{K_{GAP2} \cdot GAP(t)}{KM_{GAP} + GAP(t)}$
63	$V_{63} = 10 \cdot \frac{K_{GAP3} \cdot GAP(t)}{KM_{GAP} + GAP(t)}$
64	$V_{64} = 10 \cdot \frac{K_{DPG1} \cdot DPG(t)}{KM_{DPG} + DPG(t)}$
65	$V_{65} = 10 \cdot \frac{K_{DPG2} \cdot DPG(t)}{KM_{DPG} + DPG(t)}$
66	$V_{66} = 10 \cdot \frac{K_{P3G1} \cdot P3G(t)}{KM_{P3G} + P3G(t)}$
67	$V_{67} = 10 \cdot \frac{K_{P3G2} \cdot P3G(t)}{KM_{P3G} + P3G(t)}$
68	$V_{68} = 10 \cdot \frac{K_{P2G1} \cdot P2G(t)}{KM_{P2G} + P2G(t)}$
69	$V_{69} = 10 \cdot \frac{K_{P2G2} \cdot P2G(t)}{KM_{P2G} + P2G(t)}$

70	$V_{70} = 10 \cdot \frac{K_{PEP1} \cdot PEP(t)}{KM_{PEP} + PEP(t)}$
71	$V_{71} = 10 \cdot \frac{K_{PEP2} \cdot PEP(t)}{KM_{PEP} + PEP(t)}$
72	$V_{72} = 10 \cdot \frac{K_{PYR1} \cdot PYRUVATE(t)}{KM_{PYR} + PYRUVATE(t)}$
73	$V_{73} = 10 \cdot \frac{K_{PYR2} \cdot PYRUVATE(t)}{KM_{PYR} + PYRUVATE(t)}$
74	$V_{74} = 10 \cdot \frac{K_{LAC1} \cdot LACTATE(t)}{KM_{LAC} + LACTATE(t)}$
75	$V_{75} = 10 \cdot \frac{K_{LAC2} \cdot LACTATE(t)}{KM_{LAC} + LACTATE(t)}$
76	$V_{76} = 10 \cdot \frac{K_{PYR3} \cdot PYRUVATE(t)}{KM_{PYR} + PYRUVATE(t)}$
77	$V_{77} = 10 \cdot \frac{K_{AA1} \cdot AA(t)}{KM_{AA} + AA(t)}$
78	$V_{78} = 10 \cdot \frac{K_{AA2} \cdot AA(t) \cdot \varepsilon \cdot GM(t)}{KM_{AA} + AA(t)}$
79	$V_{79} = 10 \cdot \frac{K_{P3G3} \cdot P3G(t)}{KM_{P3G} + P3G(t)}$

The Cell Cycle And Metabolism Model ODEs:

$$AA'(t) = V_{76} + V_{79} - V_{77} - V_{78}$$

$$CA'(t) = V_6 - V_{10} - V_{12} - V_{18}$$

$$CD'(t) = V_7 - V_8 - V_{17} - V_3$$

$$CDc20'(t) = V_{26} - V_{27} - V_{28}$$

$$CDc20T'(t) = V_{35} - V_{25}$$

$$CDhI'(t) = V_{20} - V_{21}$$

$$CE'(t) = V_5 - V_{11} - V_{13} - V_{16}$$

$$CYCA'(t) = V_{36} - V_9 - V_6 + V_{12} + V_{18}$$

$$CYCB'(t) = V_{42} - V_{19}$$

$$CYCD'(t) = V_{39} + V_{17} + V_8 - V_7 - V_4$$

$$CYCE'(t) = V_{38} - V_{14} - V_5 + V_{11} + V_{16}$$

$$DHAP'(t) = V_{57} + V_{62} - V_{58} - V_{61}$$

$$DPG'(t) = V_{63} + V_{65} - V_{64} - V_{66}$$

$$DRG'(t) = V_{41} - V_2$$

$$E2F'(t) = V_{29} + V_{45} + V_{47} - V_{46} - V_{48}$$

$$E2F_Rb'(t) = V_{51} + V_{48} - V_{52} - V_{29} - V_{45}$$

$$ERG'(t) = V_{34} - V_1$$

$$F6P'(t) = V_{53} + V_{56} - V_{54} - V_{55}$$

$$FDP'(t) = V_{55} + V_{60} - V_{56} - V_{59}$$

$$G6P'(t) = V_{33} + V_{54} - V_{53}$$

$$GAP'(t) = V_{59} + V_{64} - V_{60} - V_{63}$$

$$GM'(t) = V_{31} - V_{32}$$

$$IEP'(t) = V_{23} - V_{24}$$

$$LACTATE'(t) = V_{73} - V_{74} - V_{75}$$

$$MASS'(t) = V_{78}$$

$$P27'(t) = V_{40} + V_3 + V_8 - V_{15} - V_5 - V_6 - V_7 + V_{11} + V_{12} + V_{13} + V_{10}$$

$$P2G'(t) = V_{67} + V_{70} - V_{68} - V_{69}$$

$$P3G'(t) = V_{65} + V_{68} - V_{66} - V_{67} - V_{79}$$

$$p_E2F'(t) = V_{30} + V_{49} + V_{46} - V_{47} - V_{50}$$

$$p_E2F_Rb'(t) = V_{52} + V_{50} - V_{51} - V_{30} - V_{49}$$

$$PEP'(t) = V_{69} + V_{72} - V_{70} - V_{71}$$

$$pp_Rb'(t) = V_{29} + V_{30} + V_{43} - V_{44}$$

$$PPX'(t) = V_{37} - V_{22}$$

$$PYRUVATE'(t) = V_{71} + V_{74} + V_{77} - V_{73} - V_{76}$$

$$Rb'(t) = V_{44} + V_{45} + V_{49} - V_{48} - V_{50} - V_{43}$$

The Cell Cycle And Metabolism Model Definitions And Steady-state Relations:

$$PP1A = \frac{PP1T}{1 + K21 \cdot (FE \cdot (CYCA(t) + CYCE(t)) + FB \cdot CYCB(t))}$$

$$V2 = K2aa \cdot CDc20(t) + K2a \cdot (1 - CDh1(t)) + K2 \cdot CDh1(t)$$

$$V4 = K4 \cdot (GA \cdot CYCA(t) + GB \cdot CYCB(t) + GE \cdot CYCE(t))$$

$$V6 = K6a + K6 \cdot (HA \cdot CYCA(t) + HB \cdot CYCB(t) + HE \cdot CYCE(t))$$

$$V8 = \frac{(YE \cdot (CYCA(t) + CYCE(t)) + YB \cdot CYCB(t)) \cdot K8}{CYCET + J8} + K8a$$

$$CYCET = CE(t) + CYCE(t)$$

$$CYCDT = CD(t) + CYCD(t)$$

$$CYCAT = CA(t) + CYCA(t)$$

$$P27T = CA(t) + CD(t) + CE(t) + P27(t)$$

$$\varepsilon(t) = 1$$

The Mammalian Cell Cycle And Metabolism Model Parameters:

k15 = 0.025
k16 = 0.025
J15 = 0.1
k17a = 0.035
k17 = 1.
J17 = 0.3
k18 = 1.
K9 = 0.25
K10 = 0.5
k24 = 100.
k24r = 1.
K7a = 0.
K7 = 0.06
K8a = 0.01
K8 = 0.2
K25 = 100.
K25R = 1.
J8 = 0.1
YE = 1.
YB = 0.05
K29 = 0.005
K30 = 2.
K1a = 0.01
K1 = 0.06
J1 = 0.1
K2a = 0.005
K2 = 2.
K2aa = 0.1
K5 = 2.
K6a = 1.
K6 = 10.
HE = 0.5
HB = 1.
HA = 0.5
LD = 3.3
LE = 5.
LB = 5.
LA = 3.
K20 = 1.
K19a = 0.
K19 = 2.
K21 = 1.
PP1T = 1.
FE = 25.

FB = 2.
K3a = 0.75
K3 = 14.
J3 = 0.01
J4 = 0.01
K4 = 4.
GE = 0.
GB = 1.
GA = 0.3
K33 = 0.005
K34 = 0.005
K31 = 0.07
K32 = 0.18
J31 = 0.01
J32 = 0.01
K11a = 0.
K11 = 0.15
K12 = 0.15
K13 = 0.5
K14 = 0.25
J13 = 0.005
J14 = 0.005
K22 = 0.1
K23a = 0.0005
K23 = 0.1
K26 = 1000.
K26R = 20.
K27 = 0.02
K28 = 0.02
K₃₃ = 0.0047
K_{GLU} = 0.00012
K_{PYR1} = 1.
K_{AA1} = 1.
K_{AA2} = 1.
K_{PYR2} = 1.
K_{PYR3} = 1.
K_{G6P1} = 1.
K_{F6P1} = 1.
K_{F6P2} = 1.
K_{FDP1} = 1.
K_{DHAP1} = 1.
K_{DHAP2} = 1
K_{FDP2} = 1.
K_{GAP1} = 1.
K_{GAP2} = 1.
K_{GAP3} = 1.
K_{DPG1} = 1.
K_{DPG2} = 1.
K_{P3G1} = 1.
K_{P3G2} = 1.

$K_{P2G1} = 1.$
 $K_{P2G2} = 1.$
 $K_{PEP1} = 1.$
 $K_{PEP2} = 1.$
 $K_{LAC1} = 1.$
 $K_{FDP3} = 1.$
 $K_{LAC2} = 1.$
 $K_{P3G3} = 1.$
 $KM_{G6P} = 1.$
 $KM_{F6P} = 1.$
 $KM_{FDP} = 1.$
 $KM_{DHAP} = 1.$
 $KM_{GAP} = 1.$
 $KM_{DPG} = 1.$
 $KM_{P3G} = 1.$
 $KM_{P2G} = 1.$
 $KM_{PEP} = 1.$
 $KM_{PYR} = 1.$
 $KM_{LAC} = 1.$
 $KM_{AA} = 1.$

Initial Concentrations For The Mammalian Cell Cycle And Metabolism Model:

$$AA(0) = 0.001$$

$$CA(0) = 0.0356927$$

$$CD(0) = 0.010976$$

$$Cdc20(0) = 0.00220177$$

$$Cdc20T(0) = 2.36733$$

$$CDh1(0) = 0.000653278$$

$$CE(0) = 0.000542587$$

$$CYCA(0) = 1.4094$$

$$CYCB(0) = 2.72898$$

$$CYCD(0) = 0.43929$$

$$CYCE(0) = 0.0229112$$

DHAP(0) = 0.001
DPG(0) = 0.001
DRG(0) = 0.900533
ERG(0) = 0.0121809
F6P(0) = 0.001
FDP(0) = 0.001
G6P(0) = 0.001
GAP(0) = 0.001
GLU(0) = 0.001
GM(0) = 1.35565
IEP(0) = 0.154655
LAC(0) = 0.001
MASS(0) = 1.687759919
P27(0) = 0.00922806
P2G(0) = 0.001
P3G(0) = 0.001
PEP(0) = 0.001
PPX(0) = 1
pp-RB(0) = 9.97574
E2F(0) = 0.989986
p-E2F(0) = 3.98594
PYR(0) = 0.001
Rb(0) = 0.000190871
E2F-Rb(0) = 0.00478911
p-E2F-Rb(0) = 0.0192822

APPENDIX

C

A CANCER MODEL OF CELL GROWTH AND
DIVISION – SUPPLEMENTARY INFORMATION

Gene Expression Data For Non-Essential Amino Acids:

Table C.1. Gene expression data for the non-essential amino acids (non-essential amino acids originally synthesised from pyruvate highlighted in blue, non-essential amino acids originally synthesised from P3G highlighted in red).

Reaction	Probe Set	Gene Title	Normal Breast (N) Expression Intensity (RMA Values)	MDA-MB-231 (C)	C/N	Mean C/N Value
Pyruvate ↔ Acetyl-CoA	212568_s_at	dihydrolipoamide S-acetyltransferase	7.03334	8.99998	1.280	
	209095_at	dihydrolipoamide dehydrogenase	8.66898	10.18015	1.174	
	200980_s_at	pyruvate dehydrogenase (lipoamide) alpha 1	9.3186	10.07881	1.082	
	214518_at	pyruvate dehydrogenase (lipoamide) alpha 2	6.75406	6.63949	0.983	1.024
	211023_at	pyruvate dehydrogenase (lipoamide) beta	9.33612	10.35314	1.109	
Pyruvate ↔ Oxaloacetate	204476_s_at	pyruvate carboxylase	8.71777	7.83095	0.898	
Acetyl-CoA ↔ Citrate	208660_at	citrate synthase	9.62448	10.12605	1.052	

Oxaloacetate ↔ Citrate	210337_s_at	ATP citrate lyase	8.67534	10.02305	1.155
Citrate ↔ 2-Oxoglutarate	201193_at	isocitrate dehydrogenase 1 (NADP+), soluble	9.59213	9.11944	0.951
	210046_s_at	isocitrate dehydrogenase 2 (NADP+), mitochondrial	9.24751	7.45937	0.807
	202069_s_at	isocitrate dehydrogenase 3 (NAD+) alpha	6.33698	8.41867	1.328
	210418_s_at	isocitrate dehydrogenase 3 (NAD+) beta	9.75743	9.90988	1.016
	202471_s_at	isocitrate dehydrogenase 3 (NAD+) gamma	8.46486	8.87557	1.049
	207071_s_at	aconitase 1, soluble	8.80744	9.72109	1.104
	200793_s_at	aconitase 2, mitochondrial	9.29468	8.4346	0.907
Oxaloacetate ↔ Malate	200978_at	malate dehydrogenase 1, NAD (soluble)	10.53948	10.84643	1.029
	209036_s_at	malate dehydrogenase 2, NAD (mitochondrial)	9.89076	10.24875	1.036
Malate ↔ 2-Oxoglutarate	201093_x_at	succinate dehydrogenase complex, subunit A, flavoprotein (Fp)	10.15333	9.77263	0.963
	202675_at	succinate dehydrogenase complex, subunit B, iron sulfur (Ip)	7.66838	8.40596	1.096
	210131_x_at	succinate dehydrogenase complex, subunit C, integral membrane protein, 15kDa	9.58464	9.71337	1.013
	202026_at	succinate dehydrogenase complex, subunit D, integral membrane protein	8.51777	9.81056	1.152
	203033_x_at	fumarate hydratase	8.59815	9.81297	1.141
	217874_at	succinate-CoA ligase, alpha subunit	10.07123	9.81537	0.975
	212459_x_at	succinate-CoA ligase, GDP-forming, beta subunit	8.89277	9.0229	1.015
	202930_s_at	succinate-CoA ligase, ADP-forming, beta subunit	7.63021	7.52115	0.986
	215210_s_at	dihydrolipoamide S-succinyltransferase (E2 component of 2-oxoglutarate complex)	9.29238	9.10458	0.980

	201282_at	oxoglutarate (alpha-ketoglutarate) dehydrogenase (lipoamide)	8.02853	8.90386	1.109
	219277_s_at	oxoglutarate dehydrogenase-like	7.30291	6.28699	0.861
	209095_at	dihydrolipoamide dehydrogenase	8.66898	10.18015	1.174
2-Oxoglutarate ↔ Glutamate	200946_x_at	glutamate dehydrogenase 1	9.50983	9.47876	0.997
	210447_at	glutamate dehydrogenase 2	5.19602	3.91696	0.754
Glutamate ↔ Glutamine	217202_s_at	glutamate-ammonia ligase	9.55751	7.18976	0.752
	221510_s_at	glutaminase	7.54051	8.58004	1.138
	205531_s_at	glutaminase 2 (liver, mitochondrial)	6.35534	5.23404	0.824
Glutamate ↔ Glutamate-5-semialdehyde	211552_s_at	aldehyde dehydrogenase 4 family, member A1	7.49449	6.95062	0.927
	217791_s_at	aldehyde dehydrogenase 18 family, member A1	9.33656	9.87035	1.057
Glutamate-5-semialdehyde ↔ Proline	206259_at	protein C (inactivator of coagulation factors Va and VIIIa)	6.38476	5.61183	0.879
Pyruvate ↔ Alanine	206709_x_at	glutamic-pyruvate transaminase (alanine aminotransferase)	7.16144	6.2815	0.877
	208284_x_at	gamma-glutamyltransferase 1	7.90119	6.30516	0.798
Oxaloacetate ↔ Aspartate	208813_at	glutamic-oxaloacetic transaminase 1, soluble (aspartate aminotransferase 1)	7.52059	9.51707	1.265
Aspartate ↔ Asparagine	205047_s_at	asparagine synthetase (glutamine-hydrolyzing)	7.87405	10.29289	1.307
3-Phosphoglycerate ↔ 3-Phosphohydroxypyruvate	201397_at	phosphoglycerate dehydrogenase	8.51304	6.72785	0.790
3-Phosphohydroxypyruvate ↔ 3-Phosphoserine	220892_s_at	phosphoserine aminotransferase 1	6.43545	10.33315	1.606
3-Phosphoserine ↔ Serine	205194_at	phosphoserine phosphatase	5.1804	7.50377	1.448
Serine ↔ Glycine	214096_s_at	serine hydroxymethyltransferase 2 (mitochondrial)	9.81703	10.35204	1.054
Serine ↔ Cysteine	212816_s_at	cystathionine-beta-synthase	5.52769	4.73133	0.856
	217127_at	cystathionase (cystathionine gamma-lyase)	5.57663	8.36191	1.499

1.209

The Mammalian Cell Cycle And Metabolism Model Parameters For Cancer Cells:

k15 = 0.025
k16 = 0.025
J15 = 0.1
k17a = 0.035
k17 = 1.
J17 = 0.3
k18 = 1.
K9 = 0.25
K10 = 0.5
k24 = 100.
k24r = 1.
K7a = 0.
K7 = 0.06
K8a = 0.01
K8 = 0.2
K25 = 100.
K25R = 1.
J8 = 0.1
YE = 1.
YB = 0.05
K29 = 0.005
K30 = 2.
K1a = 0.01
K1 = 0.06
J1 = 0.1
K2a = 0.005
K2 = 2.
K2aa = 0.1
K5 = 2.
K6a = 1.
K6 = 10.
HE = 0.5
HB = 1.
HA = 0.5
LD = 3.3
LE = 5.
LB = 5.
LA = 3.
K20 = 1.
K19a = 0.
K19 = 2.
K21 = 1.
PP1T = 1.

FE = 25.
FB = 2.
K3a = 0.75
K3 = 14.
J3 = 0.01
J4 = 0.01
K4 = 4.
GE = 0.
GB = 1.
GA = 0.3
K33 = 0.005
K34 = 0.005
K31 = 0.07
K32 = 0.18
J31 = 0.01
J32 = 0.01
K11a = 0.
K11 = 0.15
K12 = 0.15
K13 = 0.5
K14 = 0.25
J13 = 0.005
J14 = 0.005
K22 = 0.1
K23a = 0.0005
K23 = 0.1
K26 = 1000.
K26R = 20.
K27 = 0.02
K28 = 0.02
K₃₃ = 0.0065
KM_{GLU} = 0.00012
K_{PYR1} = 9.26
K_{AA1} = 1.024
K_{AA2} = 1.
K_{PYR2} = 3.71
K_{PYR3} = 1.024
K_{G6P1} = 1.85
K_{F6P1} = 1.85
K_{F6P2} = 3.41
K_{FDP1} = 3.41
K_{DHAP1} = 1.66
K_{DHAP2} = 3.84
K_{FDP2} = 1.66
K_{GAP1} = 1.66
K_{GAP2} = 3.84
K_{GAP3} = 2.79
K_{DPG1} = 2.79
K_{DPG2} = 7.02
K_{P3G1} = 7.02

$K_{P3G2} = 1.81$
 $K_{P2G1} = 1.81$
 $K_{P2G2} = 4.35$
 $K_{PEP1} = 4.35$
 $K_{PEP2} = 9.26$
 $K_{LAC1} = 3.71$
 $K_{FDP3} = 1.66$
 $K_{LAC2} = 1.$
 $K_{P3G3} = 1.209$
 $KM_{G6P} = 1.$
 $KM_{F6P} = 1.$
 $KM_{FDP} = 1.$
 $KM_{DHAP} = 1.$
 $KM_{GAP} = 1.$
 $KM_{DPG} = 1.$
 $KM_{P3G} = 1.$
 $KM_{P2G} = 1.$
 $KM_{PEP} = 1.$
 $KM_{PYR} = 1.$
 $KM_{LAC} = 1.$
 $KM_{AA} = 1.$