Computer Simulation of a Large-Scale Signalling Network Regulating Translation Initiation in the Mammalian Cell

By

David James Taylor

Submitted in Partial Fulfilment of the Requirements for the Degree of

Doctor of Philosophy

January 2017

School of Biosciences & Medicine
Faculty of Health & Medical Sciences
University of Surrey
Declaration of Originality

I confirm that the submitted work is my own work and that I have clearly identified and fully acknowledged all material that is entitled to be attributed to others (whether published or unpublished) using the referencing system set out in the programme handbook. I agree that the University may submit my work to means of checking this, such as the plagiarism detection service Turnitin® UK. I confirm that I understand that assessed work that has been shown to have been plagiarised will be penalised.
Abstract

The process of translation initiation in mammalian systems is complex and not fully understood. It is regulated by an intricate network of signalling pathways and is a significant energetic burden to the cell. Although models of initiation are available for yeast, to date, such models do not include the regulation of this process, nor do they exist for mammalian systems. Existing literature was used to reconstruct the process of translation initiation and the regulatory signalling networks in the Petri Net formalism within the software Snoopy. The final version of the model was altered to incorporate the effects of Murine Norovirus. The model was converted to a binary form and the software QSSPN was used to run Gillespie algorithm-based stochastic simulations. The predictive power of the model was established by incorporating commonly used chemical inhibitors. Using the Matthews’ Correlation Coefficient, a quantitative measure of predictive power was established by comparing the model behaviour to the effects of each inhibitor recorded in existing literature. A qualitative model containing 584 reactions was constructed. The predictive power of the model was raised to MCC = 0.4558 through a series of refinements. Two predicted behaviours, an increase in eIF4E phosphorylation and a reduction of AKT phosphorylation both in response to Rapamycin, were validated with the Immunoblotting techniques, Western Blotting and Human Phospho-MAPK Arrays, in the murine monocyte/macrophage RAW 264.7 cell line. The model incorporating the effects of Murine Norovirus infection generated five testable predictions. Of these, four were verified with the Human Phospho-MAPK Arrays. The model presented here demonstrates the value of generating large-scale models using the binary model formalism and performing simulations with QSSPN. The model of the regulation of translation initiation has shown that it is capable of generating experimentally verifiable predictions. Furthermore, the incorporation of viral effects demonstrates that the model has a range of potential future uses.
Acknowledgements

For funding my work, I need to thank the University of Surrey.

I would like to thank my supervisors, Prof Andrzej Kierzek, Dr Nicolas Locker and Prof Lisa Roberts for their help and support over the course of this project. This work would not have been possible without their guidance and patience over the last three years. It has been a pleasure to work with all of you.

I also need to thank everyone who has helped me over the last three years (Dr Elizabeth Royall, Dr Nicole Doyle, Dr Margaret Carter, Dr Belinda Hall, Salmah Zaini, Majid Al-Sailawi, Dr Melvin Leteane, Dr Nargs Alfaqih, Mariam Sulaiman, Nefeli Karataraki, Joy Ogbechi, Dr Adeola Fowotade, Katja Ebert-Keel, Dr Valerie Odon, Dr Manuel Salvador de Lara, Joe Holley, Dr Juliet Jones, Dr Huihai Wu and Dr Mehdi Khoury to name but a few). Your friendship and support has made the last three years so much easier.

I must also thank Dr Ben Neuman (University of Reading) for inspiring me to pursue a career in molecular biology. Secondly, thank you to Dr Ditte Hobbs (University of Reading) for her friendship and help during my undergraduate degree. Additional thanks must go to Dr Alessia Ruggieri (Universität Heidelberg) for a very thought provoking conversation that made me re-evaluate some of my positions on the experimental work undertaken in this project.

I must thank my two best friends, Matthew Fletcher and Anthony Kendall, for all of the good times we have had! You guys have always managed to help me make the best of even the worst situation.
My mum, Patricia, has always supported and encouraged me. Thank you for never doubting me. You have always been there for me and I can never thank you enough. Although they are no longer here, I must thank my Nan, Jean, and my Granddad, Joe. I miss you both very much and hope that you would both be proud of me. The three of you made me into the person I am today and I owe everything I achieve to you all.
# Table of Contents

Declaration of Originality  i
Abstract  ii
Acknowledgements  iii
Table of Contents  v
List of Figures  xiii
List of Tables  xvii
List of Equations  xviii
List of Abbreviations  xx
List of Models & Variants  xxviii

Chapter 1 – Introduction

Section 1.1 – Cellular Translation Initiation  2
Section 1.1.1 – Post Termination Ribosomal Complex Recycling  4
Section 1.1.2 – Ternary Complex Formation  5
Section 1.1.3 – eIF4F Components & mRNA Circularisation  6
Section 1.1.4 – Start Codon Recognition  8
Section 1.1.5 – 60S Subunit Joining  9

Section 1.2 – Key Points of Regulation & Upstream Signalling Pathways  10
Section 1.2.1 – eIF4E-Binding Proteins & The mTORC1 Signalling Pathway  12
Section 1.2.2 – eIF4E Phosphorylation and The Role of ERK1/2 & p38  20
Section 1.2.3 – Ribosomal Protein S6 Kinases  27
Section 1.2.4 – Phosphorylation of Eukaryotic Initiation Factor 2α  30
Section 1.2.5 – Emerging Roles for JNK1 in the Regulation of Translation Initiation

Section 1.3 – Overview of the Viruses Modelled in this Project and Viral Strategies Used to Alter Host Cell Translation Initiation

Section 1.3.1 – General Features & Epidemiology of Viruses Incorporated in the Model

Section 1.3.1.1 – Arenaviruses & Junin Virus

Section 1.3.1.2 – Caliciviruses & Murine Norovirus

Section 1.3.1.3 – Hantaviruses & Andes Virus

Section 1.3.2 – General Overview of Viral Effects on Translational Regulation

Section 1.3.2.1 – Internal Ribosome Entry Site (IRES)-Mediated Translation & Host mRNA Degradation

Section 1.3.2.2 - VPg Proteins

Section 1.3.2.3 – Host Factor Degradation & Viral Mimics of Host Factors

Section 1.3.2.4 – Alteration to Cell Signalling Pathways

Section 1.4 – Background to Systems Biology Techniques

Section 1.4.1 – Exact Stochastic Simulations with the Gillespie Algorithm

Section 1.4.2 – Ordinary Differential Equations (ODEs)

Section 1.4.3 – Flux Balance Analysis

Section 1.4.4 – The Petri Net Formalism

Section 1.4.5 – The Quasi-Steady State Petri Net Algorithm

Section 1.4.6 – Translation of Petri Net Terminology for a Molecular Biologist

Section 1.5 – Aims & Objectives and Novel Contributions to Knowledge

Section 1.5.1 – Aims & Objectives
Section 1.5.2 – Novel Contributions to Knowledge

Chapter 2 – Model Creation & Refining

Section 2.1 – Materials & Methods

Section 2.1.1 – Qualitative Petri Nets
Section 2.1.2 – The Binary Model Formalism
Section 2.1.3 – Model Construction
Section 2.1.3.1 – General Model Construction Process
Section 2.1.3.2 – Specific Examples
Section 2.1.3.3 – Creation of the Benchmark Datasets
Section 2.1.4 – Simulation Method
Section 2.1.4.1 – Monte Carlo Simulations
Section 2.1.4.2 – Evaluation of the Predictive Power of the Model
Section 2.1.4.3 – Synchronous Simulation Methods
Section 2.1.4.3.1 – Case Study Identifying Differences in the Asynchronous Simulation Approach & the Synchronous Methods
Section 2.1.5 – Investigation of the Interaction Between eIF4E Phosphorylation and mTORC1 Inhibition

Section 2.2 – Results

Section 2.2.1 – Binary Model of LPS Stimulation
Section 2.2.1.1 – Updated Benchmark Dataset
Section 2.2.2 – Re-Evaluation of the Binary Model
Section 2.2.3 – Mechanism by Which mTORC1 Inhibition Affects eIF4E Phosphorylation
Section 2.2.4 – Evaluation of the Predictive Power of the Final Model

Section 2.3 – Discussion

Section 2.3.1 – Mechanism of Increasing eIF4E Phosphorylation in Response to mTORC1 Inhibition

Section 2.3.2 – Predictive Power of the Final Version of the Model

Section 2.3.3 – Problems Encountered & Troubleshooting

Section 2.3.4 – Concluding Remarks

Chapter 3 – Experimental Validation of Model Predictions

Section 3.1 – Introduction

Section 3.1.1 – Rapamycin

Section 3.1.2 – eIF4E Phosphorylation During mTORC1 Inhibition

Section 3.1.3 – Interplay Between mTORC1 & AKT

Section 3.2 – Materials & Methods

Section 3.2.1 – Growth & Passage of RAW 264.7 Murine Macrophage Cells

Section 3.2.2 – Determination of Rapamycin Cytotoxicity

Section 3.2.3 – Cell Harvesting & Bicinchoninic Acid (BCA) Assay

Section 3.2.4 – Human Phospho-MAPK Arrays

Section 3.2.5 – eIF4E/Phosphorylated eIF4E & AKT/Phosphorylated AKT Western Blot Analysis

Section 3.2.6 – Re-Evaluation of the Model Based Upon the Results of the Human Phospho-MAPK Arrays

Section 3.2.7 – Statistical Analysis
Section 3.3 – Results

Section 3.3.1 – Cytotoxicity of Rapamycin

Section 3.3.2 – Evaluation of the Human Phospho-MAPK Arrays

Section 3.3.3 – Alterations to eIF4E Phosphorylation Levels

Section 3.3.4 – Confirmation of the Effect Rapamycin has Upon AKT Phosphorylation

Section 3.3.5 – Comparison of Model Results to Array Data

Section 3.4 – Discussion

Section 3.4.1 – Validation of the eIF4E Phosphorylation Prediction

Section 3.4.2 – Validation of the AKT Phosphorylation Prediction

Section 3.4.3 – Problems Encountered & Troubleshooting

Section 3.4.4 – Potential Future Work

Section 3.4.5 – Concluding Remarks

Chapter 4 – Modelling the Effects of Viral Infection

Section 4.1 – Introduction

Section 4.2 – Model Changes & Simulation Methods

Section 4.2.1 – Junin Virus Model

Section 4.2.2 – Murine Norovirus Model

Section 4.2.3 – Andes Virus Model

Section 4.2.4 – Experimental Validation of the Murine Norovirus Model

Section 4.3 – Results

Section 4.3.1 – Simulation of the Junin Virus Models

Section 4.3.2 – Simulation of the Murine Norovirus Models
Chapter 6 – Conclusion

Section 6.1 – Introduction 248

Section 6.2 – Model Construction & Predictive Power Evaluation 250

Section 6.3 – Experimental Validation 254

Section 6.4 – Mechanism for Rapamycin-Induced Increases in eIF4E Phosphorylation 256

Section 6.5 – Models of Viral Infection 257

Section 6.6 – Integration of the Translational Regulation Model with a GSMN 261

Section 6.7 – Novel Modelling Strategies 262

Section 6.8 – General Conclusion 264

References 265

Appendices 349

Supplementary Information 1 - References for the Construction of the Literature 353

Datasets

Supplementary Information 1a - Benchmark Dataset References 354

Supplementary Information 1b - Benchmark Dataset References for the Binary Model with Lipopolysaccharide (LPS Benchmark) 383

Supplementary Information 2 - Dataset against Which the Binary Model was Tested (Final Benchmark Dataset) 396

Supplementary Information 3 - Dataset Against Which the Model Integrated with Metabolism was Tested (Metabolic Benchmark Dataset) 411

Supplementary Information 4 – Ingredients List for Solutions Produced In-House 415
The Files Below are Saved to the USB Drive Enclosed

Supplementary Information E1 – Standardization of the Nomenclature of Model Components

Supplementary Information E2 - List of Model Transitions & References in the Binary Model

Supplementary Information E3 – Final Version of the Binary Model

Supplementary Information E4 – MNV Model

Supplementary Information E5 – LPS Benchmark Dataset

Supplementary Information E6 – Revised Original Benchmark Dataset

Supplementary Information E7 - Frequency Data Generated by Reachfq with Associated 99.99% Binomial Confidence Intervals
# List of Figures

## Chapter 1 – Introduction

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1a</td>
<td>Overview of Mammalian Translation Initiation</td>
<td>3</td>
</tr>
<tr>
<td>1.2a</td>
<td>Position of Translation Initiation within the Signalling Network</td>
<td>11</td>
</tr>
<tr>
<td>1.2.1a</td>
<td>Schematic of the mTORC1 Pathway &amp; the Downstream Components</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Regulating Translation Initiation</td>
<td></td>
</tr>
<tr>
<td>1.2.1b</td>
<td>Summary of the Role of AKT1 in the Activation of mTORC1</td>
<td>16</td>
</tr>
<tr>
<td>1.2.1c</td>
<td>Regulation of Translation Initiation by the Sequestration of eIF4F</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Components</td>
<td></td>
</tr>
<tr>
<td>1.2.2a</td>
<td>Signalling Cascade Resulting in the Activation of ERK1/2</td>
<td>22</td>
</tr>
<tr>
<td>1.2.2b</td>
<td>Effects of ERK1/2 Signalling on Translation Initiation</td>
<td>24</td>
</tr>
<tr>
<td>1.2.2c</td>
<td>The Activation of p38α &amp; The Effects of this Kinase on Translation and Gene Expression</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Protein Kinases</td>
<td></td>
</tr>
<tr>
<td>1.2.3a</td>
<td>Overview of Pathways Resulting in the Activation of the Ribosomal S6</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Phosphorylation of eIF2α</td>
<td></td>
</tr>
<tr>
<td>1.4.1a</td>
<td>Schematic View of Propensity Functions</td>
<td>47</td>
</tr>
<tr>
<td>1.4.4a</td>
<td>Example of a Simple Petri Net</td>
<td>57</td>
</tr>
<tr>
<td>1.4.4b</td>
<td>Extended Petri Net Notation</td>
<td>60</td>
</tr>
<tr>
<td>1.4.5a</td>
<td>Summation of the QSSPN Algorithm</td>
<td>70</td>
</tr>
</tbody>
</table>

## Chapter 2 – Model Creation & Refining

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1.2a</td>
<td>Overview of the Binary Model Formalism</td>
<td>82</td>
</tr>
<tr>
<td>2.1.3.1a</td>
<td>Petri Net Representation of Translation Initiation</td>
<td>88</td>
</tr>
<tr>
<td>2.1.3.1b</td>
<td>Petri Net Representation of the ERK1/2 Signalling Pathway</td>
<td>90</td>
</tr>
</tbody>
</table>
Figure 2.1.3.1c – Petri Net Representation of the p38α Signalling Pathway

Figure 2.1.3.1d – Petri Net Representation of the JNK1 Signalling Pathway

Figure 2.1.3.1e – Petri Net Representation of the PI3K and mTORC1 Signalling Pathways

Figure 2.1.3.2 – Example of Model Construction Process

Figure 2.1.4.3.1a - Schematic of the Example Model of eIF4E Phosphorylation and Availability

Figure 2.1.4.3.1b - Example Trajectory from the Simple Model Generated Using the Asynchronous Simulation Approach

Figure 2.1.4.3.1c - Example Trajectory from the Simple Model Generated Using the Synchronous Simulation Approach

Figure 2.2.1a – Example of the LPS Model Comparing the Effects of Inhibitors to a Wild Type Model

Figure 2.2.1b – Overview of the Predictive Power of the Model Incorporating the Effects of Lipopolysaccharide Signalling

Figure 2.2.2a – Overall Predictive Power of the Model Through the Four Stages of Refinement

Figure 2.2.2b – The Model Refinement Workflow

Figure 2.2.2c – Evaluation of the Predictive Power of the Fifth Iteration of the Model

Figure 2.2.4a – Confirmation of the Binary Nature of the Model Formalism

Figure 2.2.4b – Components of the Model Common to Human Phospho-MAPK Array

Figure 2.2.4c – Evaluation of the Predictive Power of the Final Version of the Model

Figure 2.2.4d – Stability Analysis of the Final Model
Chapter 3 – Experimental Validation of Model Predictions

Figure 3.1.3a – Overview of the Feedback Mechanism between S6K1 and the Insulin Receptor Substrate 1 (IRS-1) 163

Figure 3.1.3b – Overview of the Feedback Mechanism that Exists between S6K1 and mTORC2 165

Figure 3.3.1a – Evaluation of the Effects of Three Hour Treatments with Differing Concentrations of Rapamycin (Rapa) or 1% DMSO on the Murine Macrophage RAW 264.7 Cells 176

Figure 3.3.2a – Evaluation of the Broad Early Effects of 10μM Rapamycin Treatment to the Murine Macrophage RAW 264.7 Cells 178

Figure 3.3.3a – Rapamycin-Induced Increase in Phosphorylation of eIF4E. 182

Figure 3.3.4a – Effect of Rapamycin on the Levels of AKT1 Phosphorylation (Ser-473) 184

Figure 3.3.5a – Comparison of Array Data with the Behaviour of the Binary Model 186

Figure 3.4.1a – Hypothesis Linking mTORC1 to the MNK-Dependent Stress Response 190

Chapter 4 – Modelling the Effects of Viral Infection

Figure 4.3.1a – Modelling the Effects of JUNV Infection 210

Figure 4.3.2a – Modelling the Effects of MNV Infection 213

Figure 4.3.2.1a – Effect MNV Infection has Upon CREB Phosphorylation 216

Figure 4.3.3a - Modelling the Effects of ANDV Infection 218
Chapter 5 - Integration of the Translational Control Model with a Genome Scale Metabolic Network of a RAW 264.7 Cell

Figure 5.1.1a – Overview of the Role Played by S6K1 in Regulating Multiple Metabolic Pathways 231

Figure 5.3.1a – Predictive Power of the Model of Translational Regulation Integrated with a GSMN of RAW 264.7 Cell Metabolism and Simulated with the Current, Asynchronous Method 237

Figure 5.3.2.1a – Summary of the Results of the Ordinary Differential Equation Simulation Method 240

Figure 5.3.2.1b – Overview of the Predictive Power of the Model when Simulated with the Ordinary Differential Equation Method 242

Chapter 6 – Conclusion

Figure 6.5a – Schematic of the Effects of MNV-Induced ERK1/2 and p38 Activation 259
List of Tables

Chapter 2 – Model Creation & Refining

Table 2.1.3.3a - Chemical Inhibitors used in the Model 104

Chapter 3 – Experimental Validation of Model Predictions

Table 3.2.5a – Summary of the Antibody Solutions Used in Testing eIF4E and AKT Phosphorylation 172
# List of Equations

## Chapter 1 - Introduction

<table>
<thead>
<tr>
<th>Equation 1.4.1a</th>
<th>45</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equation 1.4.1b</td>
<td>45</td>
</tr>
<tr>
<td>Equation 1.4.1c</td>
<td>45</td>
</tr>
<tr>
<td>Equation 1.4.1d</td>
<td>48</td>
</tr>
<tr>
<td>Equation 1.4.1e</td>
<td>48</td>
</tr>
<tr>
<td>Equation 1.4.2a</td>
<td>50</td>
</tr>
<tr>
<td>Equation 1.4.3a</td>
<td>51</td>
</tr>
<tr>
<td>Equation 1.4.3b</td>
<td>52</td>
</tr>
<tr>
<td>Equation 1.4.3c</td>
<td>52</td>
</tr>
<tr>
<td>Equation 1.4.3d</td>
<td>52</td>
</tr>
<tr>
<td>Equation 1.4.4a</td>
<td>56</td>
</tr>
<tr>
<td>Equation 1.4.4b</td>
<td>56</td>
</tr>
<tr>
<td>Equation 1.4.4c</td>
<td>58</td>
</tr>
<tr>
<td>Equation 1.4.4d</td>
<td>58</td>
</tr>
<tr>
<td>Equation 1.4.4e</td>
<td>58</td>
</tr>
<tr>
<td>Equation 1.4.4f</td>
<td>58</td>
</tr>
<tr>
<td>Equation 1.4.4g</td>
<td>58</td>
</tr>
<tr>
<td>Equation 1.4.4h</td>
<td>58</td>
</tr>
<tr>
<td>Equation 1.4.4i</td>
<td>58</td>
</tr>
<tr>
<td>Equation 1.4.4j</td>
<td>58</td>
</tr>
<tr>
<td>Equation 1.4.4k</td>
<td>58</td>
</tr>
</tbody>
</table>
Chapter 2 – Model Creation & Refining

Equation 2.1.4.2a 109
Equation 2.1.4.2b 110
Equation 2.1.4.2c 110
List of Abbreviations

3' UTR 3' Untranslated Region
40S 40 Svedberg Ribosomal Subunit
43S 43 Svedberg Complex
48S 48 Svedberg Complex
4E-BP1 Eukaryotic Initiation Factor 4E-Binding Protein 1
4E-BP2 Eukaryotic Initiation Factor 4E-Binding Protein 2
4E-BP3 Eukaryotic Initiation Factor 4E-Binding Protein 3
4E-BPs Eukaryotic Initiation Factor 4E-Binding Proteins
60S 60 Svedberg Ribosomal Subunit
80S 80 Svedberg Ribosome
ABCE1 ATP-Binding Cassette 1
AC Approximate Coefficient
Acetyl-CoA Acetyl-Coenzyme A
ACP Average Conditional Probability
ADP Adenosine Diphosphate
AGC Protein Kinase A, Cyclic Guanosine Monophosphate-Dependent Protein Kinase and Protein Kinase C
AHF Argentine Haemorrhagic Fever
AKT1 Protein Kinase B
ANDV Andes Virus
ANOVA Analysis of Variance
ATF-2 Cyclic AMP-Dependent Transcription Factor
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP</td>
<td>Adenosine Trisphosphate</td>
</tr>
<tr>
<td>AUG</td>
<td>Adenosine-Uracil-Guanosine</td>
</tr>
<tr>
<td>BCA</td>
<td>Bicinchoninic Acid</td>
</tr>
<tr>
<td>CAD</td>
<td>Carbamoyl-Phosphate Synthetase 2, Aspartate Transcarbamylase and Dihydroorotase</td>
</tr>
<tr>
<td>Cas</td>
<td>CRISPR-Associated Protein</td>
</tr>
<tr>
<td>CREB</td>
<td>Cyclic AMP-Responsive Element-Binding Protein</td>
</tr>
<tr>
<td>CRISPR</td>
<td>Clustered, Regularly Interspaced, Short Palindromic Repeats</td>
</tr>
<tr>
<td>CUG</td>
<td>Cytosine-Uracil=Guanosine</td>
</tr>
<tr>
<td>DEAD</td>
<td>Asp-Glu-Ala-Asp</td>
</tr>
<tr>
<td>DEPTOR</td>
<td>DEP-Domain-Containing mTOR-Interacting Protein</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl Sulphoxide</td>
</tr>
<tr>
<td>DT</td>
<td>Dynamic Transitions</td>
</tr>
<tr>
<td>EBOV</td>
<td>Ebola Virus</td>
</tr>
<tr>
<td>EBV</td>
<td>Epstein-Barr Virus</td>
</tr>
<tr>
<td>eIF</td>
<td>Eukaryotic Initiation Factor</td>
</tr>
<tr>
<td>eIF1</td>
<td>Eukaryotic Initiation Factor 1</td>
</tr>
<tr>
<td>eIF1A</td>
<td>Eukaryotic Initiation Factor 1A</td>
</tr>
<tr>
<td>eIF2</td>
<td>Eukaryotic Initiation Factor 2</td>
</tr>
<tr>
<td>eIF2B</td>
<td>Eukaryotic Initiation Factor 2B</td>
</tr>
<tr>
<td>eIF2α</td>
<td>Eukaryotic Initiation Factor 2, Subunit α</td>
</tr>
<tr>
<td>eIF2β</td>
<td>Eukaryotic Initiation Factor 2, Subunit β</td>
</tr>
<tr>
<td>eIF2βγ</td>
<td>Eukaryotic Initiation Factor 2, Complex of Subunits β and γ</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>eIF3</td>
<td>Eukaryotic Initiation Factor 3</td>
</tr>
<tr>
<td>eIF4A</td>
<td>Eukaryotic Initiation Factor 4A</td>
</tr>
<tr>
<td>eIF4B</td>
<td>Eukaryotic Initiation Factor 4B</td>
</tr>
<tr>
<td>eIF4E</td>
<td>Eukaryotic Initiation Factor 4E</td>
</tr>
<tr>
<td>eIF4F</td>
<td>Eukaryotic Initiation Factor 4F</td>
</tr>
<tr>
<td>eIF4G</td>
<td>Eukaryotic Initiation Factor 4G</td>
</tr>
<tr>
<td>eIF4GI</td>
<td>Eukaryotic Initiation Factor 4G, Isoform I</td>
</tr>
<tr>
<td>eIF4GII</td>
<td>Eukaryotic Initiation Factor 4G, Isoform II</td>
</tr>
<tr>
<td>eIF4H</td>
<td>Eukaryotic Initiation Factor 4H</td>
</tr>
<tr>
<td>eIF5</td>
<td>Eukaryotic Initiation Factor 5</td>
</tr>
<tr>
<td>eIF6</td>
<td>Eukaryotic Initiation Factor 6</td>
</tr>
<tr>
<td>eRF1</td>
<td>Eukaryotic Release Factor 1</td>
</tr>
<tr>
<td>eRF3</td>
<td>Eukaryotic Release Factor 2</td>
</tr>
<tr>
<td>ERK1/2</td>
<td>Extracellular Signal-Regulated Kinase 1/2</td>
</tr>
<tr>
<td>ERK5</td>
<td>Extracellular Signal-Regulated Kinase 5</td>
</tr>
<tr>
<td>FBA</td>
<td>Flux Balance Analysis</td>
</tr>
<tr>
<td>FCV</td>
<td>Feline Calicivirus</td>
</tr>
<tr>
<td>FKBPs</td>
<td>FKS06-Binding Proteins</td>
</tr>
<tr>
<td>FKBP12</td>
<td>FKS06-Binding Protein 12</td>
</tr>
<tr>
<td>FKBPS1</td>
<td>FKS06-Binding Protein 51</td>
</tr>
<tr>
<td>FKBPS2</td>
<td>FKS06-Binding Protein 52</td>
</tr>
<tr>
<td>FN</td>
<td>False Negative</td>
</tr>
<tr>
<td>FP</td>
<td>False Positive</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>GAP</td>
<td>GTPase Activating Protein</td>
</tr>
<tr>
<td>GCN2</td>
<td>General Control Nonderepressing 2</td>
</tr>
<tr>
<td>GDP</td>
<td>Guanosine Diphosphate</td>
</tr>
<tr>
<td>GEF</td>
<td>Guanine Nucleotide Exchange Factor</td>
</tr>
<tr>
<td>GPC</td>
<td>Glycoprotein C-Terminus</td>
</tr>
<tr>
<td>GSK3</td>
<td>Glycogen Synthase Kinase 3</td>
</tr>
<tr>
<td>GSMN</td>
<td>Genome-Scale Metabolic Network</td>
</tr>
<tr>
<td>GTP</td>
<td>Guanosine Trisphosphate</td>
</tr>
<tr>
<td>HCMV</td>
<td>Human Cytomegalovirus</td>
</tr>
<tr>
<td>HCPS</td>
<td>Hantavirus Cardiopulmonary Syndrome</td>
</tr>
<tr>
<td>HIF-1α</td>
<td>Hypoxia Inducible Factor 1α</td>
</tr>
<tr>
<td>HIV-1</td>
<td>Human Immunodeficiency Virus Type 1</td>
</tr>
<tr>
<td>hnRNP A1</td>
<td>Heterogenous Nuclear Ribonucleoprotein A1</td>
</tr>
<tr>
<td>hpi</td>
<td>Hours Post Infection</td>
</tr>
<tr>
<td>HRI</td>
<td>Haem-Regulated eIF2α Kinase</td>
</tr>
<tr>
<td>HSV</td>
<td>Herpes Simplex Virus</td>
</tr>
<tr>
<td>HSV-1</td>
<td>Herpes Simplex Virus Type 1</td>
</tr>
<tr>
<td>HTLV-1</td>
<td>Human T-Cell Leukaemia Virus 1</td>
</tr>
<tr>
<td>HuNV</td>
<td>Human Norovirus</td>
</tr>
<tr>
<td>IHC</td>
<td>Immunohistochemistry</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal Ribosome Entry Site</td>
</tr>
<tr>
<td>IRS-1</td>
<td>Insulin Receptor Substrate 1</td>
</tr>
<tr>
<td>JNK</td>
<td>c-Jun N-Terminal Kinase</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>JUNV</td>
<td>Junin Virus</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilo Daltons</td>
</tr>
<tr>
<td>LPS</td>
<td>Lipopolysaccharide</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen-Activated Protein Kinase</td>
</tr>
<tr>
<td>MCC</td>
<td>Matthews' Correlation Coefficient</td>
</tr>
<tr>
<td>MEK</td>
<td>MAPK/ERK Kinase</td>
</tr>
<tr>
<td>MettRNA</td>
<td>Methionine-Aminoacylated Initiator tRNA</td>
</tr>
<tr>
<td>MK2</td>
<td>MAPK-Activated Protein Kinase-2</td>
</tr>
<tr>
<td>MKK3</td>
<td>Mitogen-Activated Protein Kinase Kinase 3</td>
</tr>
<tr>
<td>MKK6</td>
<td>Mitogen-Activated Protein Kinase Kinase 6</td>
</tr>
<tr>
<td>MKP-7</td>
<td>MAPK Phosphatase 7</td>
</tr>
<tr>
<td>mLST8</td>
<td>Mammalian Lethal with Sec13 Protein 8</td>
</tr>
<tr>
<td>MNK1</td>
<td>MAPK-Interacting Kinase 1</td>
</tr>
<tr>
<td>MNK2</td>
<td>MAPK-Interacting Kinase 2</td>
</tr>
<tr>
<td>MNV</td>
<td>Murine Norovirus</td>
</tr>
<tr>
<td>MOI</td>
<td>Multiplicity of Infection</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger Ribonucleic Acid</td>
</tr>
<tr>
<td>mSIN1</td>
<td>Mammalian Stress-Activated Protein Kinase Interacting Protein</td>
</tr>
<tr>
<td>MSK1</td>
<td>Mitogen- and Stress-Activated Kinase 1</td>
</tr>
<tr>
<td>MSK2</td>
<td>Ribosomal Protein S6 Kinase α-4</td>
</tr>
<tr>
<td>mTOR</td>
<td>Mammalian Target of Rapamycin</td>
</tr>
<tr>
<td>mTORC1</td>
<td>Mammalian Target of Rapamycin Complex 1</td>
</tr>
<tr>
<td>mTORC2</td>
<td>Mammalian Target of Rapamycin Complex 2</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>ns</td>
<td>Non-Significant</td>
</tr>
<tr>
<td>ODE</td>
<td>Ordinary Differential Equation</td>
</tr>
<tr>
<td>PABP</td>
<td>Poly(A) Binding Protein</td>
</tr>
<tr>
<td>PAM</td>
<td>Protoscaler Adjacent Motif</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-Buffered Saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PDCD4</td>
<td>Programed Cell Death 4</td>
</tr>
<tr>
<td>PDK1</td>
<td>3-Phosphoinositide-Dependent Protein Kinase-1</td>
</tr>
<tr>
<td>PERK</td>
<td>Double-Stranded RNA-Activated Protein Kinase-Like ER Kinase</td>
</tr>
<tr>
<td>PI3K</td>
<td>Phosphoinositide-3-Kinase</td>
</tr>
<tr>
<td>PKC</td>
<td>Protein Kinase C</td>
</tr>
<tr>
<td>PKCα</td>
<td>Protein Kinase C, Isoform α</td>
</tr>
<tr>
<td>PKCβII</td>
<td>Protein Kinase C, Isoform β II</td>
</tr>
<tr>
<td>PKCδ</td>
<td>Protein Kinase C, Isoform δ</td>
</tr>
<tr>
<td>PKCζ</td>
<td>Protein Kinase C, Isoform ζ</td>
</tr>
<tr>
<td>PKCθ</td>
<td>Protein Kinase C, Isoform θ</td>
</tr>
<tr>
<td>PKR</td>
<td>Double-Stranded RNA-Dependent Protein Kinase</td>
</tr>
<tr>
<td>PMID</td>
<td>PubMed Identifier</td>
</tr>
<tr>
<td>Poly(A)</td>
<td>Poly Adenosine Tract</td>
</tr>
<tr>
<td>Poly(I:C)</td>
<td>Polynosinic-Polycytidylic Acid</td>
</tr>
<tr>
<td>Post-TC</td>
<td>Post Termination Ribosomal Complex</td>
</tr>
<tr>
<td>PP2A</td>
<td>Protein Phosphatase 2A</td>
</tr>
<tr>
<td>PP5</td>
<td>Protein Phosphatase 5</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>PRAS40</td>
<td>Proline-Rich AKT Substrate 40kDa</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene Fluoride</td>
</tr>
<tr>
<td>QSSF</td>
<td>Quasi-Steady State Fluxes</td>
</tr>
<tr>
<td>QSSPN</td>
<td>Quasi-Steady State Petri Net</td>
</tr>
<tr>
<td>RAPTOR</td>
<td>Regulatory-Associated Protein of mTOR</td>
</tr>
<tr>
<td>Rheb</td>
<td>Ras-Homologue Enriched in Brain</td>
</tr>
<tr>
<td>Rictor</td>
<td>Rapamycin-Insensitive Companion of mTOR</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>RSK1</td>
<td>p90 S6 Kinases, Isoform 1</td>
</tr>
<tr>
<td>RSK2</td>
<td>p90 S6 Kinase, Isoform 2</td>
</tr>
<tr>
<td>S6K1</td>
<td>p70 S6 Kinases, Isoform 1</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium Dodecyl Sulphate</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>Ser</td>
<td>Serine</td>
</tr>
<tr>
<td>siRNA</td>
<td>Small Interfering RNA</td>
</tr>
<tr>
<td>SLBP</td>
<td>Stem Loop Binding Protein</td>
</tr>
<tr>
<td>SREBP1</td>
<td>Sterol-Regulatory Element Binding Protein 1</td>
</tr>
<tr>
<td>SREBP1/2c</td>
<td>Proteolytic Processed form of SREBP1/2</td>
</tr>
<tr>
<td>SREBP2</td>
<td>Sterol-Regulatory Element Binding Protein 2</td>
</tr>
<tr>
<td>STAT3</td>
<td>Signal Transducer and Activator of Transcription 3</td>
</tr>
<tr>
<td>TAK1</td>
<td>Transforming Growth Factor-β-Activated Kinase 1</td>
</tr>
<tr>
<td>TC</td>
<td>Ternary Complex</td>
</tr>
<tr>
<td>TCID_{50}</td>
<td>50% Tissue Culture Infectious Dose</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Name</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>TGF-β</td>
<td>Transforming Growth Factor-β</td>
</tr>
<tr>
<td>Thr</td>
<td>Threonine</td>
</tr>
<tr>
<td>TLR</td>
<td>Toll-Like Receptor</td>
</tr>
<tr>
<td>TN</td>
<td>True Negative</td>
</tr>
<tr>
<td>TP</td>
<td>True Positive</td>
</tr>
<tr>
<td>Tpl2</td>
<td>Mitogen-Activated Protein Kinase Kinase Kinase 8</td>
</tr>
<tr>
<td>tRNA</td>
<td>Transfer Ribonucleic Acid</td>
</tr>
<tr>
<td>Trp</td>
<td>Tryptophan</td>
</tr>
<tr>
<td>TSC</td>
<td>Tuberous Sclerosis Complex</td>
</tr>
<tr>
<td>TSC1</td>
<td>Hamartin</td>
</tr>
<tr>
<td>TSC2</td>
<td>Tuberin</td>
</tr>
<tr>
<td>TTP</td>
<td>Tristetraprolin</td>
</tr>
<tr>
<td>Tyr</td>
<td>Tyrosine</td>
</tr>
</tbody>
</table>
List of Models and Variants

Case Study Model -

<table>
<thead>
<tr>
<th>Petri Net Component</th>
<th>Count</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place/Molecular Species</td>
<td>70</td>
<td>This model contained only molecular interactions relating to eIF4E phosphorylation and availability. This information was taken from existing literature on mammalian translation initiation and the signalling pathways regulating this process. The interactions in this model were extracted from the large-scale model (see Final Version of the Binary Model)</td>
</tr>
<tr>
<td>Transitions/Reaction</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>Edges</td>
<td>132</td>
<td></td>
</tr>
<tr>
<td>Read Edges</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Inhibitor Edges</td>
<td>43</td>
<td></td>
</tr>
</tbody>
</table>

Final Version of the Binary Model -

<table>
<thead>
<tr>
<th>Petri Net Component</th>
<th>Count</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place/Molecular Species</td>
<td>511</td>
<td>This model contained information from existing literature on mammalian translation initiation and the signalling pathways regulating this process. Additionally, this model contained the model of ERK1/2 activation created by Schilling et al. (2009) and, which was treated as a binary model during the course of this project.</td>
</tr>
<tr>
<td>Transitions/Reaction</td>
<td>582</td>
<td></td>
</tr>
<tr>
<td>Edges</td>
<td>1236</td>
<td></td>
</tr>
<tr>
<td>Read Edges</td>
<td>515</td>
<td></td>
</tr>
<tr>
<td>Inhibitor Edges</td>
<td>545</td>
<td></td>
</tr>
</tbody>
</table>

Metabolic Model -

<table>
<thead>
<tr>
<th>Petri Net Component</th>
<th>Count</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place/Molecular Species</td>
<td>699</td>
<td>This version of the Final model included information on the regulation of metabolism by the mTORC1 signalling pathway. In this regard, enzymes expression was added along with any post-translational modifications dependent upon mTORC1. Furthermore, these enzymes were also linked to the corresponding reactions in the FBA GSMN model of a RAW 264.7 murine macrophage cell line. This model was simulated with the Asynchronous method and the ODE method. This was used in Chapter 5 – Integration of the Translational Control Model with a Genome Scale Metabolic Network of a RAW 264.7 Cell.</td>
</tr>
<tr>
<td>Transitions/Reaction</td>
<td>781</td>
<td></td>
</tr>
<tr>
<td>Edges</td>
<td>1576</td>
<td></td>
</tr>
<tr>
<td>Read Edges</td>
<td>656</td>
<td></td>
</tr>
<tr>
<td>Inhibitor Edges</td>
<td>706</td>
<td></td>
</tr>
</tbody>
</table>
## Junin Virus Model -

<table>
<thead>
<tr>
<th>Petri Net Component</th>
<th>Count</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place/Molecular Species</td>
<td>539</td>
<td>Used in Chapter 4 – Modelling the Effects of Viral Infection, this variant of the Final model included extra Petri Net components to model the translational effects of JUNV infection and the effects viral factors have upon cellular signalling pathway components.</td>
</tr>
<tr>
<td>Transitions/Reaction</td>
<td>611</td>
<td></td>
</tr>
<tr>
<td>Edges</td>
<td>1301</td>
<td></td>
</tr>
<tr>
<td>Read Edges</td>
<td>532</td>
<td></td>
</tr>
<tr>
<td>Inhibitor Edges</td>
<td>557</td>
<td></td>
</tr>
</tbody>
</table>

## Murine Norovirus Model -

<table>
<thead>
<tr>
<th>Petri Net Component</th>
<th>Count</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place/Molecular Species</td>
<td>533</td>
<td>Used in Chapter 4 – Modelling the Effects of Viral Infection, this variant of the Final model was used to model MNV infection. Various Petri Net components were added to model the effects MNV has upon cellular signalling components. Furthermore, the alterations MNV induces in the cellular translational machinery were included. Such changes to the model were based upon existing literature, such as Royall et al. (2015).</td>
</tr>
<tr>
<td>Transitions/Reaction</td>
<td>604</td>
<td></td>
</tr>
<tr>
<td>Edges</td>
<td>1287</td>
<td></td>
</tr>
<tr>
<td>Read Edges</td>
<td>521</td>
<td></td>
</tr>
<tr>
<td>Inhibitor Edges</td>
<td>542</td>
<td></td>
</tr>
</tbody>
</table>

## Andes Virus Model -

<table>
<thead>
<tr>
<th>Petri Net Component</th>
<th>Count</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place/Molecular Species</td>
<td>540</td>
<td>Used in Chapter 4 – Modelling the Effects of Viral Infection, this variant of the Final model concerned the effects of ANDV infection. It was necessary to include additional Petri Net components to model how viral translation is carried out. Furthermore, the effects of ANDV on signalling factors was included.</td>
</tr>
<tr>
<td>Transitions/Reaction</td>
<td>606</td>
<td></td>
</tr>
<tr>
<td>Edges</td>
<td>1292</td>
<td></td>
</tr>
<tr>
<td>Read Edges</td>
<td>529</td>
<td></td>
</tr>
<tr>
<td>Inhibitor Edges</td>
<td>554</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 1 – Introduction
Section 1.1 – Cellular Translation Initiation

The process of translation initiation was central to the model produced during the course of this project. As a consequence of this, a complete understanding of this process is necessary.

The ability to synthesise nascent proteins is fundamental for all organisms. Translation is the cytoplasmic process of forming complex polypeptides from amino acids using messenger RNA (mRNA) transcripts to convey information encoded by genes within the nucleus (Brenner et al., 1961). The process of translation can be divided into three key stages: initiation, elongation and termination. A number of key signalling pathways act to regulate these stages; of which initiation is the most tightly regulated. The importance of this ‘energetically demanding process’ (Laxman et al., 2013) is highlighted by the presence of over 30,000 protein encoding regions in the human and mouse genomes (Lander et al., 2001; Waterston et al., 2002). Each of these genes, would in theory, need to be expressed and translated. The process of translation, particularly in higher eukaryotic and mammalian systems, is complex and not fully understood. The currently accepted view of translation initiation is shown in Figure 1.1a.
Figure 1.1a – Overview of Mammalian Translation Initiation. Adapted from Jackson et al. (2010), this illustration gives a complete overview of the currently accepted mechanism of translation initiation in mammalian cells.
As can be seen in Figure 1.1a, the process of translation can be regarded as cyclic. The end product of the termination stage is a Post-Termination Ribosomal Complex (Post-TC). Recycling of this complex must occur to allow the loading of various eukaryotic initiation factors (eIFs) onto the small 40 Svedberg (40S) ribosomal subunit to form the 43S complex (Jackson et al., 2010). Prior to translation, mature mRNA is exported from the nucleus. Such mRNA contains a 5’ methylguanosine cap and 3’ poly(A) tail with both 5’ and 3’ untranslated regions and an open reading frame (Sahin et al., 2014). Once the mRNA, and associated eIFs, are loaded onto the the 43S complex, scanning of the mRNA to find a start codon sequence within the 5’ region of the mRNA begins. This scanning process requires the unwinding of the mRNA by the action of a helicase. Once this has been found, the 48S complex is formed and this complex is committed to initiating translation at this site (Ibid). The formation of the 80S ribosome, capable of carrying out translation, occurs through the joining of the large, 60S ribosomal subunit (Jackson et al., 2010).

Section 1.1.1 – Post-Termination Ribosomal Complex Recycling

The first stage of initiation is the recycling of Post-TCs. These are formed by the release of the nascent polypeptide and the action of the eukaryotic Release Factors (eRF) 1 and 2 (Alkalaeva et al., 2006). Although only eIF3 displayed any inherent activity towards carrying out Post-TC recycling, and the subsequent formation of the 40S and 60S subunits, eIF1 and eIF1A were required for efficient removal of deactetylated tRNA and mRNA (Pisarev et al., 2007). The release of the nascent polypeptide from the 80S ribosome is mediated by the actions of eRF1 and eRF3 (Alkalaeva et al., 2006) both of which are retained on the Post-TC (Pisarev et al., 2007). However, it is possible that the retained eRF1 plays a later role in the recycling of the Post-TC (Pisarev et al., 2010). The ATP-Binding Cassette, ABCE1, has been
shown to mediate recycling of Post-TCs in a manner that requires the presence of eRF1
(Ibid).

Section 1.1.2 – Ternary Complex Formation

The Ternary Complex (TC) consists of an amino-acetylated tRNA charged with a Methionine
residue (Met-tRNA) and eIF2 bound to Guanosine Diphosphate (GDP) and an inorganic
phosphate group (Algire et al., 2005). Prior to a new round of translation initiation, the GDP
molecule is exchanged for a Guanosine Trisphosphate (GTP) molecule through the
conserved action of eIF2B in both yeast and mammals (Cigan et al., 1993; Williams et al.,
2001).

Given that eIF2 has around a 400-fold greater affinity for GDP than GTP (Panniers et al.,
1988), eIF2B can only catalyse this exchange if GTP is in excess relative to GDP ((Dholakia &
Wahba, 1989; Oldfield & Proud, 1992; Nika et al., 2000 (As cited by Williams et al., 2001)).
eIF2B functions in the formation of the TC by promoting the release of eIF2 from the
complex with eIF5 (Jennings et al., 2013). The context of this complex between eIF2 and
eIF5 will be discussed in Section 1.1.5 – 60S Subunit Joining. Of the five eIF2B subunits, the
ε-subunit is required for the exchange of GDP for GTP (Fabian et al., 1997).

eIF2 consists of α, β and γ subunits (Hashem et al., 2013). The γ-subunit of eIF2 may be
responsible for interacting with both the Met-tRNA and the 40S ribosomal subunit (Shin et
al., 2011). The γ-subunit of eIF2 links the TC to the small ribosomal subunit through eIF3
within the 43S Complex (Chaudhuri et al., 1997; Chaudhuri et al., 1999). The eIF2 β-subunit
mediates an interaction between eIF2 and eIF2B (Kimabill et al., 1998). The ability of eIF2B
to exchange GDP for GTP is partially dependent upon phosphorylation of eIF2β by Protein
Kinase A (Ibid). This interaction between eIF2β and eIF2B appears to be altered under
conditions in which eIF2α is phosphorylated (Kimball et al., 1998; Krishnamoorthy et al., 2001). The role of the α-subunit of eIF2, and the mechanism by which it is capable of inhibiting translation initiation, will be discussed in Section 1.2.4 – Phosphorylation of Eukaryotic Initiation Factor 2α.

**Section 1.1.3 – eIF4F Components & mRNA Circularisation**

The eIF4F complex is comprised of eIF4A, eIF4E and eIF4G and mediates the interaction between the mRNA and the assembling ribosome. The components of the eIF4F complex anchor mRNA to the 43S complex and allow mRNA scanning. eIF4F carries out the function of binding mRNA through the cap binding protein, eIF4E, and the unwinding of mRNA during scanning by the helicase, eIF4A (Jackson et al., 2010).

eIF4G is a scaffold protein of around 200kDa (Yan et al., 1992; Hernández et al., 1998). Two functional isoforms of eIF4G, designated eIF4GI and eIF4GII, are found in human cells (Gradi et al., 1998) and display 46% similarity at the amino acid level with some regions of the middle domain being completely conserved (Ibid). These two isoforms may be required for targeting specific mRNAs for preferential expression under certain environmental conditions (Gradi et al., 1998).

eIF4G recruits mRNA to the 48S complex through an interaction with eIF3 (Korneeva et al., 2000; LeFebvre et al., 2006; Villa et al., 2013). In addition to multiple initiation factors, eIF4G can interact with several additional proteins, including the Mitogen-Activated Protein Kinase (MAPK)-Interacting Kinase (MNK) 1 (Pyronnet et al., 1999) and a protein required for the translation of histone mRNAs, Stem Loop Binding Protein (SLBP)-Interacting Protein 1 (Cakmakci et al., 2007).
A second component of the eIF4F complex is primarily responsible for scanning mRNA in order to identify start codon sequences. This component is the DEAD (Asp-Glu-Ala-Asp)-Box RNA helicase, eIF4A (Rogers Jr. et al., 1999; Rogers Jr. et al., 2001a). The helicase activity of eIF4A is dependent upon ATP binding (Ray et al., 1985; Abramson et al., 1987; Blum et al., 1992). However, binding of either ATP or mRNA to eIF4A, in the first instance, does not preclude the other from interacting with the helicase (Lorsch & Herschlag, 1998a). The conformational changes in eIF4A occurring upon an interaction with either ATP or mRNA result in an increase in affinity for the other (Lorsch & Herschlag, 1998b). Furthermore, ADP binding results in lower affinity binding to mRNA suggesting that a succession of ATP hydrolysis events enable further mRNA unwinding (Ibid).

Two additional initiation factors, namely eIF4B (Abramson et al., 1987) and eIF4H (Richter et al., 1999; Rogers Jr. et al., 2001b; Rozovsky et al., 2008) are required to achieve optimal eIF4A activity. While a combination of both factors has been shown to increase the activity of eIF4A by around 30% (Richter-Cook et al., 1998), the effects of eIF4B and eIF4H are not simply additive with the increase being more dependent upon eIF4B (Özeş et al., 2011). eIF4B, together with eIF4G, couples the unwinding of mRNA to ATP hydrolysis by maintaining eIF4A in a state in which both ATP and mRNA are bound (Andreou & Klostermeier, 2014). Furthermore, the amount of ATP required for unwinding mRNA is markedly reduced in the presence of eIF4B (Harms et al., 2014).

The final component of eIF4F is the 5’ cap binding protein, eIF4E. This protein acts to anchor the 5’ end of mRNA to the eIF4F complex and the 43S complex. As discussed later in Section 1.2.1 – eIF4E-Binding Proteins & The mTORC1 Signalling Pathway, eIF4E availability
is tightly regulated by the 4E-Binding Proteins (4E-BPs) which function to sequester eIF4E and prevent an interaction with eIF4G (Marcotrigiano et al., 1999).

The mRNA 3’ Poly(A) tail interacts with the Poly(A) Binding Protein (PABP) (Sachs et al., 1986 (as cited by Tarun Jr. & Sachs, 1996); Görlach et al., 1994). PABP has also been shown to interact with human and yeast eIF4G via the the N-terminal domain (Tarun Jr. & Sachs, 1996; Imataka et al., 1998; Kessler & Sachs, 1998; Wakiyama et al., 2000). Given these interactions, it has been postulated that the mRNA, linked by eIF4E and PABP, and bridged by eIF4G, at the 5’ and 3’ ends, respectively, becomes circularised (Wells et al., 1998).

Section 1.1.4 – Start Codon Recognition

The start codon refers to an AUG (Adenosine-Uracil-Guanosine) nucleotide sequence found within the 5’ region of mRNA. Kozak (1987) established that in addition to this AUG sequence, several other nucleotides, both upstream and downstream, are required to ensure that this sequence is decoded by the 48S complex, only when it occurs in an appropriate context. The optimal and conserved context consists of the sequence: AXXaugG where X represents any nucleotide (Kozak, 1987; Kozak, 1997; Rogozin et al., 2001).

In addition, alternative initiation sequences, such as CUG (Cytosine-Uracil-Guanosine), are commonly used. It has been shown that the use of ‘Kozak Sequences’ as initiation sites maybe a low as a third of initiation sequences (Lee et al., 2012).

Following the eIF4A-mediated scanning, recognition of the mRNA start codon can occur. This stage of translation initiation has not been well studied in mammalian systems and, as a consequence, much of the summary here relies on work conducted in yeast. eIF1 and eIF1A both act to ensure the recognition of a start codon (Saini et al., 2010) by maintaining the
ability of the 48S complex to scan through mRNA for the start codon (Passmore et al., 2007; Hussain et al., 2014). When a start codon is recognised, eIF1 is displaced from the codon-anti-codon binding site and this converts the 48S complex to a state where scanning is no longer possible (Hussain et al., 2014).

Consistent with this view, in mammalian systems eIF1, eIF1A and eIF5 interact through the C-terminal domain of eIF5 and N-terminal domain of eIF1A (Luna et al., 2013). This interaction prevents eIF5 from accessing eIF2 in advance of start codon recognition. Once the Met-tRNA and start codon interact, this interaction between eIF1A and eIF5 is prevented and scanning ceases, allowing eIF1 to dissociate (Ibid).

Section 1.1.5 – 60S Subunit Joining

Upon recognition of the start codon, eIF5, a GTPase-Activating Protein (Paulin et al., 2001), mediates eIF2-driven GTP hydrolysis to release the tRNA. It has been proposed that eIF5 interacts with eIF2β which allows the release of the inorganic phosphate molecule from eIF2 (Luna et al., 2013). eIF5B subsequently mediates the 60S subunit joining through a second GTP hydrolysis (Pestova et al., 2000) and may be dependent upon eIF1A (Acker et al., 2006), while the bound eIFs are released (Unbehaun et al., 2004).

To prevent re-association of the 40S and 60S subunits after recycling, eIF6 acts by binding to the 60S subunit (Valenzuela et al., 1982; Raychaudhuri et al., 1984; Si et al., 1997; Wood et al., 1999). The ability of eIF6 to prevent the formation of the 80S complex is regulated by PKC β-mediated phosphorylation (Ceci et al., 2003). The yeast Elongation Factor-Like 1 protein may also regulate the interaction between eIF6 and the 60S subunit (Senger et al., 2001).
Section 1.2 – Key Points of Regulation & Upstream Signalling Pathways

The regulation of initiation occurs through the coordinated and complex activity of a number of cellular signalling pathways, including the Mitogen-Activated Protein Kinase (MAPK) pathways (Extracellular Signal-Regulated Kinase (ERK) 1/2, p38, c-Jun N-Terminal Kinase (JNK1) and ERK5), the Phosphoinositide-3-Kinase (PI3K) and the mammalian Target of Rapamycin Complex 1 (mTORC1) pathways, as shown in Figure 1.2a.
Figure 1.2a – Position of Translation Initiation within the Signalling Network. The cellular signalling network is intimately involved with the regulation of translation initiation and while the initiation phase of translation is key to the synthesis of nascent proteins it is only a small part of the model produced during the course of this project, as shown by the area within the red box. This signalling network must interact with the translational machinery at key points in order to regulate the process of initiation. The aim of this project was to produce a qualitative model of this system in which translation initiation is regulated by a signalling network.
In order to accurately model the regulation of the initiation phase of translation it is necessary to understand the points at which these signalling pathways converge on the translational machinery. Furthermore, the mechanisms by which translation initiation is affected by these pathways must also be known in order to correctly model this process. A further level of complexity is added by the large degree of cross-talk and feedback mechanisms existing in and between these pathways resulting in complex interaction networks. This section will therefore focus upon introducing the key signalling pathways and how they affect the translational machinery.

Section 1.2.1 – eIF4E-Binding Proteins & The mTORC1 Signalling Pathway

As introduced in Section 1.1.3 – eIF4F Components & mRNA Circularisation, the activity of a key initiation factor, eIF4E, is regulated by the phosphorylation of 4E-BP1, which is controlled by the mammalian Target of Rapamycin (mTOR) Complex 1 (mTORC1).

Two discrete mTOR complexes are found in mammalian cells, each with distinct biological functions (Laplante & Sabatini, 2009). mTORC1 is composed of mTOR, Regulatory-Associated Protein of mTOR (RAPTOR), DEP-Containing mTOR-Interacting Protein (DEPTOR), mammalian Lethal with Sec13 Protein 8 (mLST8), and the Proline-Rich AKT1 Substrate of 40kDa (PRAS40) (Ibid). By comparison, mTOR Complex 2 (mTORC2) contains mTOR, the Rapamycin-Insensitive Companion of mTOR (Rictor), mammalian Stress-Activated Protein Kinase Interacting Protein (mSIN1), DEPTOR and mLST8 (Leplante & Sabatini, 2009). The activation on mTORC1 is not fully understood but several discrete mechanisms have been observed.

Upstream of mTORC1 lies the Phosphoinosite 3-Kinase (PI3K) pathway component Protein Kinase B (AKT1) (summarised in Figure 1.2.1a). This Ser/Thr kinase is a member of the AGC
(Protein Kinase A, Cyclic Guanosine Monophosphate-Dependent Protein Kinase and Protein Kinase C) kinase family and expressed in three isoforms: AKT1, AKT2 & AKT3 (Hanks & Hunter, 1995; Scheid & Woodgett, 2001; Zinda et al., 2001). AKT1 is phosphorylated and activated in response to several stimuli, including insulin (Alessi et al., 1996; Kulik et al., 1997; Hermann et al., 2000), by 3-Phosphoinositide-Dependent Protein Kinase-1 (PDK1) (Alessi et al., 1997b) and mTORC2 (Hresko & Mueckler, 2005; Sarboassov et al., 2005; Guertin et al., 2006; Ikenoue et al., 2008). Feedback mechanism within this pathway are discussed in Section 3.1.3 – Interplay Between mTORC1 & AKT1.
Figure 1.2.1a – Schematic of the mTORC1 Pathway & the Downstream Components Regulating Translation Initiation. Upon binding of Insulin to the Insulin Receptor, the Insulin Receptor Substrate 1 (IRS-1) becomes phosphorylated (Gual et al., 2005). Downstream of IRS-1 is Phosphatidylinositol 3-Kinase (PI3K) (Shaw, 2001; Khamzina et al., 2003). Downstream of PI3K lies the 3-Phosphoinositide-Dependent Protein Kinase (PDK1) and Protein Kinase B (AKT1) (Cantley, 2002; Hemmings & Restuccia, 2015). A further signalling component downstream of PI3K is the Mammalian Target of Rapamycin Complex (mTORC) 2 (Liu et al., 2015). The activation of AKT1 requires phosphorylation by both PDK1 and mTORC2 (Alessi et al., 1997b; Hresko & Mueckler, 2005; Sarbassov et al., 2005). Once activated, AKT1 serves to activate, via several mechanisms, mTORC1 (discussed later in this Section). Downstream of mTORC1, two proteins involved in the regulation of translation initiation are phosphorylated, p70 Ribosomal S6 Kinase 1 (S6K1) (Seufferlein & Rozengurt, 1996; Burnett et al., 1998) and eIF4E-Binding Protein 1 (4E-BP1) (Beretta et al., 1996; von Manteuffel et al., 1996).
The first mechanism of mTORC1 activation involves the phosphorylation, and inhibition of the protein Tuberin (TSC2) by AKT1 (Inoki et al., 2002; Potter et al., 2002; Cai et al., 2006) and several other key signalling components including, ERK1/2 (Arvisais et al., 2006; Winter et al., 2011), the p90 Ribosomal S6 Protein Kinase (RSK1) (Roux et al., 2004) and the p38 substrate MAPK-Activated Protein Kinase-2 (MK2) (Li et al., 2003). TSC2, along with Hamartin (TSC1), composes the Tuberous Sclerosis Complex (Carbonara et al., 1994; Green et al., 1994 (as cited by Tee et al., 2003)). In the unphosphorylated state, TSC2 interacts with the GTPase protein, Ras-Homologue Enriched in Brain (Rheb) and the hydrolysis of Rheb-bound GTP occurs, which maintains mTORC1 in an inactive state (Garami et al., 2003; Inoki et al., 2003; Zhang et al., 2003). However, once phosphorylated, the GTPase Activating Protein (GAP) activity of TSC2 is inhibited and the ability of Rheb to hydrolyse bound GTP is diminished (Inoki et al., 2002; Potter et al., 2002).

AKT1 further activates mTORC1 through the phosphorylation of PRAS40 (Kovacina et al., 2003; Nascimento et al., 2006; Vander Haar et al., 2007) (as shown in Figure 1.2.1b). PRAS40 interacts, via RAPTOR, with mTORC1 with PRAS40 preventing mTORC1 substrates interacting with RAPTOR (Sancak et al., 2007; Wang et al., 2007c). Such inhibition is prevented by AKT1-mediated phosphorylation of PRAS40 (Sancak et al., 2007).
AKT1 is known to mediate the activation of mTORC1 through two distinct, complementary mechanisms. The first is the phosphorylation of the Tuberin (TSC2) component of the Tuberous Sclerosis Complex 1/2 complex. AKT1-dependent TSC2 phosphorylation is believed to occur on multiple Serine residues (Inoki et al., 2002). The result of this phosphorylation is suppression of the GTPase Activating Protein (GAP) activity of TSC1/2 toward Ras-Homologue Enriched in Brain (Rheb) (Garami et al., 2003). This results in the activation of mTORC1 via a mechanism that is not clearly defined. The second mechanism concerns the direct phosphorylation of Proline-Rich AKT Substrate 40kDa (PRAS40) (Kovacina et al., 2003). In the unphosphorylated form, PRAS40 prevents substrate recognition by the Regulatory-Associated Protein of mTOR (RAPTOR) (Wang et al., 2007a). Once PRAS40 is phosphorylated the mammalian Target of Rapamycin (mTOR) Complex 1 (mTORC1) becomes activated as RAPTOR is free to interact with mTORC1 substrates, such as 4E-BP1 and S6K1. The two mechanisms are believed to act in a synergistic manner (Sancak et al., 2007).
A further mechanism requires the localisation of mTORC1 to the lysosome (Sancak et al., 2008). This localisation is mediated through the actions of four small heterodimeric GTPase proteins, termed RagA through to RagD, and a five protein complex, termed Ragulator, which serves as a Guanine Nucleotide Exchange Factor (GEF) for the Rag GTPases (Sancak et al., 2008; Sancak et al., 2010; Bar-Peled et al., 2012). The GATOR1 complex opposes the GEF activity of the Ragulator complex by providing Rag-directed GAP activity (Bar-Peled et al., 2013). The activation of mTORC1 by the Ragulator complex requires amino acids and the lysosomal vacuolar H⁺/ATPase (Zoncu et al., 2011; Bar-Peled et al., 2012; Jewell et al., 2015). This mechanism of activation is further complicated by the fact that different amino acids activate mTORC1 through different Rag proteins and that additional proteins are required to allow Rag GTPases to respond to specific amino acids (Jewell et al., 2015; Resamen et al., 2015; Wang et al., 2015).

Both mechanisms of activation serve to allow mTORC1 to respond to a diverse array of stimuli and may in fact work to complement each other. In support of this idea, it is worth noting that Rheb is localised to the lysosomal membrane (Saito et al., 2005 (as cited by Groenewald & Zwartkruis, 2013)).

In a tissue-specific manner, vertebrates encode three proteins that sequester eIF4E from an interaction with eIF4G: 4E-BP1, 4E-BP2 and 4E-BP3 (Marcotrigiano et al., 1999; Graber et al., 2013). The sequence to which eIF4E interacts with 4E-BP1, or eIF4G, is Tyr-x-x-x-x-Leu-h (in which x represents any amino acid whilst h denotes a hydrophobic residue) (Marcotrigiano et al., 1999; Niedzwiecka et al., 2002). Interactions with 4E-BP1 prevent an interaction between eIF4E and the mRNA cap structure (Tomoo et al., 2005; Volpon et al., 2006). Hypophosphorylated 4E-BP1 interacts with eIF4E, as shown in Figure 1.2.1c. Once
phosphorylated, by the RAPTOR subunit (Hara et al., 2002; Schalm et al., 2003) of mTORC1 (Beretta et al., 1996; von Manteuffel et al., 1996; Burnett et al., 1998; Gingras et al., 1998; Marcotrigiano et al., 1999), 4E-BP1 dissociates from eIF4E and allows the eIF4F complex to form. Multiple residues on 4E-BP1, Thr-37, Thr-46, Thr-70 and Ser-65, are targeted for phosphorylation (Gingras et al., 2001) with phosphorylation of positions 37 and 46 occurring first (Gingras et al., 1999).
Figure 1.2.1c – Regulation of Translation Initiation by the Sequestration of eIF4F Components. The eIF4F complex consists of the scaffold protein, eIF4G, the cap binding protein eIF4E and the mRNA helicase, eIF4A. The availability of eIF4E and eIF4A are tightly regulated by interactions with the binding proteins 4E-BP1 and Programmed Cell Death 4 (PDCD4), respectively. Once sequestered, these initiation factors are unable to interact with eIF4G to form the eIF4F complex. Consequently, the 43S complex is then unable to form. However, the activity of these binding proteins is tightly regulated by the mTORC1 and PI3K signalling pathways. Once these pathways are activated, 4E-BP1 becomes phosphorylated by mTORC1 (Beretta et al., 1996) and PDCD4 is phosphorylated by AKT1 (Palamarchuk et al., 2005) and S6K1 (Dorrello et al., 2006). These post-translational modifications have an inhibitory effect on the interactions with either eIF4E or eIF4A (as shown by the red round-headed arrows) and, thus, these factors are free to form the eIF4F complex.
Unlike the regulation of 4E-BP1, the interaction between eIF4E and 4E-BP2 is only partially regulated by mTORC1-dependent phosphorylation. 4E-BP2, within the brain of mammals, interacts with RAPTOR at higher affinity when two Asparagine residues, not conserved in 4E-BP1 or 4E-BP3 at positions 99 and 102, are deamidated (Bidinosti et al., 2010). The regulation of 4E-BP3 is not fully understood (Kleijn et al., 2002).

Additionally, truncation of 4E-BP1 (Tee & Proud, 2002; Constantinou et al., 2008; Dennis et al., 2011) and 4E-BP2 (Wollenhaupt et al., 2012), at the N-terminal phosphorylation sites, results in sequestration of eIF4E. Such truncation promotes the long-term, stable repression of translation and occurs primarily during cellular stress (Constantinou et al., 2008).

The availability of eIF4A is similarly regulated by the binding protein, Programmed Cell Death 4 (PDCD4) (Yang et al., 2003; Zakowicz et al., 2005; Suzuki et al., 2008). As with 4E-BP1, the phosphorylation of PDCD4 serves to negatively regulate the PDCD4-eIF4A interaction (Palamarchuk et al., 2005; Dorrello et al., 2006).

**Section 1.2.2 – eIF4E Phosphorylation and The Role of ERK1/2 & p38**

eIF4E phosphorylation occurs in response to multiple stimuli, including oxidative stress (Rao, 2000; Duncan et al., 2003; Duncan et al., 2005) and viral infection (Mizutani et al., 2004), and may contribute to the transformation process (Topisirovic et al., 2004; Bianchini et al., 2008; Furic et al., 2010; Robichaud et al., 2015). Whilst the relevance of eIF4E phosphorylation has been hard to elucidate (Minich et al., 1994; Scheper et al., 2001; Zuberek et al., 2003), recent work has shown it mediates a change in gene expression during cellular stress responses (Andersson & Sundler, 2006; Herdy et al., 2012; Royall et al., 2015).
Extracellular Signal-Regulated Kinase (ERK) 1 (44kDa) and ERK2 (42kDa) are two of the most canonical signalling components in the cell and are responsible for the phosphorylation of over 160 proteins with a variety of cellular functions (Cobb et al., 1994; Ferrer et al., 2001; Yoon & Seger, 2006 (as cited by Zampieri et al., 2007). ERK1/2 are situated downstream of MAPK/ERK Kinase (MEK) 1/2 and Raf-1 (Crews et al., 1992; MacDonald et al., 1993), as shown in Figure 1.2.2a. These two protein kinases are activated by dual phosphorylation of sites Thr-183 and Tyr-185 (Payne et al., 1991; Robbins et al., 1993; Canagarajah et al., 1997).
Figure 1.2.2a – Signalling Cascade Resulting in the Activation of ERK1/2. Extracellular Signal-Regulated Kinase 1/2 (ERK1/2) is a substrate of MAPK/ERK Kinase 1/2 (MEK1/2). In order to phosphorylate MEK1/2, RAS, when bound to Guanosine Trisphosphate (GTP), phosphorylates Raf-1 (MacDonald et al., 1993; Jelinek et al., 1996). The activation of this pathway occurs in response to a wide variety of environmental stimuli through cell surface receptors, including the Erythropoietin Receptor (EpoR) (Guillard et al., 2003; Miyake et al., 2013) and the Epidermal Growth Factor Receptor (EGFR) (Correa-Meyer et al., 2002; Pastore et al., 2005).
Several of the 160 ERK1/2 substrates have important functions in the regulation of translation initiation, as outlined in Figure 1.2.2b. These include the phosphorylation of RSK1 (Gavin & Nebreda, 1999) and the phosphorylation, and activation, of the mTORC1 component, RAPTOR (Carrière et al., 2010).
Figure 1.2.2b – Effects of ERK1/2 Signalling on Translation Initiation. Once activated, ERK1/2 induces several stimulatory effects on translation initiation. The phosphorylation of the mTORC1 component Raptor brings about the full activation of the mTORC1 pathway, resulting in the phosphorylation of 4E-BP1. Translation is also enhanced by ERK1/2 through the RSK1-mediated phosphorylation of helicase accessory factor, eIF4B and through the suppression of 4E-BP1 expression. The other main mechanism by which ERK1/2 can affect translation is through the phosphorylation of eIF4E via the ERK1/2 substrate, MNK2. The phosphorylation of eIF4E is viewed as allowing the selective translation of specific mRNAs. Through these pathways, ERK1/2 can be seen as a central regulator of translation initiation.
Four isoforms of p38 (α, β, γ and δ) are expressed in a tissue-specific manner in mammals and share between 45% and 75% amino acid sequence homology (Jiang et al., 1996; Mertens et al., 1996; Goedert et al., 1997; Hu et al., 1999; Korb et al., 2006; Cuenda & Rousseau, 2007). Three kinases have been implicated in the activation of p38: Mitogen-Activated Protein Kinase Kinase (MKK) 3 (Derijard et al., 1995; Raingeaud et al., 1996; Makeeva et al., 2006), MKK4 (Winston et al., 1997; Ganiatsas et al., 1998; Blüher et al., 2009) and MKK6 (Raingeaud et al., 1996; Bode et al., 2001; Brancho et al., 2003). It is believed that MKK3 is the main kinase responsible for activation of p38 (Galan-Moya et al., 2011). Furthermore, MKK4 is only responsible for phosphorylating p38 in vitro (Winston et al., 1997; Ganiatsas et al., 1998; Blüher et al., 2009).

The activation of p38 produces considerable effects on the cell and a number of substrates have been identified as being downstream of this kinase, including MK2 (Alessi et al., 1996; Zu et al., 1996; Krump et al., 1997), Cyclic AMP-Dependent Transcription Factor (ATF-2) (Raingeaud et al., 1995; Waas et al., 2001; Ouwens et al., 2002) and Mitogen- and Stress-Activated Kinase 1 (MSK1) (Deak et al., 1998; McCoy et al., 2005; van der Heide et al., 2011). Downstream of MK2, p38 plays a role in the regulation of translation, as shown in Figure 1.2.2c, by the phosphorylation of Tristetraprolin (TTP) which has a considerable affect on mRNA stability by negatively regulating the degradation of mRNAs containing Adenosine and Uridine-Rich Elements (Clement et al., 2011; Tiedje et al., 2012).
The activation of p38α occurs in response to cellular stresses, such as Lipopolysaccharide (LPS) treatment. The main cellular kinase responsible for the activation of this MAPK is MAPK Kinase 3 (MKK3) (Galan-Moya et al., 2011). Upstream of MKK3 lies several kinases, including Transforming Growth Factor-β1-Activated Kinase 1 (TAK1) (Kim et al., 2007a; Xin et al., 2010). Other kinases also shown to be responsible for MKK3 phosphorylation include Mixed Lineage Kinase 3 (Tibbles et al., 1996; Abi Saab et al., 2012; Zhou et al., 2014), MAPK Kinase Kinase (MEKK) 3 (Uhlik et al., 2003) and MEKK4 (Abell et al., 2007). Once phosphorylated by MKK3, p38α is capable of not only mediating a change in gene expression, through Cyclic AMP-Dependent Transcription Factor (ATF-2) phosphorylation, but it plays a key role in controlling translation. The first such role involves the phosphorylation of MAPK-Interacting Kinase (MNK) 1. This kinase mediates an increase in eIF4E phosphorylation. The second role in controlling translation is found in regulating the stability of specific mRNA. This occurs through the p38α substrate, MK2. Once active, MK2 phosphorylates Tristetraprolin (TTP) and the degradation of mRNAs containing Adenosine and Uridine Rich Elements.

Figure 1.2.2c – The Activation of p38α & The Effects of this Kinase on Translation and Gene Expression.
Despite p38 and ERK1/2 having different roles in the regulation of translation initiation, both kinases share roles in regulating the expression of 4E-BP1 and the phosphorylation of eIF4E. Both p38 and ERK1/2 have been shown to negatively regulate 4E-BP1 expression through the phosphorylation of Elk-1 (Gillé et al., 1996; Cruzalegui et al., 1999; Guha et al., 2001; Rolli-Derkinderen et al., 2003).

Both MNK1 and MNK2 have been shown to phosphorylate eIF4E at position Ser-209 (Scheper et al., 2001; Ueda et al., 2004). The activity of MNK2 is constitutive and maintains the phosphorylation (Ueda et al., 2004), whilst MNK1 activity is inducible (Scheper et al., 2001). These two kinases are regulated by the ERK1/2 and p38 mitogen activated protein kinases cascades (Fukunaga & Hunter, 1997; Waskiewicz et al., 1997). MNK1 is regulated by p38 whilst ERK1/2 is solely responsible for MNK2 phosphorylation (Fukunaga & Hunter, 1997; Waskiewicz et al., 1997). The ability of MNK1 to phosphorylate eIF4E is not exclusively dependent upon p38. It must first interact with eIF4G, within the 43S complex, and requires the phosphorylation of eIF4G at position Ser-1186 by Protein Kinase C (PKC) α (Dobrikov et al., 2011; Walsh & Mohr, 2014).

**Section 1.2.3 – Ribosomal Protein S6 Kinases**

The activities of the kinases responsible for the phosphorylation of the ribosomal S6 protein are a good example of convergence of biological activity. The two classes of enzyme responsible for this have been termed the p70 Ribosomal S6 Kinases (S6K1) (Shima et al., 1998) and RSK1 (Erikson & Maller, 1985). Both of these kinases are also able to phosphorylate eIF4B (Shahbazian et al., 2006; van Gorp et al., 2009), thereby enhancing the activity of eIF4A. Despite these similarities, both S6K1 and RSK1 have distinct cellular functions. The S6K1 acts to negatively regulate mammalian Target of Rapamycin Complex 2
(Julien et al., 2010) and provide negative feedback to Insulin receptor signalling (Tremblay et al., 2007). RSK1 can act to stimulate mTORC1 via suppression of Tuberous Sclerosis Complex function (Roux et al., 2004).

RSK1 is activated by ERK1/2 (Gavin & Nebreda, 1999), ERK5 (Ranganathan et al., 2006) and the PDK1 (Jensen et al., 1999), as shown in Figure 1.2.3a. Interestingly there is some evidence that RSK1 is able to phosphorylate ERK5, suggesting that dual regulation may exist here (Le et al., 2013). The regulation of S6K1 is more straight-forward requiring dual phosphorylation. Both mTORC1 (Seufferlein & Rozengurt, 1996; Burnett et al., 1998) and PDK1 (Alessi et al., 1997a) are required for activation of this kinase by phosphorylating Ser-412 (Isotani et al., 1999) and Thr-252 (Alessi et al., 1997a).
The activation of the p70 Ribosomal S6 Kinase (S6K1) and p90 Ribosomal S6 Kinase (RSK1) involves phosphorylation (Magnuson et al., 2012). While 3-Phosphoinositide-Dependent Protein Kinase (PDK1) can phosphorylate both RSK1 and S6K1 (Alessi et al., 1997a; Jensen et al., 1999), RSK1 can also be activated by Extracellular-Signal Regulated Kinase (ERK) 1/2 (Gavin & Nebreda, 1999). However, in order to fully activate S6K1, an additional mTORC1-dependent phosphorylation is also required (Seufferlein & Rozengurt, 1996; Burnett et al., 1998). In terms of biological function, RSK1 and Protein Kinase B (AKT1) both function in the phosphorylation of Tuberous Sclerosis Complex (TSC) 2 protein (Inoki et al., 2002; Potter et al., 2002; Roux et al., 2004) to bring about Mammalian Target of Rapamycin Complex 1 (mTORC1). Both S6K1 and RSK1 function to phosphorylate eukaryotic Initiation Factor 4B (not shown here) (Shahbazian et al., 2006; van Gorp et al., 2009).
Section 1.2.4 – Phosphorylation of Eukaryotic Initiation Factor 2α

Eukaryotic initiation factor 2α (eIF2α) phosphorylation occurs in response to a variety of cellular stresses, including viral infection (Connor & Lyles, 2005; McInerney et al., 2005; Krähling et al., 2009; Elbahesh et al., 2011; Welnowska et al., 2011), amino acid starvation (Thiaville et al., 2008; Jiménez-Díaz et al., 2013) and hypoxia (Koumenis et al., 2002; Owen et al., 2005; Zhu et al., 2009; Liu et al., 2010). Phosphorylation of the Serine residue at position 51 (Rhoads, 1999) (Ser-51) is an efficient mechanism of inhibiting translation as this interaction inhibits the ability of eIF2B to exchange GDP for GTP (Krishna et al., 1997; Ramaiah et al., 1994; Kimball et al., 1998; Sudhakar et al., 2000), thus preventing a new round of translation initiation occurring.

While different mechanisms have been proposed in yeast and mammalian cells, the phosphorylation of eIF2α leads to a situation in which the interaction between the β- and γ-subunits of eIF2 and eIF2B, that allows the exchange of GDP for GTP, to be removed (Kimball et al., 1998; Krishnamoorthy et al., 2001).

This mechanism of inhibiting translation is mediated by four kinases capable of phosphorylating eIF2α: Double-Stranded RNA-Dependent Protein Kinase (PKR) (Romano et al., 1995; Romano et al., 1998; Ung et al., 2001; Vattem et al., 2001), Double-Stranded RNA-Activated Protein Kinase-Like ER Kinase (PERK) (Shi et al., 1998; Harding et al., 1999; Sood et al., 2000; Kouroku et al., 2007), Haem-Regulated eIF2α Kinase (HRI) (Crosby et al., 1994; Han et al., 2001; Liu et al., 2008) and General Control Nonderepressing 2 (GCN2) (Dever et al., 1992; Ramirez et al., 1992; Dever et al., 1993; Berlanga et al., 1999; Hao et al., 2005; Maurin et al., 2005). The stresses that result in the phosphorylation of eIF2α are shown in Figure 1.2.4a.
Figure 1.2.4a – Cellular Stresses Leading to the Inhibition of Translation through the Phosphorylation of eIF2α. The phosphorylation of eIF2α is a powerful mechanism by which cellular translation can be inhibited. Four cellular kinases have been shown to phosphorylate this initiation factor on Ser-51. This phosphorylation occurs in response to a wide variety of cellular stresses. It is worth noting that stresses resulting in the activation of HRI and PERK are both considered to be Endoplasmic Reticulum Stresses.
Section 1.2.5 – Emerging Roles for JNK1 in the Regulation of Translation Initiation

JNK is expressed from three gene: jnk1, jnk2 and jnk3. Ten isoforms of JNK can be expressed from alternative splicing of these genes (Davis, 2000). While JNK1 and JNK2 are expressed ubiquitously, JNK3 expression is typically expressed in the brain and testes (Ibid). JNK1 is canonically associated with stress responses and apoptosis (Dérijard et al., 1994; Minden et al., 1994; Paraskevas et al., 1999; Te et al., 2015). Whilst not traditionally associated with the control of translation initiation, several emerging roles have recently began to emerge.

There is some confusion of the role of JNK1 in the phosphorylation of eIF4E. Despite earlier work revealing no role for JNK1 in the phosphorylation of this initiation factor (Fukunaga & Hunter, 1997), later work demonstrated that eIF4E phosphorylation was inhibited by inhibition of JNK1 (Gandin et al., 2010; Jiang et al., 2010). It is not clear how this regulation occurs as both expression of eIF4E has shown to be reduced by JNK1 inhibition (Jiang et al., 2010) and the eIF4E kinases, MNK1 and MNK2, have been shown to be inhibited by the JNK1 chemical inhibitor, SP600125 (Bain et al., 2007).

The second role for JNK1 in the regulation of translation is found in regulating eIF2α phosphorylation. Dual phosphorylated JNK1 acts as a kinase for Protein Phosphatase 1. Such inhibitory phosphorylation prevents the dephosphorylation of eIF2α (Monick et al., 2006). ERK1/2 prevents eIF2α phosphorylation by maintaining JNK1 inactivity (Ibid).
Section 1.3 – Overview of Viruses Modelled in this Project and Viral Strategies to Alter Host Cell Translation Initiation

Viruses are obligate intracellular parasites and as such usurp the host translation machinery. To this end, viruses have evolved several strategies to divert this machinery from translating cellular to viral mRNA, and some of these rely on manipulating the signalling pathways that regulate translation. A major part of this project was to adapt the model of cellular translational regulation to investigate the effects viruses have upon, the not only translational machinery but also, the regulatory signalling network. Given this aim, it is essential to understand the commonly used strategies by which viruses commandeer the cellular translation initiation machinery and the points within signalling pathways which they affect in order to inhibit host protein synthesis and carry out processes essential for viral replication. Moreover, as a major part of this project involved modelling the effects of viral infection upon cellular translational regulation, an introduction into the three viruses with effects incorporated into the model is necessary: Junin Virus (JUNV), Murine Norovirus (MNV) and Andes Virus (ANDV).

Section 1.3.1 – General Features & Epidemiology of Viruses Incorporated into the Model

This section describes the general feature of the three viral families incorporated into the model in Chapter 4 - Modelling the Effect of Viral Infection and provides a brief introduction to the epidemiology and genome of the three viruses in question. In the case of MNV, the overview of the epidemiology will focus primarily upon the medically important Human Norovirus (HuNV) as MNV infection, at the organism level, is largely asymptomatic.
The incorporation of the effects these viruses induced in the cellular translational and signalling machinery into the model served to increase the value of the project.

**Section 1.3.1.1 – Arenaviruses & Junin Virus**

The *Arenaviridae* family contains 22 enveloped, ambisense RNA viruses (Emonet *et al.*, 2009; Emonet *et al.*, 2011) with a broad geographical distribution including: North America (Bear Canyon Virus, Whitewater Arroya Virus and Catarina Virus), the Caribbean (Tacaribe Virus) and South America (Junin Virus (JUNV), Pichindé Virus and Machupo Virus) (Cajimat *et al.*, 2007; Cajimat *et al.*, 2013). Phylogenetic analysis of the New World Arenaviruses reveals three clades (A, B and C) (Bowen *et al.*, 1997; Archer & Rico-Hesse, 2002). Clade B viruses, such as JUNV, Machupo, Guanarito and Sabiá viruses are responsible for distinct haemorrhagic fevers (Delgado *et al.*, 2008).

The genome of these cytoplasmic viruses is composed of two segments: the Large segment (7.3kb) and Small segment (3.5kb) (de la Torre, 2009; Emonet *et al.*, 2011). The Large segment encodes the 200kDa L protein, an RNA-dependent RNA polymerase (RdRp), and the 11kDa Z protein, a Zinc Finger protein. The Small segment encodes the 75kDa viral protein precursor, GPC protein and the 63kDa Nucleoprotein (de la Torre, 2009). Two stem loop structures are found on each segment in the intergenic regions of the JUNV genome (Ghiringhelli *et al.*, 1991 (as cited by Palacios *et al.*, 2010)).

JUNV is responsible for Argenitine Haemorrhagic Fever (AHF) (Arribalzaga, 1955; Parodi *et al.*, 1958 (as cited by Peters, 2006)). The fatality rate for untreated AHF lies between 10% and 30% (Maizegui *et al.*, 1975; Harrison *et al.*, 1999). Around 70% of cases are diagnosed by muscle weakness, redness of the eyes, presyncope and subcutaneous rupturing of capillaries (Schwarz *et al.*, 1970 (as cited by Harrison *et al.*, 1999)). Other symptoms include
hypotension, lymphadenopathy, irritability and tremors in the hand and tongue (Enria et al., 2008). Nearly a third of cases go on to develop haemorrhage and severe neurological complications, including convulsions and coma (Ibid). A combination of leukopenia and thrombocytopenia characterise this disease (Harrison et al., 1999).

New World Arenaviruses are zoonotic infections (Charrell & de Lamballerie, 2010). Serological investigations have revealed that predominant rodent species infected with JUNV are Calomys musculinus and Calomys laucha, however several other rodent species and mammals may also act as hosts (Mills et al., 1991; Mills et al., 1992; Mills et al., 1994).

Section 1.3.1.2 - Caliciviruses & Murine Norovirus

The Caliciviridae family contains five genera (Norovirus, Sapovirus, Lagovirus, Vesivirus and Nebovirus (ICTV Website, 2012)) of positive-sense, single-stranded RNA viruses (Pesavento et al., 2008). The name of this family is derived from the Chalice, or Calyx, shaped indentations covering the viral surface (Madely, 1979). The genomes of these viruses contain between two and four open reading frames (Alhatlani et al., 2015) with fairly conserved lengths of between 7.4kb and 7.7kb (Meyers et al., 1991; Carter et al., 1992; Rasschaert et al., 1994; Ward et al., 2007). As HuNV cannot be cultivated (Duizer et al., 2004; Leung et al., 2010), several other viruses, including MNV, are used as surrogates (Esseili et al., 2015).

The first open reading frame encodes the non-structural polyprotein (Alhatlani et al., 2015). The proteins derived from this polyprotein, in FCV, are p5.6, p32, p39 (2C-Like Nucleoside Trisphosphatase (NTPase)), p30, p13 (VPg) and p76 (Polymerase-Protease fusion) (Sosnovtseva et al., 1999; Sosnovtsev et al., 2002). In contrast with FCV, MNV requires distinct RdRp and protease proteins (Wei et al., 2001; Sosnovtsev et al., 2006). The second
and third open reading frames encode the capsid proteins (Neill et al., 1991). The fourth open reading frame of MNV encodes the virulence factor VF1, believed to regulate host cell apoptosis (McFadden et al., 2011).

Two complimentary modes of transmission, foodbourne (Fankhauser et al., 2002; Widdowson et al., 2005) and person-to-person (Gunn et al., 1980; Green et al., 2002; Isakbaeva et al., 2005), help to ensure disease dissemination. Initial cases of HuNV occur through the consumption of contaminated water or shellfish (Saitoh et al., 2007; Lowther et al., 2012; Pérez-Sautu et al., 2012). Then similar to FCV dissemination (Pedersen et al., 2000; Schorr-Evans et al., 2003; Hurley et al., 2004; Clay et al., 2006), the person-to-person spread of HuNV occurs through ingestion or inhalation of HuNV-containing fomites (Repp & Keene, 2012).

Most cases of HuNV do not require medical intervention, causing acute, self-limiting diarrhoea and vomiting lasting up to five days (Rockx et al., 2002; Lopman et al., 2004; Glass et al., 2009). It is, however, the second leading cause of viral gastroenteritis requiring medical intervention, behind Rotavirus (Medici et al., 2006; Tran et al., 2010; Lorrot et al., 2011; Payne et al., 2013; Lu et al., 2015), and causes considerable morbidity and mortality in people over 65 years (Harris et al., 2008; Trividi et al., 2012) and is responsible for hospitalising over 23,000 children, under five years, annually in the United States of America alone (Patel et al., 2008). Despite being behind Rotavirus, HuNV places an extra CAD4,000,000 burden on healthcare costs in Canada alone (Morton et al., 2015) and a multi-million pound burden on the National Health Service in the United Kingdom (Danial et al., 2011).
Section 1.3.1.3 – Hantaviruses & Andes Virus

The Hantaviruses, of the *Bunyaviridae* family, have a genome of segmented, negative-sense single-stranded RNA and cause either Haemorrhagic Fever with Renal Syndrome or Hantavirus Cardiopulmonary Syndrome (HCPS) (Plyusnin *et al*., 1996; Jonsson *et al*., 2010; Vaheri *et al*., 2013). The genome segments of Andes Virus (ANDV) are termed Large (6.6kb), Medium (3.7kb) and Small (1.9kb) (Meissner *et al*., 2002) and mainly encode the Nucleoprotein, Glycoprotein Precursor (GPC) and the RdRp (Tischler *et al*., 2003; Vera-Otarola *et al*., 2012).

Numerous Hantaviruses cause HCPS, including ANDV (Lópe *et al*., 1996; López *et al*., 1997), Sin Nombre Virus (SNV) (Elliott *et al*., 1994) and several others (Morzunov *et al*., 1995; Khan *et al*., 1996; Johnson *et al*., 1997; Torrez-Martinez *et al*., 1998; Johnson *et al*., 1999; Vincent *et al*., 2000). In contrast to the strictly zoonotic SNV (Calisher *et al*., 2007), ANDV is also transmissible through person-to-person contact (Enría *et al*., 1996; Padula *et al*., 1998; Martinez *et al*., 2005; Lázaro *et al*., 2007; Martinez-Valdebenito *et al*., 2014) with infected excretions (Castillo *et al*., 2007; Godoy *et al*., 2009). The host reservoir of ANDV appears to be members of the *Oligoryzomys* genus (Toro *et al*., 1998; Gonzalez Della Valle *et al*., 2002; Medina *et al*., 2009 (as cited by Torres-Pérez *et al*., 2010); Andreo *et al*., 2011). During ANDV infection, the host does not display symptoms, in contrast to SNV infection (Calisher *et al*., 2005).

ANDV infection has a fatality rate of around 35% and most commonly occurs in men, representing an occupational risk, with many cases being labourers in the logging or farming industries (Castillo *et al*., 2001; Unidad de Vigilancia, Dpto. de Epidemiología, Gobierno de Chile, 2014). The initial picture of ANDV infection is reminiscent of influenza. As the disease
progresses, a cough develops and hypotension, tachycardia and tachypnea are noted (Castillo et al., 2001). The majority go on to develop bleeding and renal damage or, in severe cases, haemorrhage and renal failure (Castillo et al., 2001). Extracorporeal Membrane Oxygenation has increased survival rates of HCPS in SNV infection (Dietl et al., 2008; Wernly et al., 2011 (as cited by McNulty et al., 2013)).

Section 1.3.2 – General Overview of Viral Effects on Translational Regulation

While the effects specific viruses have upon the cellular translational machinery and the signalling network were incorporated into the model, it is worth noting that many other strategies are also used by other viruses. These strategies can be broadly classified as belonging to one of several general approaches. This section describes these strategies, and the effect each has upon the translational machinery, and aims to put the specific examples incorporated into the model into a broader context.

Section 1.3.2.1 – Internal Ribosome Entry Site (IRES)-Mediated Translation & Host mRNA Degradation

Cap-dependent translation is used by the majority of cellular mRNAs and requires the 5’ cap to interact with eIF4E (Altmann et al., 1987; Carberry et al., 1989; Goss et al., 1990). In constrast, RNA secondary structures within the 5’ Untranslated Regions of viruses, termed Internal Ribosome Entry Sites (IRES) elements, can drive internal recruitment of the ribosome to engage in cap-independent translation (Kanamori & Nakashima, 2001; Nishiyama et al., 2003; Thurner et al., 2004; Willcocks et al., 2011). These elements are found in the Picornaviridae (Jang et al., 1988; Pelletier & Sonenberg, 1988; Liu et al., 1999) Flaviridae (Tsukiyama-Kohara et al., 1992; Rijnbrand et al., 1997; Fletcher &
Jackson, 2002) and Retroviruses (Ohlmann et al., 2000; Brasey et al., 2003; Herbreteau et al., 2005; Camerini et al., 2008; Weill et al., 2010; Amorim et al., 2014) and allow viral translation to continue during infection when host translation is diminished. The IRES elements of these different virus types can be, broadly, classified based upon the requirements for cellular factors (Kieft, 2008).

Whilst several viral families encode the enzymes responsible for encoding 5’ cap structures (Decroly et al., 2008; Selisko et al., 2010), others do not and require cellular cap structures to carry out viral replication. These viral families include the Bunyaviridae (Kormelink et al., 1992; Garcin et al., 1995), Arenaviridae (Raju et al., 1990; Morin et al., 2010) and the Orthomyxoviridae (Plotch et al., 1981; Dias et al., 2009). These viruses conduct cap-snatching to promote viral transcription, rather than translation, (Lelke et al., 2010; Reguera et al., 2010; Lehmann et al., 2014), while suppressing the innate interferon response of the cell (Marcus et al., 2005; Qi et al., 2010) and more generally the translational capabilities of the host.

Host mRNA degradation is also mediated by targeting the RNA binding proteins that affect mRNA stability. Members of the Alphavirus family sequester the mRNA stabilising protein, HuR, to the 3’ end of the viral genome (Sokoloski et al., 2010; Barnhart et al., 2013). This serves to not only promote the degradation of mRNAs requiring HuR, but given the large-number of HuR binding sites in intron sequences (Lebedeva et al., 2011; Mukherjee et al., 2011), this also affects the alternative splicing of mRNAs.
Section 1.3.2.2 - VP$_{g}$ Proteins

Cap-dependent translation does not specifically require the presence of a 5’ methylguanosine cap and can occur if a cap analogue is present. The VP$_{g}$ protein of the Caliciviridae family is a good example.

The viral protein, VP$_{g}$, is covalently linked to the 5’ region of the viral genome (Schaffer et al., 1980) and acts as a 5’ cap analogue (Chaudhry et al., 2006). Whilst the presence of this protein is not unique to the Caliciviridae, with VP$_{g}$ proteins being found in viral families including the Astroviridae (Jiang et al., 1993; Fuentes et al., 2012), Picornaviridae (Flanegan et al., 1977; Weitz et al., 1986) and Potyviridae (Riechmann et al., 1989; Puustinen et al., 2002), the size and exact function of this protein is fairly unique. The VP$_{g}$ function differs from that of the Picornaviruses. Whilst the Caliciviral VP$_{g}$ is required for viral protein synthesis (Daughenbaugh et al., 2003; Goodfellow et al., 2005), the VP$_{g}$ of the Picornaviridae functions only for viral genome replication (Flanegan et al., 1977; Morasco et al., 2003; Murray & Barton, 2003; Liu et al., 2007).

Much like that of the Potyviridae (Léonard et al., 2004; Hébrard et al., 2010), the Caliciviral VP$_{g}$ (Daughenbaugh et al., 2003; Daughenbaugh et al., 2006; Hosmillo et al., 2014) is responsible for the interactions with a number of cellular factors vital for the initiation of translation. However, unlike that of the Potyviridae, the VP$_{g}$ of the Caliciviridae does not interact with the PABP (Léonard et al., 2004; Beauchemin & Laliberté, 2007).

Several interactions between VP$_{g}$ and eIF3 and eIF4E are known to occur in Norwalk Virus (Daughenbaugh et al., 2003) and Feline Calicivirus (FCV) (Goodfellow et al., 2005). Differences in the requirements of eIFs exist between Caliciviruses, but overall, imply a mechanism by which the VP$_{g}$ protein is able to join the eIF4F complex to the 43S complex to
enable cap-dependent translation (Chaudhry et al., 2006). More information regarding interactions between the VP$_g$ and initiation factors is given in Section 4.2.2 – Murine Norovirus Model.

### Section 1.3.2.3 - Host Factor Degradation & Viral Mimics of Host Factors

Viral infection subjects the host cell to pronounced stress and requires the cell to respond with changes in gene expression. However, viruses are also capable of manipulating this change to confer an advantage (Villas-Bôas et al., 2009; Royall et al., 2015). An obvious viral target is the cellular translational machinery with mechanisms including the degradation of host initiation factors and signalling components and the mimicry of host factors.

To control eIF2 activity, several viruses, including Vaccinia Virus and members of the Ranavirus genera, encode functional proteins mimicking eIF2$_\alpha$, termed K3L (Essbauer et al., 2001) and vIF2$_\alpha$ (Yu et al., 1999; Grayfer et al., 2015), respectively. K3L inhibits the phosphorylation of eIF2$_\alpha$ via an interaction between the protein and PKR (Kawagishi-Kobayashi et al., 1997). vIF2$_\alpha$, however, targets the homologue of PKR in fish for degradation (Jancovich & Jacobs, 2011). PKR is also degraded during infection with the virulent Phleboviruses, such as Rift Valley Fever Virus (Habjan et al., 2009; Ikekami et al., 2009; Kalveram & Ikekami, 2013; Lihoradova et al., 2013).

### Section 1.3.2.4 - Alteration to Cell Signalling Pathways

In addition to the degradation of host factors, a commonly used strategy to alter the ability of the host to synthesise proteins is to interfere with the normal functioning of cellular signalling pathways that control translational activity (as described in Section 1.2 – Key Points of Regulation & Upstream Signalling Pathways).
Given that mTORC1 has a central role in regulating translation initiation, numerous viruses target this pathway for activation, including: ANDV (McNulty et al., 2013), Human Cytomegalovirus (HCMV) (a β-Herpesvirus) (Bai et al., 2015) and Influenza A Virus (an Orthomyxovirus) (Denisova et al., 2014).

The mechanism by which mTORC1 is activated appears to be dependent upon the family of virus that is infecting the cell. HCMV (Moorman et al., 2008) and Herpes Simplex Virus 1 (HSV-1) (Chuluunbaatar et al., 2010) both affect the activation of mTORC1 at the level of TSC2 activity. The HSV-1 protein, Us3, phosphorylates TSC2 at the site normally targeted by cellular kinases, Thr-1462 (Chuluunbaatar et al., 2010). The HCMV viral protein, pUL38, was able to interact with TSC2 in such a way that mTORC1 activity was increased with this interaction mimicking the phosphorylation of TSC2 (Moorman et al., 2008).

Section 1.4 – Background to Systems Biology Techniques

Literature reviewed in previous parts of this chapter show that translation is a fundamental process in the cell. It consumes approximately 30% of the ATP produced during oxidative phosphorylation (Buttgereit & Brand, 1995). Furthermore, this literature shows that translation initiation is regulated by a vast, interconnected network of molecular components which coordinate protein synthesis with other cellular processes in response to a diverse array of environmental cues. Due to the number of components and interactions, it is not possible to understand how the system functions by following individual scientific articles describing individual interactions. Since the beginning of this century, the Systems Biology field has robustly argued that cellular behaviour at the molecular level can only be fully understood if it is studied as a whole and that computational and mathematical modelling tools are indispensable in this process (Di Ventura et al., 2006; Chuang et al.,
The challenges of modelling biological systems relate to unprecedented complexity and the difficulty in obtaining quantitative data have led to the development of an armoury of computational systems biology techniques aimed at representing the system in different levels of detail. This being said, these approaches have not yet been fully integrated into one generally accepted methodology, with the choice of method depending upon the aim and hypothesis of individual projects.

Throughout the course of this project, several systems biology techniques were used to construct the computational model representing literature on the molecular interactions network regulating translation initiation in mammalian cells. The following sections provide the background information to these methods. The first three sections describe the three fundamental, most commonly used approaches of computational systems biology: Exact Stochastic Simulations with the Gillespie algorithm, Ordinary Differential Equations (ODEs) and Flux Balance Analysis (FBA). Following these sections, Petri Nets will be introduced as a qualitative simulation approach and a unifying framework for representing qualitative Exact Stochastic and ODE models. Finally, Quasi-Steady State Petri Nets (QSSPN) will be introduced as an attempt to integrate Petri Nets with FBA. In the final section, the general terminology of Petri Nets will be translated to a vernacular more familiar to biologists.

Section 1.4.1 – Exact Stochastic Simulations with the Gillespie Algorithm

The Gillespie algorithm for Exact Stochastic Simulations was initially developed to account for inherent, random noise in systems of coupled chemical reactions, resulting from the fact that individual reactions occur due to random collisions between molecules. In laboratory experiments, reactions are carried out in large volumes involving large numbers of molecules. A consequence of this is deterministic behaviour due to the averaging of
random effects. However, in the small volumes of living cells there may be reactions involving just a few molecules (e.g. the binding of 10 molecules per cell of the LacR transcription factor to the lactose operon) (Krebs et al., 2013). In this case, the considerable random noise in the outcome of individual reactions is expected and leads to the heterogeneity of cellular behaviours. Indeed, early simulations of prokaryotic gene expression with the Gillespie algorithm (McAdams & Arkin, 1997; Kierzek et al., 2001) predicted fundamental stochastic phenomena in gene expression. These were later experimentally demonstrated (Ozbudak et al., 2002; Elowitz et al., 2002).

The Gillespie algorithm is described in detail below as it is the workhorse of this project and used in the majority of the simulations reported here. However, the purpose of using this algorithm is different than in the majority of computational systems biology projects. The Gillespie algorithm is not used to perform quantitative simulations nor is it used to study random heterogeneity in the number of molecules in individual cells. Rather, it is used to randomly sample alternative sequences of molecular events, linking signalling network inputs and molecular effector outputs in a qualitative model of the molecular system regulating mammalian translation initiation. This important distinction will be revisited in Section 1.4.4 – The Petri Net Formalism, where the Qualitative Petri Net approach will be introduced in detail.

The Gillespie algorithm was initially formulated to model the dynamics of systems of coupled chemical reactions (Gillespie, 1977). This method is frequently referred to as an exact stochastic simulation because it proceeds by generating individual reaction events occurring at exact times. At each iteration two questions are answered: i. Which of these reactions will next occur? ii. What is the waiting time for the next reaction to occur?
In order to answer these questions, Gillespie (1977) applied a Monte Carlo approach (Ulam et al., 1947; Metropolis & Ulam, 1949). A pseudo-random number generator was used to create two numbers \( r_1 \) and \( r_2 \). These numbers are, in turn, used to calculate two additional variables, \( \tau \) and \( \mu \), respectively. The equations below (Equation 1.4.1a and Equation 1.4.1b), from Gillespie (1977), describe the calculations used to ascribe values to \( \tau \) and \( \mu \):

\[
\tau = \left( \frac{1}{\alpha_0} \right) \ln \left( \frac{1}{r_1} \right) \quad \text{Equation 1.4.1a}
\]

The value of \( \mu \) is an integer value that satisfies the following statement:

\[
\sum_{\nu=1}^{\mu-1} \alpha_\nu < r_2 \alpha_0 \leq \sum_{\nu=1}^{\mu} \alpha_\nu \quad \text{Equation 1.4.1b}
\]

In both calculations, the value of \( \alpha_0 \) is equal to the sum of the propensity functions describing all the reactions in the system. A propensity function for reaction \( \mu \) can be described as the product of the number of molecules of various substances involved in reaction \( \mu \) (\( h_\mu \)) and the stochastic rate constant for the reaction \( \mu \) (\( c_\mu \)), as shown below (Equation 1.4.1c):

\[
a_\mu = h_\mu c_\mu \quad \text{Equation 1.4.1c}
\]

In short, the answer to the question of which reaction will occur next, can be viewed as being a probability (Gillespie, 1977). The probability of instigating reaction \( \mu \) will depend upon the size of the propensity function for the reaction, \( a_\mu \), as a fraction of the sum of all propensity functions (\( \alpha_0 \)). As the propensity function of a particular reaction increases, the ratio of this to \( \alpha_0 \) also increases. Consequently, it becomes more probable that this reaction will occur. The selection of a propensity function can be exemplified with a pie chart. As
shown in Figure 1.4.1a, the probability of randomly selecting a reaction is proportional to the size of the area that this function occupies.
Figure 1.4.1a – Schematic View of Propensity Functions. The value of $\alpha_0$ would be described as the sum of all the propensity functions ($\alpha_1$ to $\alpha_5$) (Gillespie, 1977). The larger the propensity function for a particular reaction, the more likely it is to occur. As $\alpha_3$ has the largest propensity function, and so the highest fraction of $\alpha_0$, it is most likely to occur. Conversely, as $\alpha_5$ makes up the smallest fraction of $\alpha_0$, the reaction with propensity function $\alpha_5$ would occur least.
The value of $\tau$ describes the time interval taken for the next reaction to occur (Ibid) and provides an answer to the second question given above. In other words, at $t_n = t_{n-1} + \tau$, reaction $\mu$ will occur. The probability of time interval $\tau$ occurring can be described using an exponential decay (Gillespie, 1977), and it is theoretically possible for the waiting time to increase to infinity. However, the probability of this infinite waiting time interval becomes infinitesimally small. The following equation (Equation 1.4.1d) (Ibid) shows how the Gillespie algorithm calculates the probability of a waiting time interval, $\tau$, for reaction $\mu$:

$$P(\tau, \mu) = \alpha_\mu e^{-\alpha_0 \tau} \quad \text{Equation 1.4.1d}$$

Once the identity of the next reaction and associated waiting time are determined, the state of the system is updated by executing one elementary reaction event occurring at an exact time. Thus, if reaction $\mu$ is:

$$A + 2B \rightarrow C \quad \text{Equation 1.4.1e}$$

In Equation 1.4.1e, the number of molecules of $A$ will be decreased by one, the number of molecules of $B$ will decrease by two and the number of molecules of $C$ will increase by one. The simulation time is then increased by adding the waiting time, $\tau$, to the current simulation time. This way of updating the system can be described as asynchronous: in a given iteration of the simulation, only one event is executed and only the species participating in this event are updated.

The direct method of Gillespie (1977) described above, is computationally time consuming. At every iteration, two random numbers must be generated. Gibson and Bruck (2000) have proposed a method, termed the Next Reaction Method, which is equivalent to the Gillespie algorithm where only one random number per iteration needs to be generated, and where
the search for the next reaction is optimised. While very sophisticated, this method would still be very computationally expensive if the number of molecules were very large.

To address this problem, the approximate stochastic techniques have also been developed, including the τ-Leap method (Gillepsie, 2001). In this method, the discrete timestep is introduced. In each iteration, the number of firings of each reaction within this timestep is generated by a pseudo-random number from a Poisson distribution. Subsequently, new states of the system are calculated by executing all reactions capable of firing. This way of updating the system can be described as synchronous: in a given iteration of the simulation all reactions and substances are updated simultaneously. While this algorithm is much faster and practical for systems containing large numbers of molecules, it is no longer exact, as arbitrary timesteps are introduced. The synchronous updates may even lead to the generation of negative numbers of molecules if the number of reactants are not sufficiently large. This algorithm cannot be applied to systems containing reactions with small numbers of molecules (e.g. transcription factors binding to bacterial genes). As the τ-Leap method has not been used in this project, it will not be described in more detail. A possible link between these approximate stochastic methods and the exact stochastic methods is the $K$-Leap Method (Cai & Xu, 2007).

**Section 1.4.2 – Ordinary Differential Equations (ODEs)**

Deterministic modelling involves the representation of a systems as a set of coupled Ordinary Differential Equations (ODEs). This method is the most frequently used representation of biological systems (Wu et al., 2008). Individual model components are viewed as undergoing changes in concentration per unit time, as shown in Equation 1.4.2a.
A description of a system in this way requires an in depth knowledge of the system (Williamson et al., 2009).

\[ \frac{d[X]}{d[t]} = k_1[Y][Z] - k_2[A][B] \]  

Equation 1.4.2a

As shown in Equation 1.4.2a, an ODE describes changes in the concentration of a substance \([d[X]]\) per unit time \([d[t]]\). This rate of change is, therefore, the sum of all reactions that both consume and produce substance \(X\) (Wu et al., 2008; Korkola et al., 2015). In the example given in Equation 1.4.2a, the reaction involving \(Y\) and \(Z\) and described by the rate constant \(k_1\) produce \(X\), whilst that described by the constant \(k_2\), and involving \(A\) and \(B\) degrade substance \(X\). In a deterministic model, a series of ODEs is constructed to account for the synthesis and degradation of each model component. In order for ODEs to be constructed it is essential that two conditions are satisfied. Firstly, the temporal nature of the system must be known (Rangamani & Iyengar, 2008; Cowan et al., 2012). Secondly, it is necessary to have knowledge of either the volume of the compartment in which the reactions occur or the initial concentrations of each species (Cowan et al., 2012). In relation to the initial concentrations of model components, it is necessary to state that these values should be sufficiently large so as to view them as continuous variables (Rangamani & Iyengar, 2008).

Numerical simulations of ODE systems is one of the most developed areas in applied mathematics, with an array of sophisticated methods having been developed since computers became available in the second half of the 20th Century (Press et al., 2007). The common features of all algorithms is that at each iteration of the simulation a timestep is assumed. The ODEs are then used to calculate the change in concentration of each
substance during this timestep. Substance concentrations are then **synchronously** updated. Individual, highly optimised algorithms are implemented in numerical solver libraries, such as Sundials (Hindmarsh *et al.*, 2005). A detailed description of these methods is beyond the scope of this dissertation.

**Section 1.4.3 – Flux Balance Analysis (FBA)**

The application of ODE or exact stochastic simulations to large-scale, and potentially whole-cell, molecular interaction networks is hindered by a lack of sufficient, quantitative data on the values such as reaction rate constants or numbers of molecules. FBA offers a method for the analysis of whole genome-scale metabolic networks (GSMN). The cost of using this method is the introduction of two approximations. Firstly, the use of reaction fluxes rather than molecular amounts. Secondly, this method only permits the analysis of the system at steady state. As a consequence, the ODE model describing the steady state reactions between metabolites is expressed by the following equation (Equation 1.4.3a):

\[
\frac{dx}{dt} = 0 = Sv
\]  

Equation 1.4.3a

In Equation 1.4.3a, the change in concentration per unit time is equal to the product of the Stoichiometric matrix ($S$) and the flux vector ($v$) (Raman & Chandra, 2009; Orth *et al.*, 2010). The Stoichiometric matrix contains $m$ rows and $n$ columns denoting $m$ metabolites and $n$ reactions, respectively (Covert *et al.*, 2001; Benyamini *et al.*, 2010). For example, the Stoichiometric matrix $S_{ij}$ represents the stoichiometric coefficient of the $i$-th metabolite in the $j$-th reaction. Flux vector $v$ would then contain information on the metabolic reaction fluxes of $n$ reactions. At steady state this becomes a vector with all components equal to zero. (Covert *et al.*, 2001). If experimental data on some of the steady state flux values or
their ranges are available, the solution space can be further constrained by flux bounds, as shown in Equation 1.4.3b:

\[ f_l < v < f_u \]  

Equation 1.4.3b

where the flux vector \( (v) \) must fall between upper \( (f_u) \) and lower \( (f_l) \) flux bounds (Orth et al., 2010). Furthermore, it is worth noting that fluxes maybe positive or negative depending on the reversibility of the biochemical reaction being described by that particular flux (Ibid).

The boundary conditions of the FBA problem are defined by external metabolites which are not subject to the balance constraints of Equation 1.4.3a. These metabolites are treated as unlimited sources and sinks of metabolic fluxes. Usually, external metabolites are used to represent nutrients available in the medium (e.g. glucose and amino acids). The bounds of exchange reactions connecting external metabolites to the first balanced precursor in the network can be used to quantitatively constrain nutrient availability.

Flux Balance Analysis is commonly used to study full GSMNs involving thousands of reactions (Orth et al., 2010). In these systems it is never the case that enough fluxes are measured to solve Equation 1.4.3a and determine values of all fluxes. However, it is possible to determine the maximal value of any linear combination of fluxes:

\[
\max_{f_l, f_u, S} (c \cdot v), \text{ such that } \\
S \cdot v = 0 \quad \text{Equation 1.4.3c} \\
f_l < v < f_u \quad \text{Equation 1.4.3d}
\]

where, \( c \) is a vector of \( n \) coefficients set by the user. This vector defines an objective function, which expresses the metabolic functions of the cell under investigation. A
common choice is biomass, for an objective function, where coefficients corresponding to reactions producing cellular biomass components are set according to measured biomass composition and other coefficients are set to zero. The maximisation problem defined by Equation 1.5.3c is solved by linear programming algorithms, with the Simplex method being most frequently used (Orth et al., 2010). This use of linear programming guarantees finding the maximal, unique value of the objective function, but there may be many solution vectors $\mathbf{v}$ corresponding to this value. A detailed understanding of linear programming (Kantorovich, 1939 (as cited by Bentobache & Bibi, 2012)) and the Simplex algorithm (Dantzig, 1951; Dantzig, 1963 (as cited by Bentobache & Bibi, 2012)) are beyond the scope of this thesis.

One of the reasons why FBA has been widely adopted is that it offers a gradually changing level of quantitative model accuracy depending on available experimental data. If no quantitative information is available, the maximisation of the objective function can be used to check whether a certain metabolic function is possible. The conclusion drawn from a maximal value of the objective function equalling zero is that the metabolic function in question is not possible (Orth et al., 2010). Conversely, objective function values greater than zero can be interpreted as a qualitative indication of the metabolic function being feasible (Ibid). This is particularly useful in evaluating effects of perturbations and gene knockouts. For example, Beste et al. (2007) used the GSMN-TB model of Mycobacterium tuberculosis to test feasibility of growth (through a biomass objective function) for all single gene knockouts against metabolic genes. The results of which were then compared with transposon site mutagenesis (Ibid). Statistically significant predictive power has been achieved with this model (Beste et al., 2007). As quantitative information about fluxes
become available, solution space can be further constrained. For microorganisms growing in bioreactors, quantitative predictions of metabolic yields are possible and applied in metabolic engineering (Fisher et al., 2014).

Recently, Flux Balance Analysis has been widely applied to modelling metabolism of human tissues (Gille et al., 2010; Thiele et al., 2013). The capability of FBA to qualitatively predict feasible flux distributions has been further exploited in the mechanistic interpretation of transcriptome data. For example, Leoncikas et al. (2016) have analysed the transcriptomes of 2,000 individual breast cancer tumours in the context of a human GSMN (Thiele et al., 2013) and have identified metabolic pathways associated with poor prognosis.

To summarise, Flux Balance Analysis is the first approach capable of mechanistically modelling the genotype-phenotype relationship at the full genome scale. This is currently limited to metabolic genes and metabolic phenotypes only. From this perspective, one of the aims of this project was to extend the scope of mechanistic modelling to large scale signalling and gene regulatory network that cannot be studied under steady state.

**Section 1.4.4 – The Petri Net Formalism**

This project extensively used the Petri Net framework to provide a graphical language and tools for model building, integration of models developed in different formalisms and the notion of qualitative dynamic modelling. The Petri Net formalism was first developed by Carl Adam Petri in the 1962 dissertation, Kommunikation mit Automaten (Petri, 1962). While this formalism was developed to model engineering systems (Ibid), the modelling of chemical reactions has also been proposed. Thus, the formalism can be easily extended to model biochemical reactions or cellular signalling events. Before a biologically meaningful
description of this formalism is given in Section 1.4.6 – Translation of Petri Net Terminology for a Molecular Biologist, a formal definition is necessary.

Petri Nets are bipartite graphs in which two types of nodes represent places and transitions (Petri, 1962). Various types of edges link places to transitions and, given Petri Nets are bipartite graphs, no two nodes of the same type can be linked together with edges. Pre-places describe places that are linked to transitions with an edge instigated from that place, whilst a post-place is one linked to a transition by an edge originating from that transition (Wang, 2007; Breitling et al., 2008). Places within the model are assigned non-negative, integer number of tokens. The assignment of tokens to places is termed ‘marking’ (Breitling et al., 2008). The firing of transitions moves tokens from pre-places to post-places via transitions and results in the state of the system being updated (Ibid). The consumption or production of multiple tokens by one transition firing is possible. The number of tokens moved upon transition firing is specified by the weight of the edges linking pre- and post-places to transitions (Breitling et al., 2008). Each transition is enabled when the associated firing rule, concerning the marking of pre-places with sufficient tokens to account for the weighting of edges, is satisfied (Breitling et al., 2008).

Definitions of Petri Nets found in literature differ slightly in detailed notation (Wang, 2007; Breitling et al., 2008). Following Wang (2007), Petri Nets can be defined as a quintuple \( (P, T, I, O, M_0) \):

- \( P = \{p_1, p_2, \ldots, p_n\} \) and \( T = \{t_1, t_2, \ldots, t_m\} \) are finite sets of places and transitions in which \( P \cup T \neq \emptyset \) and \( P \cap T = \emptyset \).
- \( I: P \times T \rightarrow N \) describes an input function defining edges originating from pre-places, with \( N \) being a non-negative integer set

- \( O: T \times P \rightarrow N \) describes an output function defining edges linking transitions to post-places, similarly, \( N \), is a set of non-negative integers

- \( M_0: P \rightarrow N \) is the initial marking.

Changing the distribution of tokens, through the firing of enabled transitions in Petri Nets, allows the complex and dynamic behaviour of a modelled system to be represented (Wang, 2007). The transitions, which can be fired are called ‘Enabled Transitions’. The transition is enabled if the number of tokens marking all pre-places for transitions, \( t \), are greater than the weighting of edges. To put this another way, for place \( p \):

\[
M(p) \geq I(t,p)
\]

Equation 1.4.4a

Since enabled transitions are the only ones that can fire, the markings on places can only be non-negative integer values. When a transition fires, tokens marking pre-places are removed by the amount equal to the weighting of the edges connecting the place to the transition (Wang, 2007), as shown in Figure 1.4a. Concomitantly, tokens are added to the post-places with the increase in marking being equal to the weighting of edges linking the transition to the place (ibid). The new marking, \( M' \), can be described mathematically in the following way for transition, \( t \), and place, \( p \):

\[
M'(p) = M(p) - I(t,p) + O(t,p)
\]

Equation 1.4.4b
Figure 1.4.4a – Example of a Simple Petri Net. Taken from Wang (2007), this Petri Net contains three transitions \( t_1, t_2 \) and \( t_3 \) and four places \( p_1, p_2, p_3 \) and \( p_4 \). Given Petri Nets are a form of bipartite graph, places can only be linked to transitions and vice versa. Edges are used to link places to transitions and transitions to places. The edge linking the pre-place \( p_1 \) to transition \( t_1 \) has a weighting of two, meaning two tokens of species \( p_2 \) are required for the transition to fire. Once \( t_1 \) fires, two tokens of the post-place, \( p_2 \), and one token of post-place, \( p_3 \), are produced. Only one token of places, \( p_2 \) and \( p_3 \) are required to fire transitions, \( t_2 \) and \( t_3 \), respectively. The firing of either of these transitions results in the movement of one token, in each iteration of the model, to place, \( p_4 \).
In the following case study, taken from Wang (2007), the state of the Petri Net shown in Figure 1.4a, can be described in the following way (Equations 1.4.4c – 1.4.4k):

\[ P = \{p_1, p_2, p_3, p_4\}; \]  \hspace{1cm} \text{Equation 1.4.4c}

\[ T = \{t_1, t_2, t_3, t_4\}; \]  \hspace{1cm} \text{Equation 1.4.4d}

\[ I(t_1, p_1) = 2, I(t_1, p_i) = 0 \text{ for } i = 2, 3, 4; \]  \hspace{1cm} \text{Equation 1.4.4e}

\[ I(t_2, p_2) = 1, I(t_2, p_i) = 0 \text{ for } i = 1, 3, 4; \]  \hspace{1cm} \text{Equation 1.4.4f}

\[ I(t_3, p_3) = 1, I(t_3, p_i) = 0 \text{ for } i = 1, 2, 4; \]  \hspace{1cm} \text{Equation 1.4.4g}

\[ O(t_1, p_2) = 2, O(t_1, p_3) = 1, O(t_1, p_i) = 0 \text{ for } i = 1, 4; \]  \hspace{1cm} \text{Equation 1.4.4h}

\[ O(t_2, p_4) = 1, O(t_2, p_i) = 0 \text{ for } i = 1, 2, 3; \]  \hspace{1cm} \text{Equation 1.4.4i}

\[ O(t_3, p_4) = 1, O(t_3, p_i) = 0 \text{ for } i = 1, 2, 3; \]  \hspace{1cm} \text{Equation 1.4.4j}

\[ M_0 = (2\ 0\ 0\ 0)^T \]  \hspace{1cm} \text{Equation 1.4.4k}

If the enabled transition \( t_1 \) then fires, the new marking becomes:

\[ M_1 = (0\ 2\ 1\ 0) \]  \hspace{1cm} \text{Equation 1.4.4l}

The distribution of tokens following this is such that sufficient tokens are then present to enabled transitions \( t_2 \) and \( t_3 \) (Wang, 2007). If \( t_2 \) fires next, the token markings become:

\[ M_2 = (0\ 1\ 1\ 1) \]  \hspace{1cm} \text{Equation 1.4.4m}

Should \( t_3 \) then fire, the new token marking series would become:

\[ M_3 = (0\ 1\ 0\ 2) \]  \hspace{1cm} \text{Equation 1.4.4n}
In the Standard Petri Net formulation described so far, there is only one type of an edge linking places and transitions. In Extended Petri Nets, additional edge types are used to introduce additional rules. To model the action of enzymes and catalysis in biochemical systems, a read edge is used to link a pre-place to a transition (Breitling et al., 2008). If this pre-place is marked with the necessary number of tokens, the transition is enabled. The tokens marked on this pre-place are not consumed during the firing. Similarly, inhibitory edges are used to express the act of inhibition (Ibid). If a pre-place is linked to a transition through an inhibitory edge and this pre-place is marked with an adequate number of tokens to overcome the threshold value, then the firing of the transition is inhibited. An example of a Petri Net containing these differing types of edges is given in Figure 1.4.4b which also illustrates how the Extended Petri Net notation can be used to construct signalling cascades.
Figure 1.4.4b – Extended Petri Net Notation.  i. The extended Petri Net notation includes read edges to simulate the action of an enzyme. In this pre-place, linked to a transition by such an edge must be marked with sufficient tokens to overcome the threshold to satisfy the firing rule but these tokens are not consumed during the firing of the transition. An inhibitor edge is used to simulate the action of an inhibitor. When the inhibitor is marked with sufficient tokens, the transition it is linked to is inhibited. When the inhibitor is absent the transition is capable of firing. ii. Many examples of enzymatic cascades exist in molecular biology. The Extended Petri Net notation enables the construction of enzymatic cascades, including the Mitogen-Activated Protein Kinase cascades. In such a cascade, the most upstream kinase (here KinaseA) is activated by phosphorylation in response to a particular environmental stimulus activating a receptor. In the phosphorylated form, pKinaseA is able to activate via phosphorylation, a further kinase (KinaseB). In this example phosphorylated KinaseB (pKinaseB) can phosphorylate KinaseC. Phosphorylated KinaseC (pKinaseC) would then be free to act on downstream effectors, such as further kinases or transcription factors. Individual reactions, or transitions in Petri Net terminology, are denoted as R1 to R6. The forward reactions (R1, R3 and R5) are responsible for converting kinases to the phosphorylated kinases whereas reverse reactions (R2, R4 and R6) are those that mediate the dephosphorylation of kinase. Such dephosphorylation of kinases would normally require the action of a phosphatase.
**Petri Net semantics** refers to the algorithm used to fire enable transitions and the rules that govern which transitions are enabled. Simulations, within Petri Net terminology, are referred to as Token Game simulations, with the result of such a simulation being termed a ‘trajectory’ or ‘trace’ (Breitling et al., 2008). In the semantics used in the example above (Figure 1.4.4b), if sufficient numbers of tokens are present in each pre-place, a transition is enabled (*ibid*). The simulation algorithm randomly chooses and fires one transition from a list of enabled transitions, in each iteration. In such Petri Net simulations, the probability of selecting an enabled transition is equal for all enabled transitions (Breitling et al., 2008) and generates a trajectory detailing state changes but where such changes are not assigned times. However, it is possible to determine if one transition is fired before or after another (Breitling et al., 2008). Consequently, such a simulation is frequently termed **time free**.

The full dynamic simulation where integer token numbers represent physical quantities changing in real time can be performed with **Stochastic Petri Nets**, where the rules of the Gillespie algorithm are used to fire transitions and calculate waiting times. Each transition is assigned a propensity function calculated as a product of the pre-place markings and stochastic rate constant of the transition. The Gillespie algorithm (discussed in **Section 1.4.1 – Exact Stochastic Simulations with the Gillespie Algorithm**) is then used to determine which transition will occur next and after what waiting time. Subsequently, a selected transition is fired and time is increase by the waiting time.

It is possible to implement ODE simulations within Petri Nets. In a **Continuous Petri Net** real, rather than integer, numbers of tokens are used. The rate equations are then assigned to transitions and ODEs are constructed from the interactions in the Petri Net network.
(Breitling et al., 2008). State changes and simulations are carried out using the standard methods of numerically integrating ODE systems (Breitling et al., 2008).

The unique feature of the Petri Net framework, which make it exceptionally suited for the modelling of complex biological systems, is the ability to combine multiple Petri Net types into Hybrid Petri Net simulations (Breitling et al., 2007; Ghomri & Alla, 2007). Rather than defining the entire model as Stochastic or Continuous, the Hybrid Petri Nets allow the construction of a model from both stochastic and continuous transitions. Likewise, Standard Petri Net places, with integer numbers of tokens, can be used alongside continuous places with real number token marking. Furthermore, the Hybrid Petri Net algorithm combines ODE simulations with the Gillespie algorithm to integrate different types of places and transitions (Breitling et al., 2007).

Biological systems involve processes occurring at different spatial and temporal scales (Mast et al., 2014). Different levels of cellular organisations are modelled in different formalisms. For example, metabolism is modelled by FBA, while regulatory processes are usually modelled at small scale by ODE systems. As demonstrated above, Petri Nets are likely to become a commanding tool in computational systems biology.

When applied to molecular biology, Petri Nets use places to represent species such as proteins, lipids or nucleotides, and transitions to denote the interactions and reactions between places. Tokens denote the number of molecules and allow Token Game simulations to be executed. In qualitative simulations, Token Game simulations represent ‘abstract quantities whose changes, over time, correlate to changes that occur in the amounts of active proteins present in the cell’ (Ruths et al., 2008). As remarked by Ruths and colleagues (2008), the ‘abstract’ numbers of tokens are similar to the qualitative levels...
of proteins visualised in Western Blots. Should quantitative parameters be available it is possible to use Stochastic Petri Nets. Here the numbers of tokens marking a place represent the number of molecules directly. For concentrations to be modelled, Continuous Petri Nets may be used.

**Section 1.4.5 – The Quasi-Steady State Petri Net Algorithm**

The Quasi-Steady State Petri Net (Fisher *et al.*, 2013) is a recent extension to the Petri Net framework inspired by Systems Biology. The Hybrid Petri Nets are capable of integrating Gillespie algorithm and ODE simulations. The QSSPN algorithm allows further integration of Hybrid Petri Nets with Flux Balance Analysis, thus providing a unified framework for all three major algorithms of Computational Systems Biology.

This project uses the QSSPN simulation engine (Fisher *et al.*, 2013) to conduct all simulations. While the majority of the simulations presented in this dissertation could have been done with standard Petri Net tools, QSSPN was chosen to allow the integration of models developed in this project with GSMNs in Chapter 5 – Integration of the Translational Control Model with a Genome Scale Metabolic Network of a RAW 264.7 Cell. Given that QSSPN is still a new approach and that the software is used extensively as a simulation engine throughout this project, it is necessary to provide extensive background to this method.

In QSSPN models, the transitions are divided into two sets: quasi-steady state fluxes (QSSF) and dynamic transitions (DT) (Fisher *et al.*, 2013). The method then assumes quasi-steady state: following any state change resulting from DTs, the QSSF quickly reaches steady state. This requires time scale separation between two parts of the model with DTs being much slower than QSSF (*Ibid*). This assumption is motivated and satisfied by biological systems
where metabolic reactions are much faster than gene regulation and signalling processes. The assumption that following state changes in the dynamic regulatory network, the metabolic network quickly reaches steady state has already been used to integrate Boolean models and ODEs with FBA. The result of such integrations are rFBA (Covert et al., 2001) and iFBA (Covert et al., 2008). In QSSPN, the DT part of the model is described by Hybrid Petri Net semantics allowing both stochastic and continuous transitions, while QSSF is simulated in the Flux Balance Analysis framework. This makes QSSPN the first algorithm allowing integration of all three major frameworks of Computational Systems Biology: exact stochastic simulations, ODEs and FBA (Fisher et al., 2013).

The integration of QSSF and DT parts of the model is achieved through the introduction of two new types of places: Constraint Places and Objective Places (Fisher et al., 2013). The marking of the Constraint Place is determined by a dynamic simulation. The Constraint Place then translates the marking to FBA flux bounds of selected transitions in QSSF (Ibid). The Objective Place requests evaluation of the objective function and then translates the value of this objective function into a marking. For example, if an FBA model of glycolysis were to be linked to a Petri Net signalling network, the integration could be achieved by using a glycolytic enzyme regulated by the signalling network. A glycolytic enzyme, such as Phosphofructokinase, featured in the Petri Net model would serve as the Constraint Place. This place would contain information on the flux boundaries of the reaction catalysed by Phosphofructokinase in the form of an activity table. Such a table would state that if Phosphofructokinase is marked with one token, the flux boundary is set, for example, at (0.0, 100,000). However, if the enzyme is not marked with tokens, the FBA reaction would be inactive with flux bounds set at (0.0, 0.0) (Fisher et al., 2013).
simulation is carried out, the qualitative Petri Net model would mark Phosphofructokinase with a token. This would then update the flux bounds of the FBA model and, ultimately, the objective function representing the metabolic function of interest, such as maximum Acetyl Coenzyme A (Acetyl-CoA) production. This objective can then be represented as an Objective Place and objective function fed back to the dynamic part of the model. For example, one could set the activity table stating that if maximal Acetyl-CoA production is greater than zero, the Objective Place is marked with one token. Should this condition not be satisfied then no tokens would mark the Acetyl-CoA place. Given the diverse roles of glycolysis and Acetyl-CoA (Moussaieff et al., 2015), or metabolites in general, the integration of FBA models with Petri Net models would allow for a more complete understanding of how metabolic shifts are affected by cellular signalling pathways and vice versa.

QSSPN uses a general formulation of a propensity function for each transition to facilitate qualitative simulations with thresholds (Fisher et al., 2013), as shown in Equation 1.4.5a:

\[ P_t = c_t \prod_{i=1}^{N} \mu_i(x_i) \]  

Equation 1.4.5a

Where \( P_t \) and \( c_t \) describes the propensity function and rate constant of transition \( t \), respectively. \( N \) is the number of pre-places for transition \( t \), and \( x_i \) is the marking of pre-place \( i \). The variable \( \mu_i \) describes the acitivity of pre-place \( i \) in the transition as a function of its marking. The activity function, \( \mu \), is a look-up table (Equation 1.4.5b) defining the contribution of a pre-place to a propensity function of a specific transition (Fisher et al., 2013). Within this table, \( T \) thresholds, \( t_i \), and activities, \( a_i \), are contained and permit the pre-place contribution to be calculated.
\[
\mu(x) = \begin{cases} 
  x \in [t_1, t_2), & \mu(x) = a_1 \\
  x \in [t_i, t_{i+1}), & \mu(x) = a_i \\
  \ldots & \mu(x) \in [t_{T-1}, t_T), & \mu(x) = a_T 
\end{cases}
\]

Equation 1.4.5b

For example, when the marking of the pre-place \(i\) in transition \(t\) exceeds the threshold \(t_2\), but is smaller than the threshold \(t_3\), the activity function contributes value \(a_3\) (Fisher et al., 2013).

Be it stochastic, continuous or immediate, the transition type lends an interpretation of the propensity function of the transition. Regardless of the transition type, a propensity function equal to zero denotes a disabled transition (Fisher et al., 2013) as it cannot fire in a stochastic simulation scenario and does not contribute any state change in continuous simulations. In the case of stochastic transitions, the propensity function is interpreted as the probability density of a transition firing in the next infinitesimally small timestep. For continuous transitions, the propensity function is used as the reaction rate in the ODE simulation part of the hybrid algorithm. Immediate transitions are only capable of firing when the propensity function is non-zero (Ibid). Each such transition is then only fired once (Fisher et al., 2013). Both continuous and immediate transitions are fired according to the function ‘fireDeterministicTransitions(\(\Delta t\))’ (Ibid). The ODE part of the hybrid algorithm has undergone significant upgrade since the first version of QSSPN solver was published in 2013. Technical details of an adaptive timestep ODE solver used within ‘fireDeterministicTransitions(\(\Delta t\))’ function are beyond the scope of this dissertation.

Based on the marking of a constraint place, the upper- and lower-bounds of the fluxes in the QSSF model component are set using another look-up table (Equation 1.4.5c) (Fisher et al., 2013). Constraint Places are potentially linked to a number of fluxes within the QSSF model.
component. The look-up table related to a specific Constraint Place, using the number of
tokens marking this node, sets the bounds of the fluxes (Fisher et al., 2013). In the
simplified example above, such as Constraint Place would be Phosphofructokinase. The
function (setQSSFbounds()) is used to this end (Ibid).

\[
(lb(x),ub(x)) = \begin{cases} 
  x \in [t_1, t_2), & lb(x) = f_{l,1}, \quad ub(x) = f_{u,1} \\
  x \in [t_i,t_{i+1}), & lb(x) = f_{l,i}, \quad ub(x) = f_{u,i} \\
  \cdots & \cdots \\
  x \in [t_{T-1}, t_T), & lb(x) = f_{l,T}, \quad ub(x) = f_{u,T} 
\end{cases}
\]

Equation 1.4.5c

In this look-up table, the upper and lower bounds of flux are described by the variables
\(ub(x)\) and \(lb(x)\), respectively, while the \(x\) denoted the the number of tokens, or state of
the node (Fisher et al., 2013). The variables describing the number of token thresholds are
given by \(t_1, \ldots, t_T\). \(f_{l,i}\) and \(f_{u,i}\), according to the state of the node, are the lower and upper
flux bounds respectively (Fisher et al., 2013). A Boolean variable (‘UpdateRequired’) set to
TRUE is required to trigger an evaluation of the objective function of the FBA model
component if the the lower or upper flux bounds change. Once this has occurred, and prior
to the looping over of Constraint Place, the Boolean variable is re-set to FALSE (Ibid).

A QSSPN simulation, via the function evaluateObjective(), uses an Objective Place, for
example Acetyl-CoA in the simplified case given earlier in this section, to evaluate the
objective function in an FBA model. Within the QSSF model component, the objective
function may be a flux name or metabolite name. As has been stated previously, the
objective function is maximised by the Simplex algorithm (or other Linear Programming
methods) in the case of the former, or, in the latter, by maximising the sums of the fluxes
producing the objective function metabolite (Fisher et al., 2013). The most computationally
demanding stage of the QSSPN algorithm occurs when the objective function requires evaluation by linear programming-mediated maximisation of the entire QSSF model component (Ibid). Such evaluation of the objective function comes about through the variable ‘UpdateRequired’ being set to TRUE. Moreover, such a variable ensures that the evaluation only occurs when the QSSF bounds have changed (Fisher et al., 2013).

The dynamic part of the model requests information evaluation of the capabilities of steady state metabolism using the function, updateObjectiveNode() (Fisher et al., 2013). Such a function, based upon the value of the objective function, sets the state of an Objective Place. The state of the Objective Place is set using another look-up table (Fisher et al., 2013) (Equation 1.4.5d).

\[
\begin{align*}
x(o) = \left\{ \begin{array}{ll}
o \in [t_1, t_2), & x(o) = s_1 \\
o \in [t_i, t_{i+1}), & x(o) = s_i \\
o \in [t_{T-1}, t_T), & x(o) = s_T \\
\end{array} \right.
\end{align*}
\]  

Equation 1.4.5d

In this, \( x(o) \) is the state of the Objective Place. Variable \( t_1, \ldots, t_T \) defines the objective function thresholds while \( s_1, \ldots, s_T \) are integer token numbers marking an objective node (Fisher et al., 2013). The number of tokens marking this Place are determined by the magnitude of the objective function. In this look-up table, the variable \( o \) denoted the objective function (Ibid). The calculation of the objective function value is performed by the variable ‘evaluateObjective()’ (Fisher et al., 2013).

Boolean values are used to fire any transitions requiring a delay of \((t, t + t)\). TRUE or FALSE values are returned when such transitions are present or absent, respectively (Fisher et al., 2013). When a delayed transition is fired, at the scheduled time \((t_s)\), the simulation timestep \((\Delta t)\) is set to \((t_s - t)\). In a system with multiple delayed transitions, these are fired
sequentially with the order being determined by the magnitude of the $t_s$ (Fisher et al., 2013). The function ‘fireDelayedTransition($t, t$)’ is used for handling delayed transitions. Stochastic transitions with a non-zero delay time are handled by including them with delayed transitions using the function ‘scheduleDelayTransition($m$)’. The delay of stochastic transitions is governed by the waiting time calculated using the Gillespie algorithm (Fisher et al., 2013).

Now that basic operations within QSSPN method have been defined, it is possible to outline the algorithm. The algorithm begins by evaluating the number of tokens marking each Constraint Place. The state of these nodes updates the bounds of the QSSF model component. The objective function of the Objective Place is then determined using FBA maximisation (Fisher et al., 2013). Following this, using the processes described above, immediate and continuous transitions followed by stochastic transitions are executed. The firing of stochastic transitions is governed by the Gillespie algorithm (described in detail in Section 1.4.1 – Exact Stochastic Simulations with the Gillespie Algorithm). Synchronisation of stochastic immediate and continuous transitions is implemented using the Maximal Timestep Method algorithm (Puchalka & Kierzek 2004). Delayed stochastic transitions are permitted. Repeated iterations of this lead to the generation of simulation trajectories. Such trajectories are dynamic and permit the order of transition execution to be monitored. This allows identification of the effects of altering the initial marking of Constraint Places on the Objective Place and the associated objective function (Fisher et al., 2013). The algorithm is summarised in Figure 1.4.5a.
Figure 1.4.5a – Summation of the QSSPN Algorithm. Taken from Fisher et al. (2013), this figure summarises how the QSSPN algorithm orders various types of transitions and integrates the differing model types that form a QSSPN model.
Section 1.4.6 – Translation of Petri Net Terminology for a Molecular Biologist

When applied to molecular biology, Petri Nets use places to represent molecular species such as proteins, lipids or nucleotides, and transitions to denote the interactions between the molecules (Breitling et al., 2008). In this notation, marking of Petri Net places represents amounts of molecular species. In Stochastic Petri Nets, the integer number of tokens represents numbers of molecules in the volume of the cell, cellular compartment or cell micro-environment. In Continuous Petri Nets, the marking is a real number representing molecular concentrations. In qualitative simulations, the integer number of tokens represents abstract, discrete activity levels. As remarked by Ruths and colleagues (2008), the ‘abstract’ numbers of tokens are similar to the qualitative levels of proteins visualised in Western Blots.

This project capitalised on the Extended Petri Net graphical notation. The read edges used in the model of translational regulation created in this project denoted activation where, for example, there is a requirement for an enzyme or molecular co-factor that is not consumed by the reaction. Inhibitory edges are used in situations where there is a need to represent inhibition: the reaction rate is decreased by the presence of the inhibitor and, furthermore, the inhibitor is not consumed by the transition.

In the opinion of this author, one of the major challenges of interdisciplinary research is that it requires researchers, working in different, well established scientific fields, to agree on common terminology. The development of such a common language is frequently more of a social than scientific issue: the experts in individual fields have to accept that the language in which interdisciplinary research is conducted will be different than the frequently strict terminology of their field of expertise. This frequently creates barriers to interdisciplinary
work. This project is a good example of such a challenge. Should it use Petri Net terminology which is acceptable to Petri Net experts, but difficult to read for a molecular biologists or should it use molecular biology terminology that a Petri Net expert may find unacceptable as a language of algorithm description? Should the variable of interest be referred to as places or molecular species? Should mathematical terms that change variables be called transitions or interactions?

This project will use original Petri Net terminology in the above sections (Section 1.4.4 – The Petri Net Formalism and Section 1.4.5 – The Quasi-Steady State Petri Net Algorithm), where Petri Net methodology is described. However, subsequent chapters will use molecular biology terminology and will refer to the Petri Net terms in the way given at the start of this section. This project uses published computational methods to investigate biological problems. It does not introduce new algorithms. The results of this project are in the biological domain and will therefore be described in the language appropriate to molecular biology.

Section 1.5 – Aims & Objectives and Novel Contributions to Knowledge

Section 1.5.1 – Aims & Objectives

This project encompasses two main areas: first, the modelling of mammalian translation initiation and the upstream regulatory pathways and secondly, modelling the effects viral infection has upon this process.

The factors involved in translation initiation and the regulation of this process are known to play roles in the pathogenesis of medically important disease states, including heart
diseases (Lu et al., 2014), cancer (Heikkinen et al., 2013) and neurological disease (Li et al., 2004; Li et al., 2005a). A better understanding of this process, and also the regulatory pathways surrounding it, are essential to understanding the pathology of human and animal disease. To this end, the overall aim of this project can be regarded as:

**Constructing and experimentally validating a literature-based, qualitative model of translation initiation and the signalling network that regulates this process in mammalian cells.**

In terms of the second aspect of this project, viral infection, at least from a systems biology standpoint, can be viewed as ‘equivalent to a systems-level perturbation’ to the cell (Garmaroudi et al., 2010). Investigating the effects viral infection has upon the cellular translational machinery has not been attempted. The second part of the project is to model viral infection. The aim of this part of the project is to:

**Introduce the host-pathogen interactions into the model of translational regulation and establish the effects of these interactions on both the cellular translational machinery and the regulatory signalling network.**

The specific objectives of this work are:

- Construct and iteratively refine a large-scale, qualitative model representing cellular translation initiation and the signalling pathways involved in regulating this process in a mammalian system
- Validate the predictive power of this model by a comparison of model behaviours with literature data
• Apply the model to generate novel hypotheses on the regulation of translation initiation and experimentally validate these hypotheses

• Simulate the effects of viral infection and experimentally validate predicted behaviours

• Improve the predictive power of the model through the incorporation of detailed kinetic models of major regulatory hubs into the large-scale qualitative model

• Integrate the model of translational regulation with a GSMN to facilitate the study of the coordinated translational regulation with whole-cell metabolic reprogramming

Section 1.5.2 – Novel Contributions to Knowledge

This project produced a number of key outputs which can be regarded as novel. These include:

• The production of an experimentally validated model from over 1,100 pieces of peer-reviewed literature of mammalian translation initiation regulated by a large-scale signalling network

• Developing a benchmark concerning the effects of chemical inhibitors on key signalling effectors and using this to develop a method for the quantitative analysis of the predictive power of the model

• Theoretical insight into the mechanism by which Rapamycin-mediated mTORC1 inhibition can lead to an increase in eIF4E phosphorylation.

• Extending and experimentally validating the model to investigate the effects of viral infection upon the translational machinery and the regulatory signalling network
• Constructing a model setting the foundation of work looking to link the regulation of translation initiation to the regulation of metabolism
Chapter 2 – Model Creation & Refining
The construction of the model of translation initiation regulated by a large-scale signalling network was the main aim of this project. This chapter describes the work undertaken to not only construct the model from existing literature, but also to evaluate the predictive power of the model. Furthermore, detailed in this chapter is work to establish a mechanism by which the main model prediction occurs.

**Section 2.1 – Materials & Methods**

**Section 2.1.1 – Qualitative Petri Nets**

Despite recent advances in quantitative computational biology, it is still not possible to establish the rate constants of all the molecular interactions in large-scale models (Di Ventura et al., 2006). A solution to this challenge, which is gaining acceptance in the literature is to use qualitative, time-free Petri Net modelling (Ruths et al., 2008; Fisher et al., 2013). The Petri Net methodology has been described in Section 1.4.4 – The Petri Net Formalism. This Section and Section 2.1.2 – The Binary Model Formalism describe details of the specific time-free Petri Net approach used to build the translation initiation model. The terminology of the Petri Net field is translated into the vernacular more suited to molecular biology in Section 1.4.6 – Translation of Petri Net Terminology for a Molecular Biologist.

This dissertation presents an Extended Petri Net model encompassing the translational machinery and surrounding signalling pathways. Places represent molecular species and transitions denote the interactions between molecules. Read edges model the activation of interactions by certain molecular species and inhibitory edges are used to model the act of inhibition. Since there is an absence of sufficient quantitative information regarding the numbers of molecules in the cell to parameterise large-scale models, discrete activity levels
were used to represent molecular activity levels. Likewise, the time is not parameterised to time units. Time is exclusively used to state the order in which molecular interactions occurred during simulations. Therefore, the model is used to simulate the sequence of molecular events changing the activity levels of molecular species. In subsequent parts of this thesis, the sequence of molecular events will be referred to as a ‘trajectory’.

Section 2.1.2 – The Binary Model Formalism

The definition of the molecular activity levels, and the rules describing how molecular interactions change them, is the key aspect of the qualitative modelling approach. One possibility involves using multiple activity levels representing concentration ranges of molecular species. Activity levels are represented by integer numbers (tokens (or discrete activity levels)) and changed by interactions decreasing states of substrates (or pre-places in Petri Net parlance) and increasing the states of products (post-places). Calder et al. (2005) analysed different maximal numbers of discrete activity levels in qualitative models of signalling pathways and concluded that experimental data can be approximated with ten levels. Ruths and colleagues (2008) used a different strategy placing certain discrete activity levels numbers in the initial state of the signalling network and allowing Petri Net molecular interactions to exchange them without setting any upper limit. The number of discrete activity levels was then considered to be proportional to the concentration of that particular molecular species (Ibid).

The aim of this project is to build a large-scale model of the signalling network controlling protein synthesis. The resulting model is considerably larger than previously published discrete models (Calder et al., 2005; Ruths et al., 2008). For that reason it has been very difficult to calibrate the maximal number of discrete activity levels, or maximal numbers of
tokens, in the initial state. To do this, calibration would have to be repeated for every new version of the model as substantial increases to the model, in terms of scale and complexity, could potentially affect the choice of arbitrary discretisation of parameters such as maximal numbers of activity levels or the numbers of tokens in the initial states. Evaluation of the predictive power of the large-scale model is a complex exercise and it was not practical to repeat it frequently enough to optimise discretisation during model development. Therefore, more abstract, binary modelling formalisms have been developed, where molecular species were considered to be active or inactive and intermediate activity states were not considered. The original contribution of this work is the approach of setting the rules changing active/non-active species states, in such a way that biological processes are reflected with sufficient realism, to achieve statistically significant and applicable predictive power.

This work employs the following binary modelling approach. Each molecular species in the network can have either zero or one discrete activity levels. **The state of zero discrete activity levels indicates that the molecular species is not present, in a sufficient amount, to activate interactions it initiates or is a substrate in.** The state of one discrete activity level indicates that there is a sufficient amount of a molecular species to ensure that molecular interactions consuming this species, or are activated by this species, can take place. While the state of one corresponds to a higher amount than a state of zero, the proportionality between amount and discrete level is not assumed. This is justified by the fact that many, if not most, reactions in the signalling network follow a sigmoidal input/output relationship (Kothamachu et al., 2013).
Having this basic notion of a molecular species present in a sufficient amount to activate a downstream molecular process, one can proceed to the construction of the signalling network model. The elementary step of signalling transduction is an enzyme-catalysed covalent modification of a protein molecular species. Protein A is modified by the attachment of a chemical moiety, such as phosphorylation, acetylation and ubiquitination. In a modified form, Protein A can be considered to be a distinct chemical species, termed Protein B. Enzyme E is a protein that catalyses this post-translational modification of Protein A. In this signalling cascade, Protein B is an enzyme catalysing the next step of signal transduction and E is a protein activated by covalent, post-translational modification during the previous step. This signalling network could not function if the activation of B was permanent as it would lead to the permanent activation of downstream processes, even when the external signal initiating this cellular response, for example a cytokine, was no longer present. In the cell, covalently modified proteins are inactivated, either by the removal of the chemical moiety or by the degradation of a protein. Both of these processes are catalysed by specific enzymes, for example, phosphatases or a proteasome complex. In both cases the signalling system is returned to a basal level of activity and is ready to respond to a new round of stimulation. If Protein B is inactivated by the removal of the post-translational modification, then the concentration of Protein B falls and a concomitant rise in the concentration of Protein A is observed. In this case the concentration of Protein B returns to a basal level. However, in the case of the degradation of Protein B, the amount of Protein A is maintained by constitutive protein synthesis. Frequently, the literature describing signal transduction networks describes the events of covalent modification leading to forward propagation of the signal. The details of the biological mechanisms surrounding how a return to the basal levels of signalling components is achieved, upon
removal of the stimulus, remain largely undefined. In order to make a model as realistic as possible, one still needs to assume the existence of mechanisms to effectively ‘switch off’ signal propagation, and ensure the system returns to a basal state capable of reacting to repeated exposure to the stimulus.

Figure 2.1.2a shows formalism used in the construction of the binary model, using the information presented earlier in this chapter, to represent the basic signal transduction step. First, the connectivity of molecular species and biological processes are modelled using the Extended Petri Net formalism (detailed in Section 1.4.4 – The Petri Net Formalism), as shown in Figure 2.1.2ai. In reaction r1, Protein A is covalently modified to form Protein B. Common edges are used as this process must decrease Protein A levels in order to increase Protein B levels. Reaction r2 is used to represent the net effect of potentially multiple, perhaps not clearly defined, molecular mechanisms that return the signal transduction apparatus to a basal level. This molecular interaction serves to deplete the pool of Protein B and increase the level of Protein A and further requires the use of common edges. Enzyme E is connected to reaction r1 by read edges to represent catalysis. Where the mechanism underlying reaction r2 can be directly attributed to a specific enzyme, for example a phosphatase, a read edge would also be used. The situation presented here, and in the more common of situations, is where the mechanism of inactivation is not known in detail.
Despite the advances in computational systems biology over the past decade, it is still not possible to produce large-scale quantitative models of biological systems (Di Ventura et al., 2006). In order to construct a biologically meaningful model of mammalian translation initiation regulated by a large-scale signalling network it was necessary to introduce the Binary Model Formalism.

i. In a basic binary model the amount of a molecular species present can be thought of as the amount necessary to induce a particular behaviour. Consequently, one discrete activity level of sufficient amounts of Protein A and Enzyme E are necessary within a cell to enable Protein B formation, through post-translational modification, in a reaction termed r1. As Enzyme E is merely a catalyst for r1 this molecular species is linked to the reaction via a read edge. Reaction r2 is the reverse reaction in which the pool of Protein B is returned to the basal amount of Protein A, and, in a biological system may require an enzyme such as a phosphatase. Such mechanism are, in many cases, ill-defined but necessary to include in a model of a biological system.

ii. This formalism produces random oscillations which are unrealistic in a biological system as, although such systems are stochastic in nature, the majority of Protein A (shown in red) will be modified to Protein B (shown in green) until such time as Enzyme E is removed through physiological feedback mechanisms.

iii. In order to eliminate these random oscillations, rules where introduced, via an inhibitory edge between Enzyme E and reaction r2, to ensure that until the stimulus (Enzyme E) is removed, the pool of Protein B is maintained to allow potential downstream signalling to occur. iv. The Token Game simulation of this model demonstrates that oscillations are removed, with the pool of Protein B (shown in green) being maintained while the amount of Protein A (shown in red) is depleted.

v. It is common in biological systems to contain feedback mechanisms in order to allow only the appropriate activation of signalling pathways. The inclusion of rules in the model accounts for such mechanisms. When the pool of Enzyme E is depleted, possibly due to a feedback mechanism originating from Protein B, reaction r2 can then occur to return Protein A to a basal level. This system is then capable of responding, should the cell once again be exposed, to the environmental cue leading to Enzyme E activation.

vi. Following the removal of the Enzyme E-mediated inhibition of reaction r2, the depletion of Protein B (shown in green) occurs through the removal of the post-translational modification and restores the basal level of Protein A (shown in red).
Figure 2.1.2aii shows the trajectory resulting from a Token Game simulation of the model presented in Figure 2.1.2ai. This starts from the state where one discrete activity level is assigned to the molecular species representing Protein A, and no discrete activity levels are assigned to the Petri Net molecular species representing Protein B. This dynamic has the following interpretation in the modelling framework used over the course of this project: The cellular pool of Protein A is depleted below the amount sufficient to allow the enzyme-catalysed reaction r1 to occur, while the cellular pool of Protein B rises to the level required for reaction r2 to happen. This system then returns to the initial state and the process repeats. The resulting oscillatory nature is unrealistic. If Protein A and Enzyme E are present in the cell, in sufficient amounts to function, the protein pool should shift towards the covalently modified form. This situation should remain until Enzyme E is inactivated by an upstream process, in which case the system should return to the basal state. This behaviour is more realistic in a biological system. It can be achieved by the addition of the following rule to the structure of the biological process: Add an inhibitory edge between Enzyme E and reaction r2, as shown in Figure 2.1.2aiii. This rule expresses the following biological understanding: If Protein A and Enzyme E are present in sufficiently large amounts, the net rate of all elementary molecular processes covalently modifying Protein A is faster than the net rate of molecular processes returning Protein A to a basal level. The Token Game simulation of the modified model is shown in Figure 2.1.2aiv. The system has, therefore, moved to a state where the covalently modified form of the protein is constantly at a sufficient amount for the activation of downstream signalling. The system will remain in this state until Enzyme E is depleted, after such time the system will return to a basal state, as shown in Figure 2.1.2av and Figure 2.1.2avi.
During model reconstruction, care was taken to build the model using the basic motif described above, or other simple sub-networks, obviously maintaining the state in \{0,1\} range and connected to the remaining part of the model only by read and inhibitory edges that do not need discrete activity levels. This does not, however, constitute the formal proof that under no circumstances will molecular species be assigned two discrete activity levels. Obtaining such proof would require the use of invariant analysis of Petri Net theory implemented in tools such as Integrated Net Analyzer (INA) (Roch & Starke, 2003) or Charlie (Heiner et al., 2015). While these tools are integrated with Snoopy, they do not allow the analysis of networks that have read and inhibitory edges. These edges are essential for the formulation of the transition rules, required for tackling the biological complexity, present in the large-scale model. Additionally, invariant analysis is prone to a combinatorial explosion challenge (Sun & Zhang, 2011) and would be difficult for a full scale model, even assuming such a model was constructed as a standard Petri Net. Consequently, in the absence of a practical way of formally verifying that no molecular species in the model presented in this dissertation can assume a state of more than one discrete activity level, a feature within the QSSPN solver has been used to enforce this condition within the model simulations (Fisher et al., 2013). For any molecular species, the maximal number of discrete activity levels has been set to one. If the simulator attempted to increase the number of discrete activity levels to a value of more than one, the number of activity levels was reset to one. Due to the way in which the model was carefully constructed, this condition was rarely, if ever, invoked and did not influence conclusions based upon the examination of the large number of simulation trajectories.
Section 2.1.3 – Model Construction

Section 2.1.3.1 – General Model Construction Process

The binary formalism, described in Section 2.1.2 – The Binary Model Formalism, has been used to represent literature on molecular events controlling translation initiation in mammalian cells. An illustration of this is shown earlier in Figure 1.1a and again in the information detailed in Section 1.1 – Cellular Translation Initiation and Section 1.2 – Key Points of Regulation & Upstream Signalling Pathways, as an executable, computational model capable of making predictions on the effects of molecular perturbations on the initiation of protein synthesis.

A comprehensive search of the PubMed (US National Library of Medicine, Bethesda, MD (US)) database has been conducted for the articles, describing molecular events, in the system under investigation. The use of the PubMed database was to ensure that only peer-reviewed work was included and that it can be referenced in a rigorous format by giving a PubMed Identifier (PMID) numbers. Given the differences in translation that exist between yeast and mammalian cells, such as in the composition of the Multi-Factoral Complex in yeast (Fletcher et al., 1999; Bieniossek et al., 2006; Sokabe et al., 2012), the articles describing the process in mammalian species have, without exception, been prioritised over articles describing the process in yeast. However, where the process was described only in experiments conducted in yeast, this work has been used to fill gaps in the literature. The scope of the literature search has been limited to the signalling pathways that are directly involved in the regulation of translation initiation in mammals. These pathways include the PI3K, mTORC1, mTORC2 and the MAPK pathways: ERK1/2, p38α, JNK1 and ERK5. It was decided that these signalling pathways would be extended back to a cell surface receptor and a literature search has been conducted accordingly.
The experimental biology literature articles, which provide evidence for individual interactions, are used to construct the rules of the binary model. As detailed in Section 2.1.2 – The Binary Model Formalism, every effort has been made to use the basic switch motif, in which the exchange of single discrete activity levels between forms of covalently modified proteins is controlled by both read and inhibitory edges. However, the use of more complex motifs has, sometimes, been needed when this was not sufficient to represent literature data in a meaningful and accurate way. As indicated in a number of studies (Wang et al., 2013; Green et al., 2014), there is a growing need for standardised nomenclature. In order to construct the model in a transparent and open way, protein, gene and micro RNA (miRNA) components of the model are labelled with Uniprot (The Uniprot Consortium, 2014), Entrez (NCBI Resource Coordinators, 2013) and miRBase (Kozomara & Griffiths-Jones, 2014) database identifiers, respectively, as shown in Supplementary Information E1 (available on the attached Flash Drive). Similarly, a list of the papers used in the construction of the model is available in Supplementary Information E2.

The focus of this section will now turn to describing the construction of the various parts of the model in the software Snoopy (Rohr et al., 2010; Marwan et al., 2012; Heiner et al., 2012). Section 1.2 – Key Points of Regulation & Upstream Signalling Pathways details the biological particulars of each key pathway included in the model and, as a result, this section will focus on illustrating the Petri Net versions of each model pathway and on describing the rationale behind the inclusion of various aspects of the pathways.

The model details the synthesis of nascent polypeptides and describes the assembly of 80S ribosome (Figure 2.1.3.1a). This component of the model was constructed using information taken from multiple peer-reviewed literature sources. As already shown in
Figure 1.2a, this component of the model comprises only a small part of the overall model produced during the course of this work.
Figure 2.1.3.1a – Petri Net Representation of Translation Initiation. Produced in the software Snoopy, the translation initiation pathway is represented using the Binary Model Formalism. This Petri Net model was extracted from the large-scale model produced over the course of this project and extends back to include the key points of regulation. Although this figure includes key points of regulation, it is analogous to Figure 1.1a.
One of the pathways most associated with the regulation of translation initiation is that of ERK1/2. As such, it was necessary to include a model of this pathway. It was decided to attempt to use an existing model of the activation of ERK1/2. The BioModels database (Li et al., 2010) was searched for such a model and one was selected (BioModels identifier: BIOMD0000000270). The model constructed by Schilling et al. (2009) concerned the activation of ERK1/2 via MEK1/2 in response to known ligand Erythropoietin (Seong et al., 2006; Nairz et al., 2011). In order to fully integrate this model, as a qualitative component of the larger model, it was necessary to reconstruct the model in the software Snoopy using the Petri Net formalism. This reconstruction of the model involved expressing each reaction given in the working version of the model in the software Copasi (Hoops et al., 2006; Mendes et al., 2009) to a Petri Net (Figure 2.1.3.1b). To put this component in the same notation as the rest of the model it was necessary to ensure that these molecular interactions were expressed using the Binary Model Formalism. In keeping with the qualitative nature of the model, all molecular amounts were replaced with a marking of one discrete activity level, which as elsewhere in the model represented an amount of a particular molecular species needed to provoke a biological response. Similarly, all rate constants within the model were expressed as one.
Figure 2.1.3.1b – Petri Net Representation of the ERK1/2 Signalling Pathway. The ERK1/2 signalling pathway is a canonical component of the cellular signalling network and is integral to the regulation of translation initiation. The model of the activation of ERK1/2 was taken from Schilling and colleagues (2009). Once activated, ERK1/2 was able to phosphorylate numerous downstream components within several other signalling pathways. Within this model, several inhibitors affecting ERK1/2 activation were incorporated: 5Z-(7)-Oxozeanol (FR1), PD184352 (PD1) and PD98059 (PD9). This model was produced as a Petri Net using the software Snoopy.
Once reconstructed, the integration of this model with the full model was conducted. This was first achieved by linking an early point within the model to an existing model component. In this case it was achieved by linking the activation of the MEK1/2 kinase, Raf-1 to Ras (MacDonald et al., 1993; Diaz-Meco et al., 1994; Fabian et al., 1994; Leevers et al., 1994; Stokoe et al., 1994; Jelinek et al., 1996). The model already contained a number of important ERK1/2 substrates, including MNK2 and RSK1, so it was possible to ensure the model functioned within the full model. The final stage of this integration occurred during the stages of model refinement. This included the modification of the model to take into account the activation of ERK1/2 through additional pathways. These pathways included the Nuclear Factor-κB-induced kinase, Tumour Progression Locus 2 (Tpl2) (Dumitru et al., 2000; Eliopoulos et al., 2003; Rousseau et al., 2008; López-Pelaéz et al., 2012; Roget et al., 2012; Hirschhorn et al., 2014) and PKCζ (Monick et al., 2000; Lesseux et al., 2008).

The other MAPK pathways were included in the model as they form part of the network that regulates the translational machinery. The primary role of p38α in the regulation of translation is given in Section 1.2.2 – eIF4E Phosphorylation and The Role of ERK1/2 & p38. This pathway (Figure 2.1.3.1c) was reconstructed in the model from published, peer-reviewed literature. Reconstructing this pathway necessitated including not only the kinases involved in the activation of p38α, but also to extend this pathway to a cell-surface receptor. In the case of p38α, this receptor was TLR4. It was not necessary to include this receptor per se, as a basal level of p38α has been reported in a number of cell-types, including the type used to experimentally validate the predictions generated by the model produced during the course of this work (detailed in Chapter 3 – Experimental Validation of Model Predictions) (Hsieh & Papaconstantinou, 2002; Carrozzino et al., 2009; Naidu et al., 2009; Bouazza et al.,
2014). However, for completeness and to allow the effect of LPS stimulation to be studied in future, it was included.
The p38α signalling pathway regulates translation through controlling eIF4E phosphorylation and mRNA stability. This kinase plays a role in regulating many other processes in the cell. This pathway is extended back to Toll-Like Receptor 4. Key points of cross-talk between other signalling pathways are also included. These include the negative regulation of Protein Kinase C ζ (PKCζ) and the inhibition of p38α by AKT1. This model was created in the software Snoopy in the Petri Net notation.
The JNK1 pathway (Figure 2.1.3.1d) was constructed in much the same way, in that the pathway was extended back to a cell-surface receptor and that a basal level of activity (Javelaud et al., 2003; Sabapathy et al., 2004; Carrozzino et al., 2009; Cui et al., 2009; Zhu et al., 2016) was maintained. With regards to the activation of JNK1 the model included the requirement, for the presence of the metabolite, ceramide (Mackichan & DeFranco, 1999; Medvedev et al., 1999; Monick et al., 2000; Brandt et al., 2010). While not directly relevant to the regulation of translation initiation, the requirement for ceramide was included as it offered an opportunity to begin to understand the links between the cellular signalling and metabolic networks.
Figure 2.1.3.1d – Petri Net Representation of the JNK1 Signalling Pathway. JNK1 has an emerging role in the regulation of translation initiation. The main, yet not fully elucidated, role concerns regulating eIF4E phosphorylation. This pathway was extended back to the Interferon Receptor. As with the p38α pathway, this receptor was not necessary as a basal level of JNK1 was maintained in the model. This model was extracted from the large-scale model of translational regulation and was constructed in Petri Net notation using the software Snoopy.
As already stated, mTORC1 is fundamental to the regulation of translation initiation. This pathway, within the model, was extended back through the PI3K pathway to the Insulin Receptor (Figure 2.1.3.1e). While mTORC1 is activated by AKT1, it serves as a point of integration for several other key signalling pathways. It was necessary to include these pathways in the activation of this signalling component as it allowed the integration of a signal to produce more complex patterns of biological behaviour. The complexity of this pathway was further augmented by the addition of feedback loops including the S6K1-dependent inhibition of mTORC2 (Julien et al., 2010) (such mechanisms are discussed in more detail in Section 3.1.3 – Interplay Between mTORC1 & AKT1). The inclusion of such feedback served to ensure that the pathway was better able to simulate a biological system responding to an external perturbation.
Figure 2.1.3.1e – Petri Net Representation of the PI3K and mTORC1 Signalling Pathways. Created using the software Snoopy, this Petri Net model contains molecular interactions and species concerning the PI3K and mTORC1 signalling pathways. Included in this model are key points of cross-talk from other signalling pathways, including ERK1/2, p38α and JNK1. The PI3K pathway is one of the main activators of mTORC1.
One of the main features of the model, as a whole, is the implementation of cross-talk between the signalling pathways. The main kinases found to be responsible for cross-talk were AKT1 and ERK1/2. With regards to AKT1, this kinase was found, through ill-defined mechanisms, to inhibit both p38α (Gratton et al., 2001; Liao & Hung, 2003; Shi et al., 2005; Marderosian et al., 2006; Yu et al., 2006; Taniike et al., 2008) and JNK1 (Kim et al., 2001; Lu et al., 2002; Park et al., 2002; Aikin et al., 2004; Hui et al., 2005; Wang et al., 2006; Wang et al., 2007b). ERK1/2 were shown to inhibit ERK5 (Sarközi et al., 2007) and JNK1 (Carter et al., 1998; Reardon et al., 1999). These inhibitory interactions were included in the construction of the model as such interactions are responsible for the complex patterns of behaviour visible in biological systems.

Section 2.1.3.2 – Specific Examples

Literature-based, bottom-up mechanistic model building is a manual process, where expert curators make decisions about how best to represent the content of literature detailing experimental biology in a particular framework. While attempts have been made to formalise this process in particular domains, for example in GSMN (Thiele & Palsson, 2010), the key role of an expert modeller in making decisions remains. This section presents the modelling approach used in this work by giving specific examples of how particular literature articles where represented.

The bottom-up approach to model construction used in this project was achieved in three stages: i. manual creation, ii. model validation, iii. model refinement. The first stage of manual creation involved the reconstruction of the translation initiation machinery in mammalian cells. Such an approach initially required an understanding of the main steps of the process, for example the recycling of the Post-TCs, formation of the elF4F complex and
mRNA circularisation, and the joining of the 60S ribosomal subunit, as outlined in Figure 1.1. This basic understanding came from review articles such as Jackson et al. (2010). In order to represent this process as accurately as possible it was then necessary to undertake more targeted literature searches to gain an appreciation of how the various stages are undertaken. For example while Jackson et al. (2010) were able to describe the process of recycling Post-TCs, the role and regulation of eIF6 in this process is largely overlooked by this work. Consequently, it was necessary to carry out more targeted searches of existing literature to gain an appreciation of the role of this initiation factor. As eIF6 functions to prevent aberrant binding of the 60S subunit (Valenzuela et al., 1982; Raychaudhuri et al., 1984; Si et al., 1997; Wood et al., 1999) to the assembling 48S complex it has a clear role in the recycling process. However, at this stage it should be clear that some form of regulation is necessary to ensure that this factor does not prevent the formation of the 80S ribosome when required. To this end, it was necessary to investigate what is currently known about mechanisms of eIF6 regulation. Given the work of Ceci et al. (2003) has shown that eIF6 is regulated through phosphorylation by PKC βII, it was possible to represent this stage using the basic switch motif described earlier.

The second stage of model construction concerned model validation. Although this stage had little to do with the construction of the translational regulation model per se, it was necessary to validate the model against available literature in order to further refine the connectivity of the model (Pattyn et al., 2013). To this end, various chemical inhibitors were incorporated into the model (as detailed in Section 2.1.4.2 – Evaluation of the Predictive Power of the Model). The effects of these chemical inhibitors on various model components were investigated. Documented effects were recorded into benchmark
datasets. A key point of regulating the formation of the eIF4F complex involves the sequestration of eIF4E by 4E-BP1 (Marcotrigiano et al., 1999). The mTORC1-dependent phosphorylation of 4E-BP1 inhibits such interactions and allows the eIF4F complex to form (Beretta et al., 1996; von Manteuffel et al., 1996; Burnett et al., 1998; Gingras et al., 1998; Marcotrigiano et al., 1999). However, the model simulation results failed to account for a link between 4E-BP1 phosphorylation and treatment with the inhibitor SP600125 (Bain et al., 2007; Ito et al., 2011) given in the benchmark dataset, as shown in Figure 2.1.3.2. The third stage of model construction was refinement. In order to account for this effect recorded in the benchmark dataset, it was necessary to establish a mechanism by which the JNK1 inhibitor, SP600125 (Wang et al., 2004), could affect 4E-BP1 phosphorylation. A more detailed search of the literature concerning JNK1 substrates found that MSK1 was also capable of phosphorylating 4E-BP1 (Liu et al., 2002; Aggeli et al., 2006; Teng et al., 2009). The last two stages of model construction can be viewed as iterative as the model refinement process precipitated a further round of model validation.
Figure 2.1.3.2 – Example of Model Construction Process. As discussed in Section 1.1.3 – \( \text{eIF4F Components & mRNA Circularisation} \) and Section 1.2.1 – \( \text{eIF4E-Binding Proteins & The mTORC1 Signalling Pathway} \), the availability of eukaryotic Initiation Factor 4E (\( \text{eIF4E} \)) is regulated by the eIF4E-Binding Protein 1 (\( \text{4E-BP1} \)). Mammalian Target of Rapamycin Complex 1 (\( \text{mTORC1} \)) phosphorylates \( \text{4E-BP1} \) (\( \text{p4E-BP1} \)) and consequently, eIF4E is then free to form the eukaryotic Initiation Factor 4F complex. The dephosphorylation of \( \text{p4E-BP1} \) is mediated by Protein Phosphatase 2A (\( \text{PP2A} \)) (Peterson et al., 1999; Pham et al., 2000; Janzen et al., 2011). This example of an Extended Petri Net model is constructed using the Binary Model Formalism (as described in Section 1.4.4 – \( \text{The Petri Net Formalism} \) and Section 2.1.2 – \( \text{The Binary Model Formalism} \)). The frequency data given above was taken from the model produced during the course of this project and calculated using the software, Reachfq. The initial model determined that, despite mTORC1 inhibition with Rapamycin being noted as negatively regulating \( \text{p4E-BP1} \) formation in the Benchmark Dataset, no inhibition was present. Such a finding led other parts of the Benchmark Dataset to be investigated. The SP600125 model was identified and enabled a more targeted literature search to be conducted in order to elucidate a link between JNK1 and the phosphorylation of 4E-BP1. This literature search revealed a link between JNK1 and a kinase noted to be involved in 4E-BP1 phosphorylation, Mitogen- and Stress-Activated Kinase-1 (\( \text{MSK1} \)) (Liu et al., 2002). This process of initial construction followed by re-evaluation of the connectivity of the model through targeted literature searches was used throughout the project to refine the model.
Section 2.1.3.3 – Creation of the Benchmark Datasets

Evaluation of the predictive power of the model is key for future applications. Although no model is expected to provide absolute accuracy, there is always a risk that model predictions are wrong. To make a model useful, it is essential to comprehensively evaluate the predictive power to give future user’s guidance about the likelihood of obtaining correct predictions and observing different types of errors, such as false predictions of positive or negative outcomes. Ultimately, the model is a decision making tool with predictions being used to guide future experimental work. From this perspective, informing the user about the risks associated with pursuing model predictions should be regarded with equal importance as the prediction itself. For example, the user may decide that in the absence of hypotheses for the next step of experimental molecular biology work on a complex system, following model predictions is useful, even if there is a high probability of a particular prediction being false. The user may be of the opinion that predictions from a literature-based model are more likely to be accurate than a guess. Even if such a prediction is false, the user would still learn from the experimental outcome. On the other hand, a model with low predictive power would not be applicable for the design of a clinical trial, where subjects would be put at risk and where considerable financial investment was at stake.

Therefore, in this project, the development of the model was accompanied by the creation of a benchmark dataset for the evaluation of predictive power. The literature has been systematically examined to identify the effects of commonly used chemical inhibitors (as shown in Table 2.1.3.3a) on the activity of different molecular species in the model. These chemical inhibitors were chosen because usually the effects of these on cell signalling components are well documented. For each of the inhibitors, original research articles have been gathered, which describe whether the application of this inhibitor leads to an increase
or decrease in the activity of a molecular species present in the model. The PMID identifier of these articles are included in the dataset. Care was taken to ensure that data was collected on different inhibitors, applied under similar experimental conditions. For example, all data was collected in the absence, or presence, of lipopolysaccharide stimulation, to ensure that a benchmark dataset was created, suitable for evaluating the predictive power of a model. It should be noted that the systematic evaluation of the literature to create the benchmark datasets was of comparable effort to the construction of the model itself. This dataset is unique and represents a major outcome of the project. The benchmark dataset is presented in the subsequent results sections with the Final Benchmark dataset being given in Section 2.2.4 – Evaluation of the Predictive Power of the Final Model.
<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>Target</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIX02189</td>
<td>MEK5</td>
<td>Tatake et al., 2008</td>
</tr>
<tr>
<td>CGP57380</td>
<td>MNK1/2</td>
<td>Tschopp et al., 2000</td>
</tr>
<tr>
<td>GSK2334470</td>
<td>PDK1</td>
<td>Najafov et al., 2011</td>
</tr>
<tr>
<td>Ku0063794</td>
<td>mTORC1/2</td>
<td>García-Martínez et al., 2009</td>
</tr>
<tr>
<td>LY294002</td>
<td>PI3K</td>
<td>Vlahos et al., 1994</td>
</tr>
<tr>
<td>PD184352</td>
<td>MEK1/2-ERK1/2</td>
<td>Allen et al., 2003</td>
</tr>
<tr>
<td>PD98059</td>
<td>MEK1/2-ERK1/2 &amp; ERK5</td>
<td>Alessi et al., 1995;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dudley et al., 1995</td>
</tr>
<tr>
<td>Rapamycin</td>
<td>mTORC1</td>
<td>Sabers et al., 1995</td>
</tr>
<tr>
<td>SB203580</td>
<td>p38</td>
<td>Badger et al., 1996</td>
</tr>
<tr>
<td>SP600125</td>
<td>JNK1</td>
<td>Wang et al., 2004</td>
</tr>
<tr>
<td>XMD8-92</td>
<td>ERK5</td>
<td>Yang et al., 2010</td>
</tr>
<tr>
<td>5Z-7-Oxozeaenol</td>
<td>TAK1, MEK1/2</td>
<td>Ninomiya-Tsuji et al., 2003;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Choo et al., 2006;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Windheim et al., 2007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Acuña et al., 2012</td>
</tr>
</tbody>
</table>

Table 2.1.3.3a - Chemical Inhibitors used in the Model. The target given above is the marketed cellular target and does not take into account off-target effects. Note that 5Z-7-Oxozeaenol was only used in the LPS model and Ku0063794 was replaced by Rapamycin in the binary model.
Section 2.1.4 – Simulation Method

Section 2.1.4.1 – Monte Carlo Simulations

Both example simulations shown in Section 2.1.2 – The Binary Model Formalism and all simulations leading to the results presented in Section 2.2 – Results have been performed with the QSSPN solver (described in Section 1.4.5 – The Quasi-Steady State Petri Net Algorithm) (Fisher et al., 2013). While simulations presented in this chapter could have been run in other tools, such as Snoopy, the QSSPN solver was used to make the resulting model amenable to integration with the Flux Balance Analysis of GSMN (presented in Chapter 5 – Integration of the Translational Control Model with a Genome-Scale Metabolic Network of a RAW 264.7 Cell) – a unique feature of QSSPN not available in other tools. The QSSPN solver performs qualitative simulations using the Gillespie algorithm (described in Section 1.4.1 – Exact Stochastic Simulations with the Gillespie Algorithm) with all reaction propensity functions set equal to one (Ibid). It should be noted that this algorithm was used to generate alternative sequences of molecular interactions occurring in the binary formalism model, rather than the modelling of the dynamics of molecular amounts in a quantitative model. Consequently, the propensity function could only assume two values: zero if the molecular interaction was not enabled, or one, if the molecular interaction was enabled. Furthermore, the Gillespie algorithm time was used exclusively to order the sequence of molecular interaction occurrences i.e. to determine which molecular interactions occurred either earlier or later. It was not interpreted as physical time. Moreover, the Gillespie algorithm used time as a stop condition for the simulations. For example, trajectories were run for 100 arbitrary time units. Again, this number was not interpreted as a physical time.
The qualitative behaviour of the model has been explored by the simulation of an ensemble of trajectories. Each trajectory was started from the same initial state and the only difference between individual trajectories was exclusively due to the pseudo-random numbers generated in the Gillespie algorithm simulation. Therefore, trajectories represent the alternative sequences of molecular interactions possible, given the literature knowledge expressed by the rules of the binary model formalism. These alternative scenarios of model behaviour were then analysed by two different approaches implemented in Reachfq and Eventfq scripts (developed by A. Kierzek as part of the QSSPN project).

The Reachfq script calculates the number of trajectories where a particular molecular species reached a state of \( k \) discrete activity levels. In simulations with the binary model, \( k \) was set to one. This script reads every trajectory in the ensemble and moves sequentially through the states from time 0 to the end. If this condition has been fulfilled, the script stops analysing that particular trajectory, increases the counter, and moves to the next trajectory. If the end of the trajectory is reached and the condition was not satisfied, the counter remains unchanged and the script analyses the next trajectory in the ensemble.

The Eventfq script calculates, for each molecular interaction in the model, the total number of firings in all trajectories in the ensemble. Each trajectory is read in turn. The first column of the trajectory file indicates which molecular interaction firing changed state. The script reads each field in this column, in turn, and whenever the reaction name is encountered, the counter for this molecular interaction is increased. The result is the total number of firings, across all trajectories calculated, for each molecular interaction in the model.
Section 2.1.4.2 – Evaluation of the Predictive Power of the Model

Predictive power of the model was evaluated by simulating the application of inhibitors, accounted for in the benchmark datasets, and comparing the simulation results with experimentally observed effects. First, the Wild Type model, representing a situation where no inhibitors were applied, was simulated and 1,000 simulation trajectories were collected. The Reachfq script was then used to calculate the number of trajectories reaching each of the key effectors from the important signalling pathways featured in the model, and described in Section 1.2 – Key Points of Regulation & Upstream Signalling Pathways. The activity changes of these effectors, as a result of the application of inhibitors, were recorded in the benchmark datasets. Subsequently, for each inhibitor, another variant of the model representing the application, was created. The Monte Carlo simulation was then run and 1,000 simulation trajectories were collected. The effect of the inhibitor on a particular model effector has been evaluated by comparing the number of trajectories reaching this effector, in the Wild Type model, and the model accounting for the action of a particular inhibitor. In the simulation, the inhibitor was considered to have had an effect on a particular effector, if the number of trajectories reaching this molecular species were significantly different between the Wild Type and inhibitor models. A difference in the number of trajectories reaching a particular effector was considered significantly different if the Wild Type and inhibitor numbers displayed no overlap in the 99.99% Binomial Confidence Intervals (Marshall & Olkin, 1985). These Intervals were calculated in the software R (R Core Development Team, 2013) using the binconf function, which is part of the Hmisc package (Brant et al., 1991-2016). The application of confidence intervals, in comparing the numbers of Monte Carlo trajectories exhibiting behaviours of interest, is an
accepted methodology in the Statistical Model Checking field of computer science (Kwiatkowska et al., 2011).

The predictive power of the model was evaluated using the well accepted approach of constructing Confusion Matrices. The Confusion Matrix classifies comparisons of the model and simulation results to one of the four possible outcomes:

- True Positive (TP) – Positive Experiment & Positive Model
- False Positive (FP) – Negative Experiment & Positive Model
- True Negative (TN) – Negative Experiment & Negative Model
- False Negative (FN) – Positive Experiment & Negative Model

In these outcomes, a Positive result denotes a difference in the frequency of trajectories, between the ‘Wild Type’ and inhibitor models, resulting in the activation of a particular effector. A Negative outcome connotes that the incorporation of the inhibitor resulted in no change in the frequency of Monte Carlo simulation trajectories bringing about the activation of a particular effector. A True or False designation for each effector is given based upon whether or not the simulation result is in agreement with the benchmark dataset.
Using the results of the Confusion Matrices, the Matthews’ Correlation Coefficient (MCC) (Matthews, 1975) was calculated for each model and each inhibitor. This calculation was done using the equation (Equation 2.1.4.2a) shown in Gowthaman & Agrewala (2008) and Schloss & Westcott (2011):

\[
MCC = \frac{((TP \times TN) - (FP \times FN))}{\sqrt{((TP + FN) \times (TN + FP) \times (TP + FP) \times (TN + FN))}}
\]

Equation 2.1.4.2a

The MCC value ranges from between +1 and -1 and has three points for interpretation (Shi et al., 2010; Jurman et al., 2012). As the MCC value approaches +1, this indicates increasing agreement between the model predictions and what has been experimentally shown. A value of +1 would, therefore, indicate complete agreement. Conversely, a value approaching -1 indicates increasing disagreement between the model predictions and the experimental data contained within the benchmark datasets (Shi et al. 2010). An illustration of this point can be found in a system in which a decrease in the phosphorylation of a particular model component is predicted, while the experimental evidence demonstrates an increase. The final point for interpretation is 0. As the MCC value approaches this, it indicates that the predictions generated by the model are essentially random (Shi et al., 2010; Jurman et al., 2012). One point to note with regard to MCC values is that in a system yielding limited numbers of both FP and TP values, the resulting value tends to be high (Baldi et al., 2000).

An overall MCC value was also calculated, based upon the total numbers, pooled from each inhibitor model Confusion Matrix. This was necessary to ensure that predictive power was calculated from sufficient data points to overcome the classic argument that small sample
sizes lead to the validity of a correlation coefficient being called into question (Quenouille, 1949). Therefore, this global MCC value was used to define the overall predictive power of the model, while the MCC values for the individual inhibitor models were used to re-assess the connectivity of the model, through targeted literature searches, during refinement.

In order to confirm the validity of the overall MCC values, two additional tests were also carried out. The first involved the calculation of the Approximate Coefficient (AC) (Burset & Guigó, 1996) and the calculation was carried out using the variables contained within Confusion Matrices using the following equations (Burset & Guigó, 1996; Baldi et al., 2000) (Equation 2.1.4.2b and Equation 2.1.4.2c):

\[
ACP = \frac{1}{4} \times \left( \frac{TP}{TP + FN} + \frac{TP}{TP + FP} + \frac{TN}{TN + FP} + \frac{TN}{TN + FN} \right)
\]

Equation 2.1.4.2b

\[
AC = (ACP - 0.5)
\]

Equation 2.1.4.2c

The second additional test involved subjecting the overall Confusion Matrix of each binary model to a Chi Square Test. This was done by comparing the observed number of Confusion Matrix outcomes to an expected number for each outcome (Baldi et al., 2000). In this case the expected number of TP, TN, FP and FN outcomes can be viewed as being the total number of data points divided by four (as the number of possible outcomes). In simple terms, if the total number of data points is 100, then it would be reasonable to assume that each of the four possible outcomes in the Confusion Matrix would occur 25 times. If this were the case, the resulting MCC value would be equal to zero. By performing this Chi Square analysis, the MCC value obtained is tested for a significant difference from an MCC
value of zero \textit{(Ibid)}. Both this and the calculation of the AC value were limited to the evaluation of the models constructed using the binary model formalism as these were inherently more meaningful, as discussed in Section 2.1.2 – The Binary Model Formalism.

Section 2.1.4.3 – Synchronous Simulation Methods

The Gillespie algorithm (described in Section 1.4.1 – Exact Stochastic Simulations with The Gillespie Algorithm), which is the main simulation approach used in this work (detailed in full in Section 2.1.4.1 – Monte Carlo Simulations), is a stochastic and asynchronous method. This means that the firing of individual Petri Net molecular interactions is not synchronised within the simulation timesteps. In each iteration of the algorithm, only one molecular interaction is chosen randomly and fired. The qsspn solver implements an alternative, synchronous simulation approach in which firing of molecular interactions is synchronised. Rather than firing only one molecular interaction in a particular timestep, the algorithm attempts to fire all molecular interactions, which are enabled at this timestep. This however, frequently leads to a situation where two or more molecular interactions attempt to update the same molecular species. If this molecular species contains only one discrete activity level, the synchronised execution of more than one molecular interaction would lead to negative numbers of discrete activity levels, which must not occur. This conflict between more than one molecular interaction requesting discrete activity levels from the same molecular species is resolved by the stochastic method. The order in which molecular interactions connected to the same molecular species remove discrete activity levels from that molecular species is randomised. Molecular interactions consume discrete activity levels until none are left. In other words, if multiple enabled molecular interactions are connected to the same molecular species, and the number of discrete activity levels
associated with this species is smaller than the number of molecular interactions, not all of
these interactions will be fired and those which are fired are chosen randomly. Therefore,
\textbf{this method is stochastic and synchronous.}

The $\tau$-Leap method (already discussed in \textbf{Section 1.4.1 – Exact Stochastic Simulations with
the Gillespie Algorithm}), proposed by Gillespie (2001), is a stochastic, synchronous
algorithm frequently used in Systems Biology. Briefly, in each timestep $\tau$, the number of
firings, $k_\tau$, is randomly generated from a Poisson distribution for each molecular
interaction, $0\tau$, in the system (Gillespie, 2001; Gillespie, 2007). Subsequently, each
molecular interaction is fired $k_\tau$ times. This method does not resolve the problem of
molecular interactions simultaneously accessing the same molecular species and, ultimately,
it is likely that such a situation would lead to negative numbers of discrete activity levels in
the qualitative model produced during the course of this project (Cao \textit{et al.}, 2006).
Consequently, this method is not applicable to the binary model produced in this project
and so will not be discussed further.

The synchronous stochastic approach is attractive because it is less computationally
demanding than the asynchronous method. Since multiple molecular interactions are
executed simultaneously, smaller number of simulation timesteps may be required to reach
a biological outcome of interest. Where multiple enabled molecular interactions compete
for a substrate, one interaction is selected at random. Furthermore, simulation trajectory
files are smaller, requiring less storage space. This reduced file size makes the analysis with
the Reachfq script faster. Given this, the synchronous stochastic approach implemented in
the qsspn solver has been considered as an alternative to the Gillespie algorithm. However,
it has been established that the asynchronous method is better in terms of generating
biologically meaningful hypotheses and, therefore, the synchronous algorithm was not used. A case study detailed in Section 2.1.4.3.1 – Case Study Identifying Differences in the Asynchronous Simulation Approach & the Synchronous Methods gives an example to show why this was the case.

Section 2.1.4.3.1 – Case Study Identifying Differences in the Asynchronous Simulation Approach & the Synchronous Methods

A small model (Figure 2.1.4.3.1a) has been extracted from the full molecular network. This small model contains 65 molecular interactions and 53 molecular species, all relating to eIF4E availability and phosphorylation.
This Petri Net model was constructed in the software Snoopy and only encompasses the molecular interactions relating to the availability of eukaryotic Initiation Factor (eIF) 4E (eIF4E). All interactions present in this model can be found in the large-scale model of translational regulation. This model is constructed of 53 molecular species and 65 molecular interactions.
The simulation of 1,000 trajectories, of which an example is shown in Figure 2.1.4.3.1b, with the Gillespie algorithm, followed by the application of the Reachfq script yielded the following results:

<table>
<thead>
<tr>
<th>Model Effector</th>
<th>Frequency Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phosphorylated eIF4E</td>
<td>0.472</td>
</tr>
<tr>
<td>Phosphorylated MNK1</td>
<td>0.472</td>
</tr>
<tr>
<td>Phosphorylated MNK2</td>
<td>0.502</td>
</tr>
<tr>
<td>Phosphorylated AKT1</td>
<td>0.511</td>
</tr>
<tr>
<td>Phosphorylated JNK1</td>
<td>0.556</td>
</tr>
</tbody>
</table>
Figure 2.1.4.3.1b - Example Trajectory from the Simple Model Generated Using the Asynchronous Simulation Approach. This approach uses the Gillespie algorithm to generate a sequence of enabled molecular interactions. As this model incorporated molecular interactions only concerning eukaryotic Initiation Factor (eIF) 4E phosphorylation and availability, this trajectory ends with eIF4E phosphorylation (marked with the green L). This trajectory starts with the Raf-1-mediated phosphorylation of the Mitogen-Activated Protein Kinase (MAPK) Kinase, MEK1 (marked with the green A). B concerns the activation of ERK1. A prerequisite for MNK2 phosphorylation (D) is the activation of JNK1 through phosphorylation (C). The phosphorylation of eIF4E occurs through MNK2 and requires the PKCα-dependent eIF4G phosphorylation (E and F). The availability of eIF4E is regulated through sequestration by 4E-BP1. This sequestration is negative regulated by both mTORC1 and MSK1 (K). The activation of mTORC1 (J) is dependent upon Rheb while MSK1 activation is carried out through p38α (H). p38α phosphorylation occurs through MKK3 (G). The model includes the phosphorylation of AKT1 (I). In this model, AKT1 serves to inactivate JNK1. This model is shown in detail in Figure 2.1.4.3.1a.
By comparison, the simulation of 1,000 trajectories with the synchronous stochastic approach, followed by the application of the Reachfq script, yielded the following results:

<table>
<thead>
<tr>
<th>Model Effector</th>
<th>Frequency Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phosphorylated eIF4E</td>
<td>0.515</td>
</tr>
<tr>
<td>Phosphorylated MNK1</td>
<td>1.000</td>
</tr>
<tr>
<td>Phosphorylated MNK2</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated AKT1</td>
<td>1.000</td>
</tr>
<tr>
<td>Phosphorylated JNK1</td>
<td>1.000</td>
</tr>
</tbody>
</table>

An example trajectory yielded by this method is given in Figure 2.1.4.3.1c
Figure 2.1.4.3.1c - Example Trajectory from the Simple Model Generated Using the Synchronous Simulation Approach. This model is shown in detail in Figure 2.1.4.3.1a. This method fires all enabled molecular interactions in each simulation time interval. In the trajectory shown above this is illustrated. All molecular species shown in red are updated first. This is followed by the updating of dark blue, green, light blue and pink molecular species. Given this trajectory, it is clear that MNK2 phosphorylation is not permitted during model simulations with this method. While phosphorylated JNK1 (pJNK1) does form (shown in red/green), it is removed once the dual-phosphorylated form of AKT1 (ppAKT1) forms. This is a result of how the synchronous simulation approach removes the temporal nature of signalling pathways imparted by the waiting time calculation of the Gillespie algorithm. Molecular species consumed during the molecular interactions are not coloured, rather only molecular species formed during the firing of interactions. An exception of this is pJNK1 which is central to the illustration of the limitations to this approach.
It can immediately be seen that the simulation with the Gillespie algorithm provided examples of the sequence of molecular events leading to MNK2 phosphorylation, while the synchronous stochastic approach predicted that the species cannot be phosphorylated. According to molecular biology literature, the phosphorylation of MNK2 is possible (Ueda et al., 2004). Thus, the synchronous stochastic approach misses biologically meaningful trajectories of molecular events and, consequently, it should not be used in this work, despite the advantages this approach has, in terms of computational performance.

Closer examination of the synchronous stochastic simulation results show that the biologically realistic behaviour arises as a result of the feedback surrounding the kinase, JNK1. This kinase is inactivated by AKT1 (Kim et al., 2001; Lu et al., 2002; Park et al., 2002; Aikin et al., 2004; Hui et al., 2005; Wang et al., 2006; Wang et al., 2007b). This synchronous approach exacerbates this feedback by removing the effects of waiting time, implemented in the asynchronous approach via the Gillespie algorithm. As all enabled molecular interactions fire as each simulation timestep occurs, this method assumes that each interaction takes the same length to occur. In actual biological systems this is not the case as reactions have different rates. While many of these rates are not known, the asynchronous approach allows the random sampling of alternative molecular sequences where molecular interactions do not occur simultaneously. The example above demonstrates that this allows the identification of a known biological outcome. Since there are many feedback loops within the large model, this example illustrates a general problem with the synchronous approach, rather than an example specific to this small example. As a result of these deficiencies, the synchronous stochastic approach was taken forward but
only in a limited capacity (detailed in Section 5.3.2 – Proposed New Approaches and Section 5.3.2.2 – Synchronous Method).

Finally, another example of a synchronous simulation algorithm is the integration of Ordinary Differential Equations, which is a workhorse of computational Systems Biology for the simulation of quantitative models. As follows from the description in Section 1.4.2 – Ordinary Differential Equations (ODEs), all molecular species in the model are updated synchronously, but random numbers are not used to resolve competition between molecular interactions consuming the same species. Instead, molecular amounts are described as concentrations rather than numbers of molecules, but competing interactions decrease the concentrations to numbers possibly smaller than one. However, these numbers remain positive. Very sophisticated adaptive timestep strategies are used, within ODE solvers, to ensure that small molecular concentrations are simulated without leading to negative numbers (Press et al., 2007). In terms of the discussion within this chapter, the ODE simulation algorithm is therefore synchronous and deterministic. It generates only one trajectory representing the time evolution of the system. While computationally even more attractive than the synchronous stochastic simulation approach, the application of this method is limited to fully parameterised, quantitative models. An attempt to run simulations of models presented here led to a trajectory which did not have a clear biological interpretation. As discussed earlier, quantitative parameterisation of the large-scale molecular network model, described here, is not possible due to a lack of quantitative measurements in existing scientific literature. However, as this approach was computationally attractive, it was applied to the full model. The results are given in Section 5.3.2.1 – Ordinary Differential Equation Method.
Section 2.1.5 – Investigation of the Interaction Between eIF4E Phosphorylation and mTORC1 Inhibition

Once the binary model had been refined and the predictive power raised to a sufficient level, it was necessary to correct the decrease in eIF4E phosphorylation noted in the binary model (as discussed in Section 2.2.2 – Re-Evaluation of the Binary Model). Moreover, a mechanism by which eIF4E phosphorylation is increased in response to the inhibition of mTORC1 was necessary.

Several unsuccessful attempts were made to elucidate such a mechanism before the necessary circumstances surrounding this increase were identified. The methods detailed in the subsequent section were used to achieve these objectives. Both efforts involved extracting information from trajectories, such as using Eventfq to establish the number of times individual molecular interactions fired and mining the sequence of molecular interactions occurring up to and including the phosphorylation of eIF4E. The latter required a script created by D. Taylor using the Python programming language (van Rossum, 2012).

Since neither Eventfq nor the use of Invariant files yielded a clear mechanism, it was decided that a new approach was necessary. A smaller model was constructed that only encompassed the molecular interactions surrounding eIF4E phosphorylation and mTORC1 activity. The small scale and scope of this model allowed for easier identification of molecular species upon which eIF4E phosphorylation is dependent. The small model was then changed in various ways so as to find a reason for the decrease in eIF4E phosphorylation.
The identification of factors in the small model permitted further, more targeted, work to be undertaken on the larger model. The critical effectors for eIF4E phosphorylation were then identified and one-by-one, or in groups, they were subject to removal. Following the removal of effectors, a new set of simulations was conducted.

**Section 2.2 – Results**

Once the model was constructed it was simulated with QSSPN calculating 1,000 trajectories. Reachfq was then used to interrogate these trajectories in order to establish the frequency of trajectories reaching a particular model effectors. The mode of model construction enabled the data generated using Reachfq to be easily correlated to the components of the model. Consequently, points of refinement were easily identified and, after going back to exiting literature, were honed without difficulty.

**Section 2.2.1 – Binary Model of LPS Stimulation**

The binary model formalism was adopted because of the arbitrary discretisation and unrealistic oscillatory nature of the initial model in preliminary work. This behaviour was produced by the unrealistic assumption that the switch-like activation and deactivation of enzymes is caused by equal kinetic parameters. The binary model formalism and rationale for using this is described above in **Section 2.1.2 – Binary Model Formalism**. In order to introduce the rule-based transitions, it was necessary to further extend the granularity of the signalling pathways contained within the model by elaborating on the annotation of these pathways back to the receptor-ligand interactions. As several of the pathways were related to stress responses, the LPS-mediated Toll-Like Receptor 4 signalling (Poltorak et al., 1998; Chow et al., 1999; Hoshino et al., 1999) was one such pathway to be modified.
Although this preliminary work is not shown, the predictive power of this initial model was found to be MCC = 0.2354.

The following table (Figure 2.2.1a) shows an excerpt of the 99.99% Binomial Confidence Intervals calculated for the Reachfq analyses of the frequency of activation of model effectors in a binary model incorporating LPS stimulation.
<table>
<thead>
<tr>
<th>Effector</th>
<th>Frequency</th>
<th>Upper</th>
<th>Lower</th>
<th>Frequency</th>
<th>Upper</th>
<th>Lower</th>
</tr>
</thead>
<tbody>
<tr>
<td>ppERK1</td>
<td>1.000</td>
<td>1.000</td>
<td>0.989</td>
<td>0.000</td>
<td>0.011</td>
<td>0.000</td>
</tr>
<tr>
<td>ppERK2</td>
<td>1.000</td>
<td>1.000</td>
<td>0.989</td>
<td>0.000</td>
<td>0.011</td>
<td>0.000</td>
</tr>
<tr>
<td>ppMEK1</td>
<td>1.000</td>
<td>1.000</td>
<td>0.989</td>
<td>1.000</td>
<td>1.000</td>
<td>0.989</td>
</tr>
<tr>
<td>ppMEK2</td>
<td>1.000</td>
<td>1.000</td>
<td>0.989</td>
<td>1.000</td>
<td>1.000</td>
<td>0.989</td>
</tr>
<tr>
<td>pRSK1</td>
<td>1.000</td>
<td>1.000</td>
<td>0.989</td>
<td>1.000</td>
<td>1.000</td>
<td>0.989</td>
</tr>
<tr>
<td>rpS6-Pi</td>
<td>1.000</td>
<td>1.000</td>
<td>0.989</td>
<td>1.000</td>
<td>1.000</td>
<td>0.989</td>
</tr>
<tr>
<td>pMNK1</td>
<td>1.000</td>
<td>1.000</td>
<td>0.989</td>
<td>0.925</td>
<td>0.948</td>
<td>0.893</td>
</tr>
<tr>
<td>pMNK2</td>
<td>1.000</td>
<td>1.000</td>
<td>0.989</td>
<td>0.000</td>
<td>0.011</td>
<td>0.000</td>
</tr>
<tr>
<td>eIF4E-Pi</td>
<td>0.864</td>
<td>0.896</td>
<td>0.824</td>
<td>0.047</td>
<td>0.074</td>
<td>0.029</td>
</tr>
</tbody>
</table>

Figure 2.2.1a – Example of the LPS Model Comparing the Effects of Inhibitors to a Wild Type Model. Inhibitors are incorporated into the model and the effects are compared to a Wild Type model in which no inhibitors are included. Reacfhq was used to interrogate the QSSPN simulations to determine, for each model effector, the fraction of the 1,000 trajectories sampled in which this effector is activated. 99.99% Binomial Confidence Intervals were used to identify regions of significant difference (shown in red).
These frequency values, and the associated 99.99% Binomial Confidence Intervals, were then used to form Confusion Matrices (as summarised in Figure 2.2.1b). The measures placed in the matrices are determined based upon the identification of perturbations resulting in significant changes in the trajectory frequencies. The frequency value for a particular model effector in an inhibitor model was considered significantly different if the 99.99% Binomial Confidence Interval for that effector if it did not overlap with the corresponding 99.99% Binomial Confidence Interval in the Wild Type model. To form these Confusion Matrices, the behaviour of the model was compared to the LPS Benchmark which contained the expected behaviour of model components in response to the various chemical inhibitors. This benchmark was composed of information taken from 122 literature articles. Supplementary Information 1b contains the bibliographic information of the articles contained with this benchmark, whilst Supplementary Information E5 contains the LPS Benchmark dataset itself.
**Figure 2.2.1b – Overview of the Predictive Power of the Model Incorporating the Effects of Lipopolysaccharide Signalling.** The Petri Net model of translational regulation was interrogated with QSSPN followed by Reachfq. The binomial confidence intervals of model effectors in the model including a chemical inhibitor were compared with those in the Wild Type model and regions of significant difference were identified. By comparing the results of this evaluation to the LPS Benchmark it was possible to produce Confusion Matrices with results being characterised as either True Positives (TP), True Negative (TN), False Positive (FP) and False Negative (FN). These qualitative measures of predictive power were used to produce the Matthews’ Correlation Coefficient (MCC). The overall predictive power was established by pooling the sums Confusion Matrix variables for each inhibitor and calculating an overall MCC value. Due to the absence of TN and FP values, an MCC value for the (5Z)-7-Oxozaenol was not quantifiable.

<table>
<thead>
<tr>
<th>Model</th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
<th>MCC Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGP57380</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>-0.447</td>
</tr>
<tr>
<td>LY294002</td>
<td>8</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>-0.302</td>
</tr>
<tr>
<td>PD184352</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0.356</td>
</tr>
<tr>
<td>PD98059</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>-0.218</td>
</tr>
<tr>
<td>Rapamycin</td>
<td>6</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>-0.289</td>
</tr>
<tr>
<td>SB203580</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>11</td>
<td>0.113</td>
</tr>
<tr>
<td>SP600125</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0.745</td>
</tr>
<tr>
<td>(5Z)-7-Oxozaenol</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>Undefined</td>
</tr>
<tr>
<td><strong>OVERALL</strong></td>
<td>30</td>
<td>9</td>
<td>8</td>
<td>35</td>
<td>-0.007</td>
</tr>
</tbody>
</table>
Whilst this overall MCC value suggests that the binary model leads to random predictions, it is worth noting that there is a smaller number of data points being used. Moreover, the maximum time limit of the simulation was such that the majority of effectors reached frequencies of one. These results prompted a re-evaluation of the simulation protocol. Therefore the binary model was re-tested against the initial, non-LPS dataset and the maximal time-step of the simulations adjusted so that the Wild Type effector frequencies should be approximately 0.50. This approach, which had already been implemented by Fisher and colleagues (2013), allowed for the discovery of increases and decreases in the reachability of the model effectors in response to perturbations. Despite this leading to larger 99.99% Binomial Confidence Intervals for the Wild Type model, it enabled more complex behaviour to be observed. The results of applying this refined protocol are shown below in Section 2.2.2 – Re-Evaluation of the Binary Model.

One of the main outcomes of the preliminary work was the finding that eIF4E phosphorylation was increased in response to mTORC1 inhibition. As such, it is noteworthy that this behaviour was conserved in the LPS model with an increase in the phosphorylation of eIF4E being seen in the Rapamycin inhibitor model.

Section 2.2.1.1 – Updated Benchmark Dataset

Proteomic differences between cell types is well established (Pontén et al., 2009; Paulitschke et al., 2013) and the question was raised as to the validity of using data from a variety of cell lines as to the cellular effects of inhibitors. The dataset constructed with publications, not using LPS stimulation, contains very few studies that use macrophage cell lines, let alone RAW 264.7 cells. It was decided that the dataset containing studies that only use LPS stimulation would be more appropriate to evaluate this effect. As LPS stimulation is
important in the innate immune system, more studies in this dataset use macrophage-derived cell lines. It is therefore possible to test this assumption with the LPS dataset, the initial results of which are given above in Section 2.2.1 – Binary Model of LPS Stimulation. Although there are undoubtedly differences in the functioning of different macrophage-derived cell lines, in order to test this, in the least conservative way, all macrophage-derived cell lines will be included. By eliminating any studies with cell types other than macrophage cell lines, the number of data points used for calculating the overall MCC and AC values was reduced from 82 to 41. By performing a Chi Square Test ($\alpha = 1.55 \times 10^{-4}$) it is clear that this is a significant reduction. The information pertaining to these datasets can be found in Supplementary Information E5.

Despite an overall increase in the global MCC value, by limiting the cell types included in the formation of the benchmark datasets, the number of possible data points was being severely reduced. By doing so, the relevance of even the global MCC values is being reduced. As such, the method of including multiple cell types in the benchmark dataset was the most appropriate. In order to further enhance the value of this dataset, verifying results across multiple studies and cell types was best practise.

**Section 2.2.2 – Re-Evaluation of the Binary Model**

The adjusted maximal time-steps for each effector in the Wild Type was applied to each of the inhibitor models. The frequency of each effector in the Wild Type model was approximately 0.5, as a result of the maximal simulation time being limited. By ensuring the Wild Type frequency took this approximate value, more complex patterns of behaviour were visible. For example, if the simulation time was extended to permit an effector to reach a frequency approaching 1.0, then it would not be possible to observe perturbations
resulting in an increase in the formation of a particular effector. The frequencies, with associated 99.99% Binomial Confidence Intervals, for this model are shown in Supplementary Information E7 – Frequency Data Generated by Reachfq with Associated 99.99% Binomial Confidence Intervals (Sheet entitled First Refinement of the Model).

By changing back to an updated version of the original dataset, more points of analysis were available against which to test the binary model. This updated version of the original dataset, termed the Revised Original Benchmark dataset (shown in Supplementary Information E6) contains information from 130 peer-reviewed articles on the effects of chemical inhibitors on key model effectors. Such a change led to a slight improvement of the predictive power of the model (MCC = -0.007 vs MCC = 0.1393 (shown in Figure 2.2.2a) in the first round of model refinement. However the predictive power of this binary model is lower than the initial model (MCC = 0.2354 vs MCC = 0.1393). This being said, the changes made to the model during the transition to the binary model formalism allowed for improvements to be made.

Further iterations of the model were made to improve the predictive power with the aim of approaching an MCC value of approximately 0.5. The workflow process for these iterations is shown in Figure 2.2.2b and summarises the process used to refine the model through five iterations. The first four iterations of are summarised in Figure 2.2.2a and show the overall MCC values increase as changes are made. Whilst global improvements were made to the MCC values, these were subject to fluctuations when looking at individual Inhibitor Models. This highlights the need to view the MCC values of the individual models with caution. Furthermore, the changes to the model have made plain the importance of identifying and
increasing cross-talk between pathways. The Revised Original Benchmark dataset was used in each comparison of the model in order to calculate the measure of predictive power.

Of interest to the experimental work of this project, the earlier prediction concerning the effect of mTORC1 inhibition on eIF4E phosphorylation is still observed. The frequency of trajectories producing eIF4E phosphorylation fell from 0.520 in the Wild Type model to 0.110 when Rapamycin was included.
<table>
<thead>
<tr>
<th></th>
<th>First Round of Model Refinement</th>
<th>Second Round of Model Refinement</th>
<th>Third Round of Model Refinement</th>
<th>Fourth Round of Model Refinement</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP</td>
<td>31</td>
<td>34</td>
<td>39</td>
<td>43</td>
</tr>
<tr>
<td>TN</td>
<td>31</td>
<td>35</td>
<td>38</td>
<td>42</td>
</tr>
<tr>
<td>FP</td>
<td>8</td>
<td>6</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>FN</td>
<td>59</td>
<td>51</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>MCC</td>
<td>0.1393</td>
<td>0.2553</td>
<td>0.3185</td>
<td>0.3918</td>
</tr>
<tr>
<td>AC</td>
<td>0.1393</td>
<td>0.2553</td>
<td>0.3185</td>
<td>0.3918</td>
</tr>
<tr>
<td>Chi Square Test α-Value</td>
<td>8.27x10^-9</td>
<td>2.78x10^-7</td>
<td>8.81x10^-6</td>
<td>5.18x10^-7</td>
</tr>
</tbody>
</table>

Figure 2.2.2a – Overall Predictive Power of the Model Through the Four Stages of Refinement. The binary model of translational regulation was compared to the Revised Original Benchmark Dataset. The comparison of the dataset to the behaviour of the model led to the construction of Confusion Matrices with behaviours being classified as being True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN). These qualitative measures were converted to quantitative measures through the calculation of both the Matthews’ Correlation Coefficient (MCC) and Approximate Coefficient (AC) values. Through four rounds of model refinement, it was possible to raise the level of predictive power from an MCC = 0.1393 to MCC = 0.3918. Each refinement stage concerned the increase in the level of cross-talk between pathways and the elimination of FP values. To accomplish these, it was necessary to return to existing literature.
Figure 2.2.2b – The Model Refinement Workflow. Over four iterations of the model, it was refined to improve the predictive power. Simulations were run using QSSPN and Reachfq once the optimal simulation times were established for each model effector. This was necessary to ensure that the Wild Type frequency of activation was approximately 0.5. The 99.99% Binomial Confidence Intervals were calculated and the results were compared to the relevant Benchmark dataset. Using these comparisons, the Matthews’ Correlation Coefficient (MCC) values were calculated to quantitatively measure the predictive power of the model. Based upon the MCC values for the individual inhibitor models, the connectivity of the model was refined using additional literature. The cross-talk of the signalling pathways was the main target for refinement so as to allow more complex patterns of behaviour.
The fifth round of model refinement was responsible for raising the predictive power of the model to an MCC value of approximately 0.5 (as shown in Figure 2.2.2c). The frequency values, along with the 99.99% Binomial Confidence Intervals, are shown in Supplementary Information E7 – Frequency Data Generated by Reachfq with Associated 99.99% Binomial Confidence Intervals (sheet entitled Fifth Refinement of the Model). Of particular interest is the finding that eIF4E phosphorylation was significantly altered by Rapamycin, as noted by a deviation in the frequency of trajectories reaching one discrete activity level (0.522 in the Wild Type model versus 0.005 in the Rapamycin model). Such an affect was also noted in the preliminary work and the LPS containing model (outlined in Section 2.2.1 – Binary Model of LPS Stimulation). Moreover, this alteration is associated with a perturbation in MNK2, as noted by Stead and Proud (2013). However, in this model eIF4E phosphorylation appears reduced despite the increase in the frequency of trajectories producing phosphorylated MNK2. Work in subsequent sections of this chapter were aimed at eliminating this discrepancy and defining the mechanism by which Rapamycin can increase the phosphorylation of this initiation factor.
### i.

<table>
<thead>
<tr>
<th>Model</th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
<th>MCC Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIX02189</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>4</td>
<td>0.3568</td>
</tr>
<tr>
<td>CGP57380</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0.8000</td>
</tr>
<tr>
<td>GSK2334470</td>
<td>2</td>
<td>12</td>
<td>0</td>
<td>2</td>
<td>0.6547</td>
</tr>
<tr>
<td>LY294002</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>0.3651</td>
</tr>
<tr>
<td>PD184352</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>0.2282</td>
</tr>
<tr>
<td>PD98059</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>0.4472</td>
</tr>
<tr>
<td>Rapamycin</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>0.4330</td>
</tr>
<tr>
<td>SB203580</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>0.1177</td>
</tr>
<tr>
<td>SP600125</td>
<td>10</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>0.4364</td>
</tr>
<tr>
<td>XMD8-92</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

### ii.

#### Confusion Matrix Values

<table>
<thead>
<tr>
<th>Model</th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
<th>MCC Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>47</td>
<td>47</td>
<td>3</td>
<td>39</td>
<td>0.4652</td>
</tr>
</tbody>
</table>

#### Measures of Predictive Power

<table>
<thead>
<tr>
<th>Model</th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
<th>MCC Value</th>
<th>AC Value</th>
<th>α-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>47</td>
<td>47</td>
<td>3</td>
<td>39</td>
<td>0.4652</td>
<td>0.4652</td>
<td>5.15x10^-8</td>
</tr>
</tbody>
</table>

**Figure 2.2.2c – Evaluation of the Predictive Power of the Fifth Iteration of the Model.**

(i) Once QSSPN and Reachfq had evaluated the model and regions of significant difference had been identified, it was possible to compare the effects of chemical inhibitors in the model to published effects of the inhibitor on particular model effectors (within the Revised Original Benchmark Dataset). The result of this comparison was the formation of Confusion Matrices. Consequently, the behaviour of the model was characterised as being either a True Positive (TP), True Negative (TN), False Positive (FP) or False Negative (FN). In order to produce a quantitative measure of predictive power, these values were used to calculate the Matthews’ Correlation Coefficient (MCC). (ii) To produce an overall measure of predictive power, the Confusion Matrices variables in (i) were pooled. The MCC and Approximate Coefficient (AC) values were calculated. In order to show the predictive power was significantly different from random, the Chi Square Test α-Value was calculated.
Section 2.2.3 – Mechanism by Which mTORC1 Inhibition Affects eIF4E Phosphorylation

Based upon the connectivity of the model, only four factors are directly linked to the phosphorylation of eIF4E: MNK1, MNK2, eIF4G phosphorylation and 4E-BP1 phosphorylation. Through the construction of a smaller model, again within Snoopy, a more detailed investigation of eIF4E phosphorylation was possible. This smaller model contained only molecular interactions included in the original model and only molecular interactions that are proximal to eIF4E phosphorylation. Given the reduced scale of the model, it was not possible to include the complex cross-talk between the signalling pathways present in the larger model. However, as the aim of this smaller model was to explain the behaviour of the eIF4E phosphorylation seen when Rapamycin is present, much of this cross-talk was not deemed necessary in the first instance.

Simulations with this model revealed that eIF4G phosphorylation and 4E-BP1 phosphorylation appeared to act to reduce eIF4E phosphorylation in a model incorporating Rapamycin. This raised three possible, but not necessarily mutually exclusive, mechanisms to explain how an increase in eIF4E phosphorylation was possible:

i. The level of 4E-BP1 in the cell must lower than that of eIF4E so as to allow sequestration of eIF4E by hypophosphorylated 4E-BP1 and at the same time permit an increase in eIF4E phosphorylation.

ii. mTORC2 cannot be the only kinase responsible for the activation of the eIF4G kinase, PKCα.
iii. The activity of a protein phosphatase may be suppressed by mTORC1 activity. Such a phosphatase may act on either MNK1 or MNK2, eIF4G, PKCα or even eIF4E directly.

The first hypothesis is in agreement with the work of Eckerdt and co-workers (2014) who established that the concentration of 4E-BP1 had little influence on the increased phosphorylation of eIF4E. Although difficult and costly, proteomic evaluation of RAW 264.7 cells could reveal if this is the case and show if and to what extent the levels of eIF4E and 4E-BP1 differ. Several studies have employed proteomic approaches with RAW 264.7 cells (Martínez-Solano et al., 2009; Shi et al., 2009; Eichelbaum & Krijgsveld, 2014). However to date no information regarding the relative levels of either protein have been reported in these studies. Cai et al. (2014) and Decarlo et al. (2015) both remark that the ratio of eIF4E to 4E-BP1 is critical in the early transformation events of cells. Quantitative immunoblotting was not considered here as the method is often regarded as pseudo-quantitative or, in the case of advanced quantitative immunoblotting, requires the expression of recombinant proteins in control samples to be used to aid the quantification (Janes, 2015). Moreover, the quantification of Western Blots with densitometry is affected by a great number of factors that may arise during the immunoblotting procedure, the scanning of blots or even during the quantification itself and consequently has little in the way of scientific foundation (Gassmann et al., 2009). Taylor and co-workers (2013) even go so far as to state that the use of radiographic film is inferior to camera-based development systems when it comes to quantitative immunoblotting. For these reasons, a proteomics based method would be preferred for the quantification of eIF4E and 4E-BP1 levels.
Whilst there is no direct evidence to support the hypothesis that 4E-BP1 is expressed at a lower level than eIF4E, there is reason to believe this is the case. Data taken from The Cell Line Atlas database, within The Human Protein Atlas (Uhlén et al., 2005; Uhlén et al., 2015), has revealed that in the majority of cases, across multiple human cell cancer lines, that the level of eIF4E (Analysis of Expression of EIF4E Website, taken from http://www.proteinatlas.org/ENSG00000151247-EIF4E/cell/CAB004077 [last accessed 22/05/2015]) is higher than that of 4E-BP1 (Analysis of Expression of EIF4EBP1 Website, taken from http://www.proteinatlas.org/ENSG00000187840-EIF4EBP1/cell/CAB005032 [last accessed 22/05/2015]). Using Immunohistochemistry (IHC) data, of the 46 cell lines tested, only four (8.7%) displayed evidence of 4E-BP1 being present at higher levels than eIF4E. The data shown in The Human Protein Atlas, whilst supportive of the view that eIF4E is expressed to a higher degree than 4E-BP1, is not quantitative enough to provide conclusive support to this hypothesis. This issue of IHC being quantifiable is summarized by Taylor and Levenson (2006). This work notes that there is some contention over whether a correlation exists between the amount of antigen staining and the actual amount of the antibody-target. Samples generated by IHC are also subject to a variable interpretation by observers and, thus, the resulting scoring of sample staining is subjective and open to vast interpretation. As a result of this, and a lack of suitable experimental controls, reproducibility is often lacking even when analytical software is used (Ibid).

Although mTORC2 has previously been found to be a kinase for PKCα (Guertin et al., 2006; Facchinetti et al., 2008; Ikenoue et al., 2008), experimental work presented in Chapter 3 – Experimental Validation of Model Predictions has demonstrated that the mTORC2 phosphorylation site of AKT1, Ser-473 (Frias et al., 2006; Jacinto et al., 2006; Breuleux et al.,
2009), displays a significant reduction in AKT1 phosphorylation when using the Human Phospho-MAPK Arrays following a one hour Rapamycin treatment. It is worth re-iterating here that whilst AKT1 phosphorylation was reduced, the $p$-value calculated was only significant prior to a Holm-Sidak Correction. However, the AKT Pan level remained significantly reduced after this correction. Furthermore, immunoblotting of lysates treated for three hours displayed no significant difference in the level of Ser-473 phosphorylated AKT1. It is possible that this is as a result of reduced antibody sensitivity with this method as the signal quantified was very low. Given there is still reason to argue that AKT1 phosphorylation is reduced, along with the simulation results and the Rapamycin-mediated increase in eIF4E phosphorylation, it is reasonable to argue that mTORC2 cannot be the only kinase responsible for phosphorylating PKCα, a kinase known to be responsible for eIF4G phosphorylation (Dobrikov et al., 2011).

Although speculative, PDK1 has been shown to be capable of phosphorylating PKCα, and other PKC enzymes, in such a way as to be required for the activation of PKC (Dutil et al., 1998; Le Good et al., 1998; Sonnenburg et al., 2001). Gao and co-workers (2001) noted that, in PKCβII, PDK1 mediates phosphorylation of Thr-500 residue within the PKC activation loop. Autophosphorylation of PKCβII can then occur when PDK1 has dissociated (Ibid). Several studies have noted that PDK1 is required to maintain the level of PKCα and other PKC isoforms expressed in cells (Collins et al., 2005; Wood et al., 2007). When taken together with the work of Dutil et al. (1998) and Sonnenberg et al. (2001), which noted that PDK1 was capable of phosphorylating PKCα on the activation loop, these works suggest that PDK1 not only activates PKCα but also positively regulates the expression of this protein kinase. Although PDK1 has been shown to be the kinase responsible for activating multiple
PKC isoforms (including PKCβ (Helliwell et al., 2003), PKCδ (Deb et al., 2003), PKCζ (Chou et al., 1998) and PKCθ (Bauer et al., 2001; Villalba et al., 2002) this occurs in a manner dependent upon the PI3K pathway. It would not be unreasonable, based upon this evidence, to hypothesise that it is probable that PKCα activation would occur in an analogous fashion.

Testing this hypothesis could be achieved with either chemical inhibition of PDK1, CRISPR or with transgenic mice (these are explained in Section 3.4.4 – Potential Future Work). However, for reasons stated previously, the relevance of both to this project is somewhat unclear. At present CRISPR has not been used in RAW 264.7 cells and so it is not clear if this would be appropriate at this time. Transgenic mice have been successfully used to create PDK1 knockout mice (Tawaramoto et al., 2012), however, the cells derived from these would be likely to display behaviour different to that of RAW 264.7 cells. Chemical inhibition of PDK1 would at this time be the best option. This being said, it would require treating the cells with multiple inhibitors and the issue of inhibitor-inhibitor interactions would then be raised.

With regard to the third proposed mechanism, mTORC1 is known to regulate the functioning of several protein phosphatases, PP5 (Huang et al., 2004) and PP2A (Hartley & Cooper, 2002). As Rapamycin has been shown to lead to an increase in the Ser-437 phosphorylation of MNK2 (Stead & Proud, 2013), it follows that if a phosphatase were involved, that it must be suppressed during mTORC1 inhibition. This suggested that the increase in eIF4E phosphorylation seen during Rapamycin treatment maybe a result of alterations to protein phosphatase activity. Given that Hartley and Cooper (2002) found that Rapamycin enhanced PP2A activity, it is unlikely this phosphatase is a viable candidate.
Bishop and co-workers (2006) similarly noted that mTORC1 acts to limit PP2A activity. However, Huang and co-workers (2004) established that during Rapamycin treatment, the regulatory subunit of PP5, PR72, interacts with PP5 with reduced affinity and thus moderates PP5 activity. This phenomenon appears to be dependent upon the expression of 4E-BP1 but no mechanism has been proposed (Ibid).

Whilst this would seem to indicate that that PP5 is possible candidate here, to date, no work has linked PP5 to the regulation of eIF4E phosphorylation or to either MNK1 or MNK2. This being said, it has been reported that PP5 is capable of dephosphorylating sites targeted by PKC (Katayama et al., 2014). Whilst not conclusive, this does intimate that eIF4G may be a substrate for PP5 in this system. Furthermore, von Kriegsheim and co-workers (2006) noted that PP5 may serve to negatively regulate ERK1/2 signalling by dephosphorylating Raf-1 which suggests that downstream effectors, such as eIF4E, may also be affected.

Any potential work to elucidate a role for PP5 would be severely hampered by a lack of a suitable or specific inhibitor of PP5 (Swingle et al., 2007). Studies of the function of PP5 have been aided by RNA interference against it (Messner et al., 2006) as well as the generation of mice deficient in PP5 (Amable et al., 2011). Nevertheless, given the problems detailed in Section 3.4.3 – Problems Encountered & Troubleshooting, these do not appear to be promising avenues. However, both Gentile et al. (2006) and Messner et al. (2006) have used PP5 in which the Tyr-451 is mutated to an Alanine residue. The result of this is a considerable reduction in sensitivity to the inhibitor, Okadaic Acid. A comparison of the effects Okadaic Acid has upon wild type PP5 and the resistant PP5 may yield information on whether this phosphatase affects eIF4E phosphorylation or the activity of upstream signalling components, such as MNK1, MNK2 or eIF4G.
Whilst the only direct link between mTORC1 inhibition and PP2A appears to be one of activation, there is some evidence that PP2A may be positively regulated by negative feedback loops that exist between AKT1 and mTORC1 (as described in Section 3.1.3 – Interplay Between mTORC1 and AKT1). Although AKT1 is negatively regulated by PP2A (Hong et al., 2008; Aceto et al., 2009; Kim et al., 2009b), there is some evidence to suggest that there is some dual-regulation of PP2A and AKT1. In response to Insulin stimulation, the increase in AKT1 phosphorylation occurs concomitantly with an increase in PP2A phosphorylation (Kumar & Tikoo, 2015). This phosphorylation occurs on the Tyr-307 residue (Ibid) which is noted as suppressing the phosphatase activity of PP2A (Chen et al., 1992). This suggests that enhancing the activity of AKT1 would increase the phosphorylation of the inhibitory Tyrosine residue of PP2A. This could be achieved with, for example, low concentrations of Rapamycin. Although not demonstrated, this could act to counter the increase in PP2A activity noted by Hartley and Cooper (2002) and Bishop et al. (2006).

While both Src (Chen et al., 1992; Arif et al., 2014) and Glycogen Synthase Kinase 3β (Yao et al., 2011) are known to phosphorylate PP2A, it is not clear whether AKT1 is involved in either mechanism.

Whilst somewhat subjective, the first and second explanations given towards the start of this section have been validated by keeping eIF4E at a higher level than 4E-BP1 and the removal of a dependence on MNK1 and phosphorylated eIF4G from the model. Indeed, analogous changes in the small model had already been shown to result in the level of eIF4E phosphorylation increasing in response to Rapamycin. The removal of these was shown to bring about an increase in eIF4E phosphorylation, as shown in Section 2.2.4 – Evaluation of the Predictive Power of the Final Model. The removal of these effectors demonstrates that
MNK2 is the kinase responsible for the increase in eIF4E phosphorylation, even in the absence of any increase in ERK1 or ERK2 activity. Whilst in this case it is harder to ascribe a role for the third hypothesis mentioned above, the inhibition of a phosphatase cannot be ruled out as it is a possibility that the increase in MNK2 phosphorylation accompanying Rapamycin treatment is associated with a reduction in phosphatase activity.

**Section 2.2.4 – Evaluation of the Predictive Power of the Final Model**

The final version of the Petri Net model contained 511 molecular species and 584 molecular interactions. These species and interactions were derived from existing literature, detailing the interactions between mammalian cellular signalling components regulating a variety of eIFs and between the translation initiation machinery itself.

Before starting the next phase of the project, implementing the effects viral infection has upon the signalling network and the cellular translational machinery (as detailed in Chapter 4 – Modelling the Effects of Viral Infection), it was necessary to re-evaluate the predictive power of the model. This was of particular importance once eIF4E phosphorylation was increased in response to Rapamycin. The dataset against which this model is tested is shown in Supplementary Information 2. Furthermore, work in this section will also address potential concerns that the model was over-trained to the Benchmark Dataset against which the behaviour of the model was compared.

Prior to detailing the results of this final round of predictive power analysis, it is necessary to address a point mentioned in **Section 2.1.2 – The Binary Model Formalism**. It must be demonstrated that the Binary Model Formalism is rigorous and does not lead to a model
which permits a molecular species to be loaded with more than one discrete activity level (Figure 2.2.4a).
<table>
<thead>
<tr>
<th>Model Component</th>
<th>Frequency Reaching 2 Tokens or More</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phosphorylated MEK1</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated MEK2</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated ERK1</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated ERK2</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated RSK1</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated rpS6</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated MNK1</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated MNK2</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated eIF4E</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated JNK1</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated c-Jun</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated p38α</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated ERK5</td>
<td>0.000</td>
</tr>
<tr>
<td>Active PI3K</td>
<td>0.000</td>
</tr>
<tr>
<td>PIP₃</td>
<td>0.000</td>
</tr>
<tr>
<td>Active PDK1</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated PKCζ</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated AKT1</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated S6K1</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated 4E-BP1</td>
<td>0.000</td>
</tr>
<tr>
<td>Active mTORC1</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated TSC1/2</td>
<td>0.000</td>
</tr>
<tr>
<td>Phosphorylated eIF4B</td>
<td>0.000</td>
</tr>
<tr>
<td>Protein Synthesis</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Figure 2.2.4a – Confirmation of the Binary Nature of the Model Formalism. A QSSPN simulation was run sampling 100 trajectories. Rather than interrogating the model with Reachfq to see how many trajectories reach one discrete activity level of each molecular species, the interrogation determined how many trajectories resulted in each model effector reaching two discrete activity levels. All model effectors in this list were included in the model refinement process. Protein Synthesis was included here as the overarching aim of the project was to model the regulation of nascent protein synthesis.
Supplementary Information E7 – Frequency Data Generated by Reachfq with Associated 99.99% Binomial Confidence Intervals (sheet entitled Final Binary Model) details the frequency values, and associated 99.99% Binomial Confidence Intervals, calculated from the results of the QSSPN simulations for each of the inhibitor models. The frequency values and associated 99.99% Binomial Confidence Intervals of the common components, found in both the model and Human Phospho-MAPK Arrays (as discussed in Section 3.3.5 – Comparison of Model Results to Array Data) are shown in Figure 2.2.4b. The final model file is provided in Supplementary Information E3 (as both a Snoopy model file (E3a) and an SBML file (.xml)) (E3b).

Importantly, eIF4E phosphorylation was shown to be increased in response to Rapamycin. This was determined through an increase in the frequency of trajectories reaching one discrete activity level of phosphorylated eIF4E (0.492 in the Wild Type model versus 0.805 in the Rapamycin model). Such an increase was expected based upon the findings of a number of studies (Sun et al., 2005; Wang et al., 2007c; Jin et al., 2008a; Grosso et al., 2011; Marzec et al., 2011; Stead & Proud, 2013; Eckerdt et al., 2014). Moreover, the continued presence of an increase in trajectories reaching phosphorylated MNK2 in response to Rapamycin (0.463 in the Wild Type model versus 0.761 in the Rapamycin model), is in keeping with earlier work (Stead & Proud, 2013).
<table>
<thead>
<tr>
<th></th>
<th>Wild Type</th>
<th>Upper</th>
<th>Lower</th>
<th>Rapamycin</th>
<th>Upper</th>
<th>Lower</th>
</tr>
</thead>
<tbody>
<tr>
<td>ppERK1</td>
<td>0.520</td>
<td>0.570</td>
<td>0.468</td>
<td>0.525</td>
<td>0.576</td>
<td>0.473</td>
</tr>
<tr>
<td>ppERK2</td>
<td>0.500</td>
<td>0.552</td>
<td>0.448</td>
<td>0.512</td>
<td>0.564</td>
<td>0.460</td>
</tr>
<tr>
<td>pRSK1</td>
<td>0.577</td>
<td>0.627</td>
<td>0.525</td>
<td>0.560</td>
<td>0.611</td>
<td>0.508</td>
</tr>
<tr>
<td>ppp38a</td>
<td>0.537</td>
<td>0.588</td>
<td>0.485</td>
<td>0.544</td>
<td>0.595</td>
<td>0.492</td>
</tr>
<tr>
<td>pJNK1</td>
<td>0.594</td>
<td>0.644</td>
<td>0.542</td>
<td>0.625</td>
<td>0.674</td>
<td>0.574</td>
</tr>
<tr>
<td>mTORC1-Act</td>
<td>0.555</td>
<td>0.606</td>
<td>0.503</td>
<td>0.000</td>
<td>0.011</td>
<td>0.000</td>
</tr>
<tr>
<td>ppS6K1</td>
<td>0.153</td>
<td>0.194</td>
<td>0.119</td>
<td>0.000</td>
<td>0.011</td>
<td>0.000</td>
</tr>
<tr>
<td>ppAKT1</td>
<td>0.569</td>
<td>0.620</td>
<td>0.517</td>
<td>0.006</td>
<td>0.021</td>
<td>0.002</td>
</tr>
<tr>
<td>pMKK3</td>
<td>0.542</td>
<td>0.593</td>
<td>0.490</td>
<td>0.521</td>
<td>0.572</td>
<td>0.469</td>
</tr>
<tr>
<td>pMKK6</td>
<td>0.534</td>
<td>0.585</td>
<td>0.482</td>
<td>0.529</td>
<td>0.580</td>
<td>0.477</td>
</tr>
<tr>
<td>pMSK2</td>
<td>0.607</td>
<td>0.656</td>
<td>0.555</td>
<td>0.598</td>
<td>0.648</td>
<td>0.546</td>
</tr>
<tr>
<td>pCREB</td>
<td>0.576</td>
<td>0.626</td>
<td>0.524</td>
<td>0.573</td>
<td>0.623</td>
<td>0.521</td>
</tr>
<tr>
<td>pGSK3</td>
<td>0.546</td>
<td>0.597</td>
<td>0.494</td>
<td>0.256</td>
<td>0.304</td>
<td>0.213</td>
</tr>
</tbody>
</table>

Figure 2.2.4b – Components of the Model Common to Human Phospho-MAPK Array. Model components for which the effects of Rapamycin were determined using the Human Phospho-MAPK Arrays (the overall results of which are detailed in Section 3.3.2 – Evaluation of the Human Phospho-MAPK Arrays) were evaluated using QSSPN. Binomial Confidence Intervals, calculated in R, were used to find regions of significant difference, shown in red, by comparing the frequency of activation for each effector in the Wild Type and Rapamycin model.
Figure 2.2.4c summarises the comparison between the behaviour of the model and either the Final Benchmark dataset of 134 literature articles concerning the effects chemical inhibitors have upon model effectors, or the results of the Human Phospho-MAPK Arrays. The Final Benchmark dataset is given in Supplementary Information 2, whilst the bibliographic information contained within it is given in Supplementary Information 1a.
i.

<table>
<thead>
<tr>
<th>Model</th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
<th>MCC Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIX02189</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>4</td>
<td>0.3568</td>
</tr>
<tr>
<td>CGP57380</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0.8000</td>
</tr>
<tr>
<td>GSK2334470</td>
<td>2</td>
<td>12</td>
<td>0</td>
<td>2</td>
<td>0.6547</td>
</tr>
<tr>
<td>LY294002</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>0.3651</td>
</tr>
<tr>
<td>PD184352</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>0.2282</td>
</tr>
<tr>
<td>PD98059</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>0.4472</td>
</tr>
<tr>
<td>Rapamycin</td>
<td>6</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0.5222</td>
</tr>
<tr>
<td>SB203580</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>0.0255</td>
</tr>
<tr>
<td>SP600125</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>0.5000</td>
</tr>
<tr>
<td>XMD8-92</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1.0000</td>
</tr>
<tr>
<td>Array Data</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>0.3858</td>
</tr>
</tbody>
</table>

ii.

<table>
<thead>
<tr>
<th>Confusion Matrix Values</th>
<th>Measures of Predictive Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP</td>
<td>TN</td>
</tr>
<tr>
<td>Overall</td>
<td>51</td>
</tr>
</tbody>
</table>

**Figure 2.2.4c – Evaluation of the Predictive Power of the Final Version of the Model.**  (i) By comparing the Final Benchmark dataset with the behaviour of the model, it was possible to produce Confusion Matrices detailing a qualitative measure of predictive power. Such matrices characterise the behaviour of the model as being either True Positive (TP), True Negative (TN), False Positive (FP) or False Negative (FN). Additionally, under ‘Array Data’ it was possible to evaluate the behaviour of model effectors which are also found on Human Phospho-MAPK Arrays (as detailed in Section 3.3.2 – Evaluation of the Human Phospho-MAPK Arrays). For each inhibitor and the experimental data, a quantitative measure of predictive power is given in the form of the Matthews’ Correlation Coefficient (MCC). (ii) The overall predictive power is found by pooling the Confusion Matrix variables and calculating a final MCC value. Additionally, the Approximate Coefficient (AC) was also calculated. The Chi Square Test α-Value was calculated to ensure the level of predictive power was significantly different from random.
Overall the final version of the binary model has considerable predictive power which is conserved following the alterations to the model that corrected the discrepancy in eIF4E phosphorylation following Rapamycin inclusion. Whilst the predictive power does suffer a slight reduction (from MCC = 0.4652 (as calculated in Section 2.2.2 – Re-Evaluation of the Binary Model) to 0.4558) this is not significant (as determined by a Chi Square Test ($\alpha = 0.2747$)). A further Chi Square Test revealed that the Matthews’ Correlation Coefficient for this model is significantly different from random ($\alpha = 1.94 \times 10^{-8}$).

Whilst it is not feasible to reconstruct the model multiple times, nor is it possible to perturb the parameters of a qualitative non-parametric model, it was possible to ensure that this MCC value is representative of a model which has not been over-trained to a particular dataset. To this end, parts of the Final Benchmark Dataset were, randomly removed and the MCC value recalculated, as shown in Figure 2.2.4d. Whilst this is not a cross-validation procedure in a machine learning sense, it is an analysis of the stability of the results obtained during this work and demonstrates that the model provides stable predictions of the perturbations induced by inhibitors.
In order to demonstrate that the model can provide stable predictions of the behaviour of inhibitors, subsets of the Final Benchmark Dataset were randomly chosen. These subsets consisted of half of the data points within the Final Benchmark Dataset. Using the recorded Confusion Matrix outcomes (True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN)) of these data points, it was possible to recalculate the predictive power of the model using the Matthews’ Correlation Coefficient (MCC). Whilst the five subsets used each have different levels of predictive power, the MCC values are not dissimilar to the Overall value calculated earlier in the section (shown in Figure 2.2.4b). To demonstrate statistical significance a Chi Square Test was performed on all data. The resulting α-value provides evidence of a model having a level of predictive power which is statistically significant from random.

<table>
<thead>
<tr>
<th>Confusion Matrix Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP</td>
</tr>
<tr>
<td>OVERALL</td>
</tr>
<tr>
<td>SUBSET</td>
</tr>
<tr>
<td>First</td>
</tr>
<tr>
<td>Second</td>
</tr>
<tr>
<td>Third</td>
</tr>
<tr>
<td>Fourth</td>
</tr>
<tr>
<td>Fifth</td>
</tr>
</tbody>
</table>

**Figure 2.2.4d – Stability Analysis of the Final Model.** In order to demonstrate that the model can provide stable predictions of the behaviour of inhibitors, subsets of the Final Benchmark Dataset were randomly chosen. These subsets consisted of half of the data points within the Final Benchmark Dataset. Using the recorded Confusion Matrix outcomes (True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN)) of these data points, it was possible to recalculate the predictive power of the model using the Matthews’ Correlation Coefficient (MCC). Whilst the five subsets used each have different levels of predictive power, the MCC values are not dissimilar to the Overall value calculated earlier in the section (shown in Figure 2.2.4b). To demonstrate statistical significance a Chi Square Test was performed on all data. The resulting α-value provides evidence of a model having a level of predictive power which is statistically significant from random.
Section – 2.3 - Discussion

Section 2.3.1 – Mechanism of Increasing eIF4E Phosphorylation in Response to mTORC1 Inhibition

With regards to the proposed mechanism by which Rapamycin leads to an increase in eIF4E phosphorylation, three theoretical mechanisms have been proposed to explain this. However, it is likely that a combination of the three would be required. It seems possible in order for any phosphorylation of eIF4E to occur during Rapamycin treatment that the cell must contain a larger pool of this initiation factor than that of 4E-BP1. Following on from this, experimentally a decrease in AKT1 phosphorylation on the mTORC2-dependent site is seen, it is logical to assume that mTORC2 inhibition occurs, therefore, for eIF4E phosphorylation to increase in response to Rapamycin, mTORC2 cannot be the only kinase responsible for activating the eIF4G kinase, PKCα. Inhibition of an mTORC1-stimulated protein phosphatase, for example PP5, would then occur. The result would be a system in which eIF4E is phosphorylated by both an increase in the positive regulatory signals and a decrease in the signals that negatively regulate eIF4E phosphorylation.

While the work presented in this chapter, offered the circumstances required for the mTORC1-mediated increase in eIF4E phosphorylation, the finding that the phosphorylation of this initiation factor is increased is not novel. Several studies have similarly noted this effect (Sun et al., 2005; Wang et al., 2007c; Jin et al., 2008a; Grosso et al., 2011; Marzec et al., 2011; Stead & Proud, 2013; Eckerdt et al., 2014). A more detailed summary of the findings of these papers is given in Section 3.1.2 – eIF4E Phosphorylation During mTORC1 Inhibition.
The 99.99% Binomial Confidence Intervals infer there is no effect of Rapamycin treatment on the activation of ERK1 or ERK2 and this is in agreement with Marzec et al. (2011) and Stead and Proud (2013). The notion of Rapamycin treatment having no effect on ERK1/2 activity (Oh et al., 2001b; Omura et al., 2005; Soares et al., 2013) is somewhat controversial, with other studies noting an increase (Carracedo et al., 2008; Chaturvedi et al., 2009). However, this can be explained by differences in the concentration of Rapamycin used (Chen et al., 2010). This aside, and in agreement with the work of Stead and Proud (2013), the model does predict that MNK2 mediates this alteration to eIF4E phosphorylation in an ERK1/2-independent manner.

Section 2.3.2 – Predictive Power of the Final Version of the Model

The re-evaluation of the predictive power of the final version of the binary model displayed an increase in eIF4E phosphorylation. This increase was achieved through the elimination of molecular interactions linking eIF4G phosphorylation and MNK1 to the phosphorylation of eIF4E. The elimination of these molecular interactions goes some way towards proving that PKCα phosphorylation cannot rely solely upon mTORC2. The second change to the model was to maintain a pool of eIF4E so that it was in excess over the level of 4E-BP1 and does prove that there must be some disparity in the cellular pools of these factors. Only linking the decrease in activity of a phosphatase was not attempted here. This was not investigated as the nature of the relationship between mTORC1 and cellular phosphatases is not clear and mechanistically difficult to implement.

As the final version of this model, it was important to ensure the predictive power was maintained at a level at which the model is an accurate representation of the regulatory network that controls translation initiation. The final version of the binary model was raised
to a values of MCC = 0.4558. Given the probability of seeing this level of predictive power by chance was very low ($\alpha = 1.94 \times 10^{-8}$), the model has achieved a significant predictive power.

In addition to measuring the predictive power of the model, it was also necessary to ensure that the model had not been over-trained to the benchmark dataset against which it was tested. To this end, a form of cross-validation was used. Cross-validation is a term from the machine learning field (Bousquet & Elisseeff, 2002; Shipp et al., 2002; Arlot & Celisse, 2010).

All available data is divided into ‘training’ and ‘validation’ sets. Subsequently, the model is automatically trained on the ‘training’ set and evaluated against the ‘validation’ set. This division is then randomly changed and the process repeated. This methodology is not applicable in the field of GSMN model reconstruction, as such models are constructed manually rather than automatically. Manual reconstruction of large-scale models cannot be repeated multiple times as it takes many months to complete by expert bio-curators (Terzer et al., 2009; Henry et al., 2010 (as cited by McCloskey et al., 2013)). Consequently, none of the models in the Constraint Based Modelling field is formally cross-validated in the same way as in the machine learning field. Despite this controversy, the Constraint Based Modelling field has achieved spectacular successes in biotechnology (Scheibe et al., 2009; Fang et al., 2011) and recently has been gaining applications in the pharmaceutical industry (Trawick & Schilling, 2006). This work further extends the manual, expert reconstruction approach to the qualitative modelling of large-scale signalling networks.

With this in mind, work was conducted to demonstrate the predictive power of the model. Firstly, the Benchmark Datasets of responses to external perturbations was constructed. Such datasets were not used in the initial round of model construction.
Consequently, the MCC of 0.2354 can be used to demonstrate the predictive power of the model, prior to any refinement. This MCC value was found to be statistically significant. Furthermore, the model has been used to make predictions regarding biological behaviour which dedicated experimental work has been conducted to validate (detailed in Chapter 3 – Experimental Validation of Model Predictions). This established that the model can be used to make predictions which are useful in guiding future experimental work. This is in fact the most important demonstration of the predictive power of the model.

A further indication that the model is not over-trained in a machine-learning sense is that it is not perfect. It would be relatively easy to construct rules, such that all responses in the Benchmark Dataset would be reproduced. This has not been done. Instead, the model has been modified exclusively when there is peer-reviewed literature supporting the relevant modification to a molecular interaction. This approach did not lead to a model with perfect predictive power, but rather a model with statistically significant MCC value. Such a value indicates that many discrepancies between the current state of knowledge on molecular interactions and the behaviour of the model still exist. Such incongruities were left in the model rather than being removed by arbitrary over-training.

Whilst the points listed above make a compelling argument against the model being over-trained, the work detailed in Figure 2.2.4c provide an argument for a reader originating in a machine-learning field. Each of the five subsets of the Benchmark Dataset were randomly selected and yield a level of predictive power that was comparable to the overall level of predictive power of MCC = 0.4558. Moreover, these subsets were capable of producing levels of predictive power that were significantly better than random.
Section 2.3.3 – Problems Encountered & Troubleshooting

The main problem encountered during the completion of work that contributed towards this chapter occurred when trying to correct the decrease in eIF4E phosphorylation in the Rapamycin model. Two issues became evident here. Firstly, dissecting useful data from the large model proved difficult given that there is a high degree of cross-talk between the signalling pathways. Secondly, the interrogation of the model by either Reachfq or Eventfq did not readily lead to a potential mechanism.

These problems were overcome through the generation of a small scale model. This model allowed for the easy identification of effectors critical for eIF4E phosphorylation. From here it was possible to use the knowledge of these effectors to construct a literature-based mechanism, as detailed in Section 2.2.3 – Mechanism by Which mTORC1 Inhibition Affects eIF4E Phosphorylation.

Section 2.3.4 – Concluding Remarks

Results of this project are novel because, to date, no computational model of mammalian translation has been produced. In contrast to a yeast model of translation produced by Firczuk et al. (2013), the model presented here is entirely qualitative. Moreover, the model of Firczuk and co-workers (2013) is very much limited in scope with the model encompassing only the translational machinery. The model presented here covers not only the translational machinery but also the signalling network that regulates this process and so is not comparable to the Firczuk et al. (2013) model. A quantitative model of the signalling pathways regulating translation initiation could not have been constructed due to a lack of quantitative parameters.
The model has been subject to multiple rounds of model refinement. As a result of this rigorous process, the predictive power has been raised from a questionable MCC = 0.2354 to a more sound MCC = 0.4558. This process has largely been achieved through increasing the cross-talk between signalling pathways. Similarly, the main motivation of establishing more cross-talk has been to reduce the number of FP readings whilst simultaneously increasing the number of TP determinations. This was of considerable importance given that these have more marked effects on the MCC values than either TN or FN numbers (Baldi et al., 2000).

It has been argued that MCC values are inherently more meaningful than the AC values, as the AC does not necessarily equate random predictions with zero (Baldi et al., 2000). Furthermore, Baldi and colleagues (2000) take the view that the AC is ‘unnecessary’. However, in this case with the MCC and AC being calculated in different ways, it was included as a check to ensure that the calculated predictive power was significant.

The formation of a benchmark dataset for each inhibitor can be viewed as adding considerable value to this work. It ultimately allowed for the predictive power of the model to be quantified and improved upon. By ensuring that stimulation of cells was restricted to the relevant dataset, it was possible to ensure that the inhibitors were the causative agents producing the effects. Moreover, it has been shown that, although cell type differences may be important in some cases, it is necessary to include multiple cell types in order to ensure a large enough number of data points. The importance of cell type differences can be reduced by verifying inhibitor effects across multiple studies using multiple cell types.

Over the course of the iterations to the model, two main hypotheses were formulated about the regions of differences between the Wild Type and the Rapamycin models. There
was an alteration in eIF4E phosphorylation when the Rapamycin was included in the model. This finding was relevant as it was a prediction that linked a signalling component of the model to the control of translation initiation and provided an opportunity to experimentally validate the model. Secondly, the model inferred that AKT1 phosphorylation should also be affected by mTORC1 inhibition with Rapamycin. Both of these predictions were be tested and the results are given in Chapter 3 – Experimental Validation of Model Predictions.

While the work presented in this chapter, offered the circumstances required for the mTORC1-mediated increase in eIF4E phosphorylation, the finding that the phosphorylation of this initiation factor is increased is not novel. Several studies have similarly noted this effect (Sun et al., 2005; Wang et al., 2007c; Jin et al., 2008a; Grosso et al., 2011; Marzec et al., 2011; Stead & Proud, 2013; Eckerdt et al., 2014). A more detailed summary of the findings of these papers is given in Section 3.1.2 – eIF4E Phosphorylation During mTORC1 Inhibition.

The 99.99% Binomial Confidence Intervals infer there is no effect of Rapamycin treatment on the activation of ERK1 or ERK2 and this is in agreement with Marzec et al. (2011) and Stead and Proud (2013). The notion of Rapamycin treatment having no effect on ERK1/2 activity (Oh et al., 2001b; Omura et al., 2005; Soares et al., 2013) is somewhat controversial, with other studies noting an increase (Carracedo et al., 2008; Chaturvedi et al., 2009). However, this can be explained by differences in the concentration of Rapamycin used (Chen et al., 2010). This aside, and in agreement with the work of Stead and Proud (2013), the model does predict that MNK2 mediates this alteration to eIF4E phosphorylation in an ERK1/2-independent manner.
Chapter 3 – Experimental Validation of Model Predictions
Section 3.1 – Introduction

In Chapter 2 – Model Creation & Refining two model predictions were generated, concerning the effect Rapamyin had upon AKT1 and eIF4E phosphorylation. In the final version of the model (detailed in Section 2.2.4 – Evaluation of the Predictive Power of the Final Model) Rapamycin was found to increase eIF4E phosphorylation (as measured by an increase in the frequency of trajectories leading to the production of eIF4E phosphorylation), from 0.492 in the Wild Type model to 0.805 in the Rapamycin model. Furthermore, AKT1 phosphorylation was shown to be decreased by the incorporation of Rapamycin into the model, with the frequency of trajectories reaching phosphorylated AKT1 falling from 0.569 in the Wild Type model to 0.006 in the Rapamycin model. Both of these behaviours required experimental validation to corroborate the predictions which are detailed in this chapter.

Section 3.1.1 – Rapamycin

Rapamycin, or Sirolimus, is a macrolide agent derived from isolates of Streptomyces hygroscopicus (Sehgal et al., 1975; Vézina et al., 1975; Rani et al., 2013). Rapamycin acts as an immunosuppressive agent and potent antifungal agent (Vézina et al., 1975; Martel et al., 1977; Kino et al., 1987a; Kino et al., 1987b (as cited by Abraham & Wiederrecht (1996)). Both behaviours can be explained by Rapamycin acting to arrest the cell cycle in the G1 phase (Heitman et al., 1991; Kawamata et al., 1998; Decker et al., 2003; Zinzalla et al., 2007).

The mechanism by which Rapamycin inhibits mTOR involves many of the 14 FK506-Binding Proteins (FKBP) but FKBP-12 predominates (Wandless et al., 1991; Chiu et al., 1994; Fruman et al., 1995; Kozany et al., 2009 (as cited by März et al., 2013)). Rapamycin forms a complex.
with FKBP12 which then binds to the FKBP12-Rapamycin Binding domain of mTOR located close to the mTOR active site (Vilella-Bach et al., 1999; Banaszynski et al., 2005; Yang et al., 2013a). As a consequence of this binding, the active site is hindered and thus, mTOR is inhibited (Yang et al., 2013a).

Section 3.1.2 – eIF4E Phosphorylation During mTORC1 Inhibition

The general information regarding phosphorylation of eIF4E is discussed in Section 1.2.2 – eIF4E Phosphorylation and The Role of ERK1/2 & p38. A number of studies have noted that mTORC1 inhibition causes an increase in eIF4E phosphorylation (Sun et al., 2005; Wang et al., 2007c; Jin et al., 2008a; Grosso et al., 2011; Marzec et al., 2011; Stead & Proud, 2013). Stead and Proud (2013) noted that MNK2a appears to have greater inherent activity through a reduction in the level of phosphorylation of the Serine residue at position 437 in response to Rapamycin exposure. Furthermore, this greater basal level of activation occurs without any input from the upstream mitogen-activated protein kinase, ERK1/2 (ibid). This finding is in keeping with the work by Eckerdt and co-workers (2014). These authors noted that whilst eIF4E phosphorylation increased following Rapamycin treatment, there was no concomitant increase in MAPK activity. Eckerdt et al. (2014) went further to find that the intracellular 4E-BP1 concentration had little or no bearing on this level of increased eIF4E phosphorylation.

Other studies noted that various other signalling pathways may be involved in this increase in eIF4E phosphorylation. Both Sun et al. (2005) and Wang et al. (2007c) noted that the PI3K pathway appears to be involved in this increase in eIF4E phosphorylation, with Wang and colleagues (2007c) suggesting that PI3K is capable of influencing MNK activity through a yet undefined mechanism. It is possible that this mechanism involves affecting the
phosphorylation of the Serine residue identified by Stead and Proud (2013). However, Marzec and colleagues (2011) found that this increase was independent of PI3K, ERK1/2 and p38 activity while Grosso et al. (2011) hypothesised that p38 activity was responsible for the phosphorylation of eIF4E in Rapamycin treated cells.

By interpreting the results presented in these works, one obvious and logical conclusion can be drawn: mTORC1 inhibition leads to an increase in eIF4E phosphorylation. It would also appear, based on these studies that the increase in activity of MNK1 or MNK2 that brings about this increase is dependent upon neither ERK1/2 nor p38 activity. This suggests that an, as yet, undefined but mTORC1-dependent, pathway is responsible for this increase.

Section 3.1.3 – Interplay Between mTORC1 & AKT1

The model of translational regulation requires not only an understanding of how the mTORC1 pathway is activated, but also how feedback mechanisms regulate the upstream component AKT1. By including such mechanisms in the model, it is then possible to better model the complex patterns of behaviour observed in vivo. Such mechanisms include the degradation of the Insulin Receptor Substrate (IRS)-1 (Haruta et al., 2000) and the mTORC1-mediated inhibition of mTORC2 (Julien et al., 2010) and account for the increase in the phosphorylation of AKT1 (Ikezoe et al., 2007; Breuleux et al., 2009; Chiarini et al., 2009), observed after inhibition of mTORC1.

The first mechanism concerns the down-regulation of a protein responsible for the induction of insulin receptor signalling, IRS-1 (Paz et al., 1996). Such down-regulation of IRS-1 was initially noted as being Rapamycin-dependent (Haruta et al., 2000; Pederson et al., 2001; Takano et al., 2001), as shown in Figure 3.1.3a, with S6K1 phosphorylating IRS-1 on multiple amino acid residues, including Ser-307 and Ser-636 (Shah & Hunter, 2006;
Tremblay et al., 2007; Zhang et al., 2008). This phosphorylation leads to the ubiquitin-mediated degradation of IRS-1 (Greene et al., 2003; Xu et al., 2008; Scheufele et al., 2014). As the activation of PI3K and AKT1 signalling is downstream of IRS-1, this mechanism limits AKT1 signalling (Wan et al., 2007). A similar mechanism has also been shown for the adaptor protein, Growth Factor Receptor-Bound Protein 10 (Hsu et al., 2011; Yu et al., 2011).
Figure 3.1.3a – Overview of the Feedback Mechanism between S6K1 and the Insulin Receptor Substrate 1 (IRS-1). IRS-1, in response to Insulin, becomes phosphorylated, which results in the activation of the PI3K signalling cascade and ultimately the activation of the mTORC1 pathway. Once S6K1 is phosphorylated by mTORC1 and PDK1, this kinase is capable of phosphorylating alternative sites of IRS-1, such as Ser-307 and Ser-636. This leads to the ubiquitination of IRS-1. From this it is easy to see how inhibition of mTORC1 results in an increase in the activity of PI3K signalling components including AKT1. It is worth noting that full activation of AKT1 requires both PDK1 and mTORC2. However, for simplicity, in this case, only PDK1 has been shown.
The second mechanism by which mTORC1 regulates AKT1, involves the negative regulation of mTORC2. S6K1 is capable of phosphorylating the Thr-1135 residue of the Rictor component of mTORC2, resulting in impaired activity towards mTORC2 substrates, including AKT1 (Julien et al., 2010) as shown in Figure 3.1.3b. Additionally, S6K1 and AKT1 have been shown to phosphorylate another mTORC2 component, mSIN1, to trigger mTORC2 inhibition (Liu et al., 2013).
Figure 3.1.3b – Overview of the Feedback Mechanism that Exists between S6K1 and mTORC2. The full activation of AKT1 requires phosphorylation of two sites (Ser-473 and Thr-308) by mTORC2 (Sarbassov et al., 2005) and PDK1 (Wan & Helman, 2003; Tsuchiya et al., 2013; Dangelmaier et al., 2014), respectively. Downstream of AKT1 lies mTORC1 and the kinase S6K1. mTORC1, along with PDK1, are responsible for the full activation of S6K1. A feedback mechanism between mTORC2 and S6K1 has recently been identified by Julien et al. (2010). In this mechanism the mTORC2 component, Rictor, is phosphorylated via S6K1 on Thr-1135. A second, analogous mechanism may involve the S6K1- and AKT1-dependent phosphorylation of another mTORC2 component, mSIN1 (Liu et al., 2013). Both mechanisms would result in an inhibition of mTORC2. Phosphorylation, resulting in activation, is denoted with a blue circle whilst those having an inhibitory effect are shown by red circles.
Whilst both mechanisms could explain an increase in AKT1 phosphorylation during mTOR inhibition, neither can explain a decrease. As such, it is necessary to consider that Rapamycin may affect mTORC2. Only nascent mTOR, or mTOR from disassembled mTORC1 and mTORC2, is free to interact with the FKBP12-Rapamycin complex to be incorporated into inhibited mTORC2 (Brown et al., 1994, Sabatini et al., 1994; Sabers et al., 1995 (as cited by Sarbassov et al., 2006)). Diminished AKT1 phosphorylation has been noted after prolonged or excessive Rapamycin exposure, suggesting Rapamycin is perhaps capable of inhibiting mTORC2 in a cell-type or dose-dependent manner (Sarbassov et al., 2006; Zeng et al., 2007; Wang et al., 2008a; Chen et al., 2010). In keeping with this, decreased AKT1 phosphorylation was seen at higher Rapamycin concentrations while increased at lower concentrations (Chen et al., 2010). However, the ability of Rapamycin to inhibit mTORC2 is dependent upon the ratio of FKBP12 to FKBP51 expression in cells. Cell lines with a higher ratio of these proteins are more receptive to mTORC2 inhibition (Schreiber et al., 2015).

Alongside the aim of creating a model of the translation initiation regulatory pathways, this project must experimentally validate predicted behaviours of the model. The information presented in the introduction to this chapter highlights that the regulation of AKT1 and eIF4E by mTORC1 is not fully understood. This chapter aimed to validate the predictions of the model but not to provide a hypothesised mechanism. Work presented in Chapter 2 – Model Creation & Refining, details a computationally elucidated mechanism by which eIF4E phosphorylation is affected by Rapamycin.
Section 3.2 – Materials & Methods

Section 3.2.1 – Growth & Passage of RAW 264.7 Murine Macrophage Cells

RAW 264.7 cells are an Abelson Murine Leukaemia Virus-transformed murine macrophage cell line (Raschke et al., 1978). These cells (from European Collection of Cell Cultures, Product Number: 91062702, Lot: 13J020) were cultured in a vented T75 flask (Nunc™, Thermo Fisher Scientific Inc., Rockford, IL (USA)) using media containing 89% high glucose (4.5g per litre) Delbecco’s Modified Eagle Medium, containing Sodium Pyruvate and L-Glutamine, 10% Foetal Bovine Serum and 1% Penicillin and Streptomycin (Gibco®, Life Technologies Corp., Carlsbad, CA (USA)). The cells were incubated at 37°C with 5% Carbon Dioxide in a humidified Galaxy S incubator (Wolf Laboratories Ltd, York (UK)). Cells were passaged at around 70% confluence, with each passage requiring new media, and the cells were detached from the T75 flask using 3mm glass beads (Sigma-Aldrich Co. LLC, St. Louis, MO (USA)).

Section 3.2.2 – Determination of Rapamycin Cytotoxicity

Given the qualitative behaviour of the model, the action of inhibitors can best be described as ‘all or nothing’ with regard to the effect these chemicals have on the main target. To this end it was decided that a high concentration of Rapamycin would be used. Jin et al. (2009) had already established data concerning the viability of RAW 264.7 cells with a combination of Rapamycin, up to a concentration of 10μM, and LPS. This paper was used to guide the selection of Rapamycin concentrations to test and three concentrations of Rapamycin would be tested (10μM, 1μM and 0.1μM) in the absence of LPS.

Cytotoxicity of Rapamycin was determined by measuring the concentration of Adenosine Trisphosphate in the cell media after treatment. For this purpose the CellTiter Glo®
Luminescent Cell Viability Assay kit (Promega UK, Promega Corp., Madison, MI (USA)) was used. The Rapamycin (Tocris Biosciences, R&D Systems Inc., Minneapolis, MN (USA)) was reconstituted in Dimethyl Sulfoxide (DMSO) (Sigma-Aldrich Co. LLC, St. Louis, MO (USA)). Rapamycin was diluted to the desired concentration in RAW 264.7 cell media (the composition of which is detailed in Section 3.2.1 – Growth and Passage of RAW 264.7 Murine Macrophage Cells). Care was taken to ensure that the final concentration of DMSO was not greater than 1%.

At approximately 70% confluence, RAW 264.7 cells were seeded in a white-walled 96-well plate (Corning Incorporated, Tewksbury, MA (USA)) and incubated overnight. The cells were treated with three concentrations of Rapamycin (10μM, 1μM and 0.1μM) and two controls of 1% DMSO and cell media. Additionally, to control for potential inherent luminescent properties of Rapamycin, DMSO or indeed the media, some samples containing no cells were also included. In all cases, samples were seeded in triplicate.

After three hours of incubation, the plate was left at room temperature for 30 minutes. Reconstituted CellTiter Glo® Reagent was added to each well in a one-to-one ratio with the amount of media already present in the well. Following a ten minute incubation at room temperature, the luminescence of each well was recorded by a Glomax® 96 Microplate Luminometer (Promega UK, Promega Corp., Madison, MI (USA)).

The results of this assay were analysed by subtracting the average luminosity of the wells containing no cells from each of the readings for the corresponding treatment. The data was then normalized to the DMSO control. Statistical analysis of this data will be discussed in Section 3.2.7 – Statistical Analysis.
Section 3.2.3 – Cell Harvesting & Bicinchoninic Acid (BCA) Assay

RAW 264.7 cells, at around 70% confluence in a T75 flask, were seeded into untreated 6-well plates (Nunc™, Thermo Fisher Scientific Inc., Rockford, IL (USA)) and incubated overnight at 37°C with 5% Carbon Dioxide. The following day, the cells were treated with either 10μM Rapamycin or 1% DMSO for a period of either one or three hours (as stated in the subsequent Section 3.2.4 – Human Phospho-MAPK Array and Section 3.2.5 – eIF4E/Phosphorylated eIF4E & AKT1/Phosphorylated AKT1 Western Blot Analysis).

Prior to lysis it was necessary to remove the treatment media and ensure as much serum was removed as possible, thus ensuring accurate measurement of cellular protein concentration. To this end the cells were washed twice in chilled, sterile one-times Phosphate-Buffered Saline (PBS) solution. The cells were then lysed in 70μl of Lysis Buffer 6, taken from the Human Phospho-MAPK Array kit (R&D Systems, Inc., Minneapolis, MN (USA)), and incubated at 4°C for 30 minutes to complete lysis. The lysate was centrifuged at 4°C for 5 minutes and the pellet discarded.

The supernatant would later be used for both Western Blot analysis and the Human Phospho-MAPK Arrays. Prior to this however, it was necessary to determine the concentration of protein within the lysates. This was done using a BCA Assay kit (Pierce™, Thermo Fisher Scientific, Inc., Rockford, IL (USA)) and a 96-well Nunclon™ plate (Nunc™, Thermo Fisher Scientific Inc., Rockford, IL (USA)). The standard curve was produced according to the manufacturer’s instructions. 25μl of each sample was loaded, in duplicate, onto the plate. The resulting colour change was analysed with a Victor™ X5 Multilabel Plate Reader (Perkin Elmer, Inc., Waltham, MA (USA)). Based upon the results of this assay, the
protein concentration in the samples could be adjusted to either 400µg or 50µg of protein per ml, for the Human Phospho-MAPK Array or Western Blotting, respectively.

**Section 3.2.4 – Human Phospho-MAPK Arrays**

RAW 264.7 cells were treated for one hour with either 10µM Rapamycin or 1% DMSO and harvested, according to the method outlined in the previous section. Following the BCA assay, the protein concentrations in both lysates were adjusted to 400µg of protein per millilitre with Milli-Q® Ultrapure (EMD Millipore Corp., Merck KGaA, Darmstadt (DE)) water and carried out in accordance with the manufacturer’s instructions. This method required 20µl of the Detection Antibody Cocktail being incubated with the lysate at room temperature for one hour prior to incubation with the membrane overnight at 4°C. The array was subsequently developed onto Super-RX radiographic film (Fujifilm Corporation, Tokyo (JP)). These arrays were done in triplicate, each with new lysates.

The results of the array were analysed in the software ImageJ (Rasband, 1997-2012). This software analyses the pixel colour density of each spot on the membrane. As each protein is represented on the membrane in duplicate, the average is taken for each protein. A value for the background pixel colour density is also taken and subtracted from the average value for each protein. Normalization of all the data to the DMSO-treated controls allows for the identification of trends.

**Section 3.2.5 – eIF4E/Phosphorylated eIF4E & AKT1/Phosphorylated AKT1 Western Blot Analysis**

Once determined by the results of the BCA assay, the protein concentration of each lysate was adjusted to 50µg of protein with Milli-Q® Ultrapure water and Red Loading Buffer (New
England BioLabs Inc., Ipswich, MA (USA)). Protein samples were separated using 4-20% gradient Mini-PROTEAN® TGX (Bio Rad Laboratories, Inc., Hercules, CA (USA)) Sodium Dodecyl Sulphate (SDS) Polyacrylamide Gel Electrophoresis (SDS-PAGE) gels with an SDS-based Running Buffer (for solutions which are made in-house, the constituents are shown in Supplementary Information 4. Prestained Protein Marker (P7706) (New England BioLabs Inc., Ipswich, MA (USA)) with a resolution of between 190kDa and 11kDa was included alongside the samples. Samples were heated to 100°C for five minutes prior to loading. Proteins were transferred from the gel onto Immobilon-P (EMD Milipore Corp, Merck KGaA, Darmstadt (DE)) Polyvinylidene Fluoride (PVDF) membrane at 100V for 60 minutes in Transfer Buffer solution. Ponceau staining solution was used to confirm the transfer from the gel to the PVDF membrane. The membrane was cut to avoid the need for stripping the membrane between primary antibodies. A blocking solution was added to the membrane and incubated for one hour, or in the case of the primary α-Tubulin antibody, overnight at 4°C. The antibodies used in testing the model validity are shown in Table 3.2.5a.

Following the appropriate incubation period for each antibody, it was removed and washed in PBS-Tween solution. The secondary antibody solutions were prepared in 0.5% (w/v) Marvel® milk powder (Premier Foods PLC, St. Albans (UK)) in PBS-Tween solution. For α-Tubulin and both eIF4E and Phospho-eIF4E antibodies, the anti-mouse (Dako Denmark A/S, Glostrup (DK)) and anti-rabbit (Dako Denmark A/S, Glostrup (DK)) secondary antibodies, respectively, were incubated at room temperature for one hour. All antibodies used in the testing of the AKT1 phosphorylation prediction required anti-rabbit secondary antibodies (Dako Denmark A/S, Glostrup (DK)).
<table>
<thead>
<tr>
<th>Antibody</th>
<th>Diluting Solution</th>
<th>Amount of Antibody</th>
<th>Incubation Period</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rabbit Phospho-&lt;br&gt;eIF4E (Ser-209)&lt;br&gt;#9741</td>
<td>5% BSA in PBS-Tween</td>
<td>15μl (1:667)</td>
<td>Over Night</td>
<td>Cell Signaling Technology, Inc., Danvers, MA (USA)</td>
</tr>
<tr>
<td>Rabbit eIF4E&lt;br&gt;#9742</td>
<td>5% BSA in PBS-Tween</td>
<td>10μl (1:1000)</td>
<td>Over Night</td>
<td>Cell Signaling Technology, Inc., Danvers, MA (USA)</td>
</tr>
<tr>
<td>Rabbit AKT1&lt;br&gt;#9272</td>
<td>5% BSA in PBS-Tween</td>
<td>10μl (1:1000)</td>
<td>Over Night</td>
<td>Cell Signaling Technology, Inc., Danvers, MA (USA)</td>
</tr>
<tr>
<td>Rabbit Phospho-AKT1 (Ser-473)&lt;br&gt;#4060</td>
<td>5% BSA in PBS-Tween</td>
<td>15μl (1:667)</td>
<td>Over Night</td>
<td>Cell Signaling Technology, Inc., Danvers, MA (USA)</td>
</tr>
<tr>
<td>eIF6&lt;br&gt;10291-1-AP</td>
<td>5% BSA in PBS-Tween</td>
<td>10μl (1:1000)</td>
<td>Over Night</td>
<td>Proteintech Group, Inc., Chicago, IL (USA)</td>
</tr>
<tr>
<td>α-Tubulin&lt;br&gt;T9026</td>
<td>PBS-Tween</td>
<td>2μl (1:5000)</td>
<td>One Hour</td>
<td>Sigma-Aldrich Co. LLC, St. Louis, MO (USA)</td>
</tr>
</tbody>
</table>

Table 3.2.5a – Summary of the Antibody Solutions Used in Testing eIF4E and AKT1 Phosphorylation.
The PVDF membranes were then washed in PBS-Tween solution and incubated for one minute in SuperSignal West Pico Chemiluminescent Substrate (Pierce®, Thermo Fisher Scientific Inc., Rockford, IL (USA)). Each membrane was then developed onto Super-RX radiographic film. As with the Human Phospho-MAPK Array kit, quantification of each band was carried out in ImageJ and normalized to the DMSO lysate which acts as a control.

Section 3.2.6 – Re-Evaluation of the Model Based Upon the Results of the Human Phospho-MAPK Arrays

Using the methods described in Chapter 2 – Model Creation & Refining the binary model was interrogated against the data collected in the Human Phospho-MAPK Array kits. Simulations were run looking at the activation of effectors that are found in both the model and the array. The length of simulations was adjusted to ensure that the frequency of activation of each effector was around 0.5 in the Wild Type model. Simulations of the model containing Rapamycin was run for the same lengths of time. By comparing the 99.99% Binomial Confidence Intervals for each effector it was possible to find regions of significant difference. As previously stated, the Binomial Confidence Intervals were calculated in the statistical software R (R Core Development Team, 2013).

To ensure that the only rigorous differences in the activation are found, only array constituents displaying a difference that remains significant after the Holm-Sidak Correction (as described in Section 3.2.7 – Statistical Analysis) was considered. Given the qualitative nature of the model, only an increase, decrease or no effect were recorded.
Section 3.2.7 – Statistical Analysis

This section will cover the methods used to evaluate the significance of results obtained in the cell viability assay, array or in Western Blotting. Most of the statistical analysis was done using GraphPad Prism 6 (Version 6.05, for 64 bit Windows, GraphPad Software, Inc., La Jolla, CA (USA)). In all cases the statistical analysis applies only to normalized data.

The CellTiter Glo® Assay generated results for five independent variables (namely the three inhibitor concentrations, one DMSO treatment and the untreated, media-only, control). With this in mind, it was necessary to determine the distribution of the data. Whilst it was not possible to determine, using GraphPad Prism, if the data points for each variable were normally distributed, it is possible to determine if the Standard Deviation of each group displays no significant difference. This can be done using the Brown-Forsythe Test (Brown & Forsythe, 1974) calculated as part of the Analysis of Variance (ANOVA) function. Given that the normality of the data cannot be determined, and assuming that the Standard Deviation of each group is not significantly different, the Kruskal-Wallis Test should be used in preference of a one-way ANOVA (Kruskal & Wallis, 1952). No significant difference being found between concentrations of Rapamycin or between the concentrations and the DMSO or media controls, infers that cell viability is not affected by the presence of this inhibitor.

The data collected from both the Human Phospho-MAPK Array data and Western Blot analysis is suitable to analysis with t-tests. Once the average intensity of each phosphoprotein on the array is known and normalized (in triplicate), it can be analysed in GraphPad Prism using the multiple t-test function. In order to correct for individual tests returning significant results solely due to multiple testing, a Holm-Sidak Correction (Sidak, 1967; Holm, 1979) was applied. The Western Blot data was also corrected, but in this case
a Bonferroni Correction was applied. The \( p \)-value, calculated using an unpaired, non-parametric t-test, was multiplied by the number of t-tests it was possible to carry out on a dataset (Emily et al., 2009). In this case, the \( p \)-value was multiplied by three as the results were obtained in triplicate. This method of \( p \)-value correction is preferable as it is viewed as very orthodox and results in ‘over-corrected’ (McIntyre et al., 2000 (As cited, and quoted, by Klein et al., 2005)) test statistics.

**Section 3.3 – Results**

**Section 3.3.1 – Cytotoxicity of Rapamycin**

As the behaviour of a qualitative model treats the action of inhibitors as binary, in that the result is complete inhibition or complete activity of the target effector, it was decided that for experimental validation of predictions a high concentration of inhibitor would be used. To determine which concentration of Rapamycin does not have any effect on cell viability, the CellTiter Glo® Assay was used which measures the concentration of ATP released from cells lysed due to exposure to Rapamycin.

As shown in Figure 3.3.1a, it appears that the mTORC1 inhibitor has mild, positive effects on cell viability. As determined by the Kruskal-Wallis test though, none of these effects are significantly different from either the 1% DMSO treatment or the effects of media alone with the resulting \( p \)-values never reaching below 0.5437 (comparing 0.1\( \mu \)M Rapamycin and 1% DMSO). Although none of the concentrations tested produced a significant effect on cell viability, a concentration of 10\( \mu \)M was chosen as it represented the highest concentration tested. Consequently, this concentration would be most in keeping with the qualitative behaviour of the model in which the action of inhibitors is viewed in an all or nothing manner.
Figure 3.3.1a – Evaluation of the Effects of Three Hour Treatments with Differing Concentrations of Rapamycin (Rapa) or 1% DMSO on the Murine Macrophage RAW 264.7 Cells. The effects on viability were measured by changes to the amount of ATP present in the supernatant following cell lysis. The concentration of DMSO was never higher than 1%. All data is given as relative to the effect of 1% DMSO. The Kruskal-Wallis test was used to evaluate significance, with ‘ns’ denoting non-significant. Error bars represent Standard Error of the Mean values for each dataset.
Section 3.3.2 – Evaluation of the Human Phospho-MAPK Arrays

The binary model of translational control demonstrated that whilst eIF4E phosphorylation was affected by the presence of an mTORC1 inhibitor, the MAPKs, ERK1/2 and p38, were not. Similarly, the level of phosphorylated AKT1 is predicted to be reduced during exposure to a high concentration of this macrolide. It was important to establish if either MAPK or AKT1 was affected by treatment with Rapamycin. The Human Phospho-MAPK Array kits offered a viable method to test all of these effects in a simple way. RAW 264.7 cells were treated with either 10μM Rapamycin or 1% DMSO for one hour. The array demonstrates, as shown in Figure 3.2.2a, that Rapamycin treatment has broad implications for MAPK signalling in RAW 264.7 cells.
Figure 3.3.2a – Evaluation of the Broad Early Effects of 10μM Rapamycin Treatment to the Murine Macrophage RAW 264.7 Cells. Following treatment of RAW 264.7 cells with either 10μM Rapamycin or 1% DMSO for one hour, cells were lysed using the Lysis Buffer 6 from the Human Phospho-MAPK Array Kit (R&D Systems Inc., Minneapolis, MN (USA)). After the protein concentration of each lysate was adjusted to 400μg, the array was carried out in accordance with the manufacturers’ instructions. The results were analysed in ImageJ. T-tests were carried out on each phosphoprotein and were adjusted for multiple comparisons using the Holm-Sidak Correction. Those phosphoproteins that were significant after the t-test are denoted with *. If they remained significant after the Holm-Sidak Correction they are denoted with **. Error bars represent the Standard Error of the Mean.
It was necessary to prove that Rapamycin was having an effect on the marketed, target substrate. The Human Phospho-MAPK Array allows for measuring the level of mTOR phosphorylated on the Ser-2448 residue. The phosphorylation of this Serine residue has been linked to S6K1 and 4E-BP1 phosphorylation (Dickinson et al., 2011) and has been used as a marker of mTORC1 activity (Altman et al., 2011). This phosphorylation site has been shown to be responsive to Rapamycin treatment (Copp et al., 2009; Tyler et al., 2009; Carayol et al., 2010). In this MAPK array, it was established that Rapamycin produced a statistically significant 41% decrease ($p=1.95 \times 10^{-4}$) in the phosphorylation of Ser-2448 of mTORC1, as shown in Figure 3.3.2a. This $p$-value is that calculated after a Holm-Sidak Correction was applied.

Establishing the activity of ERK1/2 and p38 in response to Rapamycin was an important determination with regards to experimentally validating the alterations to the level of phosphorylated eIF4E in Rapamycin treatment. As shown in Figure 3.3.2a, neither kinase was affected by Rapamycin treatment. The effect of Rapamycin on both ERK1/2 and p38 is not fully understood. ERK1/2 has been shown to be subject to a diverse array of effects, with one study noting an increase (Chaturvedi et al., 2009), another a decrease (Chen et al., 2010) and several displaying no effect (Omura et al., 2005; Soares et al., 2013). Although the work of Chaturvedi et al. (2009) and Chen et al. (2010) indicated that the concentration of Rapamycin may be important in determining the direction of change, the work presented here is more in keeping with Omura et al. (2005) and Soares et al. (2013). Several studies have noted that a feedback mechanism exists between ERK1/2 and mTORC1 (Carracedo et al., 2008; Melemedjian et al., 2013), however the finding that Rapamycin does not affect
ERK1/2, after a treatment of one hour at least, suggests that this mechanism could either be absent in RAW 264.7 cells or become important much later.

As with ERK1/2, the published view of the relationship between Rapamycin and p38 is equally contradictory. Oh et al. (2001b) and Omura et al. (2005) both found that p38 activity was increased by Rapamycin. Despite these publications contradicting the work presented here, several other studies are in agreement and note that Rapamycin has no effect on p38 activity (Cuenda & Cohen, 1999; Huang et al., 2003). Whilst no definitive explanation for this can be given, it may reflect differences in cell type.

With regards to regions of significant difference, 18 regions were noted as being different and, of these, nine regions were found to remain significant after the Holm-Sidak Correction was applied. These regions are denoted with * and ** respectively in Figure 3.3.2a. Of particular interest here is the finding that both AKT1 and AKT2 display significant inhibition of 65.7% and 50.5%, respectively. Whilst the phosphorylation of both AKT1 and AKT2 was significantly reduced (with \( p = 0.018 \) and \( p = 0.014 \)) neither were found to remain significant following the Holm-Sidak Correction. However, the reduction (on average of 51.5%) of AKT Pan phosphorylation, representing the total of phosphorylated AKT1, AKT2 and AKT3, in response to Rapamycin remains significant after this correction. Other differences that are significant in this robust sense are JNK3, MKK3, MKK6, Ribosomal Protein S6 Kinase α-4 (MSK2), p38δ, p53 and RSK2.

The results outlined in Figure 3.3.2a clearly demonstrate that Rapamycin has a pronounced inhibitory effect on a broad range of cellular targets. This could reflect the central role of mTORC1 in the signalling network of the cell. The work of Caron and co-workers (2010) clearly demonstrates that mTORC1 is linked to a wide variety of cellular processes and
pathways. It is feasible that short-term exposure to Rapamycin, as used here, would have such pronounced effects on the signalling pathways examined by with these Human Phospho-MAPK Arrays.

Section 3.3.3 – Alterations to eIF4E Phosphorylation Levels

Phosphorylated eIF4E was not present on the Human Phospho-MAPK Arrays and thus it was not possible to test the effects of Rapamycin on the level of eIF4E phosphorylation using the approach described in the previous section. In order to test the hypothesis that inhibition of the mTOR pathway by Rapamycin alters the level of eIF4E phosphorylation, immunoblotting was used. Murine Macrophage RAW 264.7 cells were treated with either 10μM Rapamycin or 1% DMSO for one hour and lysed using the Lysis Buffer 6 supplied with the Human Phospho-MAPK Array kit. After this time period, the protein concentration was adjusted to 50mg per millilitre. Although this was only conducted once, there was a mild, 18% increase in phosphorylated eIF4E in response to Rapamycin treatment (data not shown).

In order to further investigate this effect, RAW 264.7 cells were treated for three hours with either 10μM Rapamycin or 1% DMSO. As expected, quantification of the pixel intensity in ImageJ revealed that the increase in treatment length resulted in a greater increase in phosphorylated eIF4E of, on average, 44.0% (range of between 36.4% and 52.2%, \( p = 0.0318 \)) in the Rapamycin-treated RAW 264.7 cell lysates (as shown in Figure 3.3.3ai-ii). The significance of this increase was determined using a two-tailed t-test with a Bonferroni Correction being applied to the resulting \( p \)-value. The Bonferroni Correction was applied and the results of the two-tailed t-test were multiplied by three to account for the three independent, biological replicates conducted.
Figure 3.3.3a – Rapamycin-Induced Increase in Phosphorylation of eIF4E. RAW 264.7 Cells were treated with 10μM Rapamycin or 1% DMSO for three hours. 50mg per ml of protein, from each lysate, were separated using SDS-PAGE and phosphorylated eIF4E (25kDa), eIF4E (25kDa) and α-Tubulin (50kDa) were detected using immunoblotting. i) A representative blot of the increase in eIF4E phosphorylation due to Rapamycin treatment. The levels of total eIF4E and the α-Tubulin loading control remain constant. ii) The pooled quantitative data of three independent replicates showing the mean ± Standard Error of the Mean. This data is expressed as the level of phosphorylated eIF4E relative to the 1% DMSO control. This quantification was tested using a two-tailed t-test after which a Bonferroni Correction was applied to the calculated p-value.
Section 3.3.4 – Confirmation of the Effect Rapamycin has Upon AKT1 Phosphorylation

Despite the fairly conclusive result generated in the Human Phospho-MAPK Arrays regarding a decrease in AKT1 phosphorylation on the Ser-473 residue, the effect was further investigated using immunoblotting. RAW 264.7 cells were treated with either 10μM Rapamycin (in 1% DMSO) or 1% DMSO for three hours. As was the case for the eIF4E/Phospho-eIF4E immunoblotting experiment, it was decided to attempt to accentuate the affect by increasing the length of Rapamycin treatment to three hours.

As shown in Figure 3.3.4a, only a small, non-significant difference in the levels of Ser-473 phosphorylated AKT1 was found between the Rapamycin-treated and control lysates. On average the level of phosphorylated AKT1 was 20.4% (range of between 3.2% and 31.5%, p = 0.4344) higher in the Rapamycin-treated lysates. Reasons for this lack of significant difference are discussed in Section 3.4.2 – Validation of the AKT1 Phosphorylation Prediction.
Cells were treated with either 10μM Rapamycin (in 1% DMSO) or 1% DMSO for three hours. 50mg per ml of protein from each lysate were separated using SDS-PAGE and phosphorylated AKT1 (60kDa), AKT1 (60kDa) and eIF6 (27kDa) were detected using immunoblotting. i) The blot shown is representative of the results obtained. Only very low levels of AKT1 phosphorylation (Ser-473) were noted. The levels of the eIF6 loading control and total AKT1 remain constant. ii) The pooled data collected from three independent replicates ± the Standard Error of the Mean. The data shown was quantified in the software ImageJ and is expressed as a ratio of the DMSO control. The p-value shown was calculated using a Bonferroni Correction applied to the results of a two-tailed t-test.
Section 3.3.5 – Comparison of Model Results to Array Data

Out of all the phosphoproteins recorded on the array, 13 were found to be present in the binary model. As shown in Figure 3.3.5a, nine effectors displayed behaviour that was in agreement with the model. Simulations performed to generate these results were run using QSSPN on the final version of the binary model, the overall results of which are detailed in Section 2.2.4 – Evaluation of the Predictive Power of the Final Model. The results shown in this section were included in the overall analysis of the final model.

The comparison of the model to the array data was done to show that the behaviour of the model correlates to experimental outcomes detailed in this chapter. Furthermore, the additional testing of the model can be seen as adding further weight behind the view that the model is capable of mimicking experimental data. The Matthews’ Correlation Coefficient (MCC) for this comparison is calculated using the equation given in Section 2.1.4.2 – Evaluation of the Predictive Power of the Model. The MCC calculated for available array data is 0.3958.
### Table

<table>
<thead>
<tr>
<th>Effector</th>
<th>Behaviour of the Effector in the Array</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>ppERK1</td>
<td>No Effect</td>
<td>TN</td>
</tr>
<tr>
<td>ppERK2</td>
<td>No Effect</td>
<td>TN</td>
</tr>
<tr>
<td>pRSK1</td>
<td>No Effect</td>
<td>TN</td>
</tr>
<tr>
<td>p38α</td>
<td>No Effect</td>
<td>TN</td>
</tr>
<tr>
<td>pJNK1</td>
<td>No Effect</td>
<td>TN</td>
</tr>
<tr>
<td>mTORC1-Act</td>
<td>Inhibition Demonstrated</td>
<td>TP</td>
</tr>
<tr>
<td>ppS6K</td>
<td>No Effect</td>
<td>FP</td>
</tr>
<tr>
<td>ppAKT</td>
<td>Inhibition Demonstrated</td>
<td>TP</td>
</tr>
<tr>
<td>pMKK3</td>
<td>Inhibition Demonstrated</td>
<td>FN</td>
</tr>
<tr>
<td>pMKK6</td>
<td>Inhibition Demonstrated</td>
<td>FN</td>
</tr>
<tr>
<td>pMSK2</td>
<td>Inhibition Demonstrated</td>
<td>FN</td>
</tr>
<tr>
<td>pCREB</td>
<td>No Effect</td>
<td>TN</td>
</tr>
<tr>
<td>pGSK3</td>
<td>Inhibition Demonstrated</td>
<td>TP</td>
</tr>
</tbody>
</table>

**Figure 3.3.5a – Comparison of Array Data with the Behaviour of the Binary Model.** In keeping with the qualitative nature of the model, the array data was analysed as showing 10μM Rapamycin having no effect or being an inhibitor or activator of an effector. The Binomial Confidence Intervals calculated for each effector allowed for differences to be found between the Wild Type or Rapamycin models. Comparing the two datasets permitted a quantitative analysis of the predictive power of the model.
Section 3.4 – Discussion

Section 3.4.1 – Validation of the eIF4E Phosphorylation Prediction

The mTORC1 pathway is integral to growth and proliferation of cells (Laplante & Sabatini, 2013) and, as of 2010, can be viewed as including close to one thousand molecular species (Caron et al., 2010). Despite this broad knowledge base regarding the interactions of mTORC1 with other signalling components and pathways, the exact nature of the mTORC1 input into many processes, for example eIF4E phosphorylation, remains unclear.

In keeping with the prediction generated by the model, a number of studies have reported an increase in eIF4E phosphorylation in response to mTORC1 inhibition (Sun et al., 2005; Wang et al., 2007c; Jin et al., 2008a; Grosso et al., 2011; Marzec et al., 2011; Stead & Proud, 2013). Marzec and colleagues (2011) noted that this increase may be independent of ERK1/2 or p38α activity. However, Grosso et al. (2011) argued that the remaining p38 activity following mTORC1 inhibition, may account for the apparent increase in eIF4E phosphorylation. The former is supported by the work of Stead and Proud (2013). In this, they argued that the interaction between ERK1/2 and MNK2 is not affected by Rapamycin-mediated mTORC1 inhibition (Ibid).

The experimental results derived in this project have shown that the Rapamycin treatment produces a notable increase in eIF4E phosphorylation after just three hours of exposure. Whilst no attempt was made to distinguish whether this was through MNK1 or MNK2 (discussed in Section 3.4.3 – Problems Encountered & Troubleshooting) the activity of both ERK1/2 and p38 were recorded in response to Rapamycin treatment. After one hour of exposure there was no indication of the activity of either kinase increasing.
It can be argued that the work here is in agreement with that shown by Stead and Proud (2013). In this work it was noted that Rapamycin was able to increase eIF4E phosphorylation through a MNK2-dependent but ERK1/2-independent manner. Here the array data has shown no effect on the levels of phosphorylated ERK1/2 or p38α. In fact, even one hour of Rapamycin treatment seemed to produce a reduction, albeit non-significantly, in the levels of these phosphorylated kinases. In keeping with this concept, the idea that at higher concentrations (at greater than 1μM) of Rapamycin there is concomitant inhibition of mTORC1 and mTORC2 which leads to inhibition of not just ERK1/2 but also AKT1 has been demonstrated (Chen et al., 2010). This suggests that increasing the treatment length would produce ERK1/2 inhibition rather than activation.

As already noted in Section 1.2.2 – eIF4E Phosphorylation and The Role of ERK1/2 & p38, eIF4E phosphorylation is now being seen as a mechanism by which genes associated with stress responses are being preferentially up-regulated (Andersson & Sundler, 2006; Herdy et al., 2012; Royall et al., 2015). It is also of interest to note the MNK kinases are also capable of phosphorylating the Sprouty proteins (DaSilva et al., 2006; Sharma et al., 2012) and the Heterogenous Nuclear Ribonucleoprotein A1 (hnRNP A1) (Buxadé et al., 2005). Both of these have roles in the physiological response to cellular stress. The Sprouty proteins are required for the negative regulation of Interferon-Stimulated Gene 15 and p38 (Sharma et al., 2012), and the response is MNK-dependent. Furthermore, phosphorylated hnRNP A1 is known to play a role in the movement of mRNA to stress granules (Guil et al., 2006).

The role of phosphorylated eIF4E in the physiological response to stress and the finding that inhibition of mTORC1 leads to increased resting activity of MNK2, suggest that mTORC1 is capable of suppressing the cellular stress response through modulation of MNK activity, as
shown in Figure 3.4.1a. This hypothesis is supported, albeit indirectly, by the finding that S6K1 is capable of inhibiting Transforming Growth Factor-β-Activated Kinase 1 (TAK1) resulting in increased signalling through Nuclear Factor-κB (Kim et al., 2014). Additionally, Shin et al. (2013) found that TAK1 was also capable of inhibiting S6K1 activation by preventing an interaction with the RAPTOR subunit of mTORC1. These findings support the hypothesis presented here by demonstrating a link between cellular stress responses and a major pathway regulating cellular growth. These works highlight that cell growth pathways and cell stress pathways appear to be able to negatively regulate each other as one becomes more necessary. Upon mTORC1 activation, there is less requirement for the cell stress responses, and thus, there is less need for eIF4E phosphorylation to facilitate the associated gene expression change. In this scenario the basal activity of MNK2, and the inducible MNK1 (Ueda et al., 2004), would be kept lower. However, if the stress response is required, the level of mTORC1 activity could be decreased and bring about MNK1 induction as well as an increase in the basal level of MNK2 activity to allow such as a response.
Figure 3.4.1a – Hypothesis Linking mTORC1 to the MNK-Dependent Stress Response. Given that eIF4E phosphorylation, via MNK2 (Stead & Proud, 2013), has been shown to increase during inhibition of mTORC1, it supports the hypothesis that inhibition of cell growth induces the cellular stress response. In such a scenario, inhibition of mTORC1 leads to an increase in the basal activity of MNK2. As a result, the phosphorylation of eIF4E and other substrates of the MNK kinases, hnRNP A1 and the Sprouty proteins occur. eIF4E phosphorylation would promote the translation of mRNAs associated with the stress response. hnRNP A1 would facilitate this by promoting the formation of Stress Granules. The Sprouty proteins may then, through a feedback mechanism, serve to return the cell to an unstressed state once the inhibition of cell growth was removed. Links experimentally shown, either in the literature or here, are denoted by the unbroken green arrows, also denoting hypothesised links. The red arrow illustrates the suppressive effect active mTORC1 has upon MNK2 to lower the level of basal activity.
Section 3.4.2 – Validation of the AKT1 Phosphorylation Prediction

Due to the presence of a variety of feedback interactions, AKT1 can be viewed as being very tightly regulated by mTORC1. One feedback loop is mediated by the phosphorylation, and subsequent degradation, of the IRS-1 by S6K1 (Tremblay et al., 2007; Xu et al., 2008; Zhang et al., 2008). In this case, AKT1 is inhibited by this phosphorylation (Shah et al., 2004; Xu et al., 2008). The importance of this mechanism of feedback inhibition can be seen in cases of Simian Virus 40 infection in which the proteolytic degradation of IRS-1 is inhibited. This results in adherent activation of AKT1 and the ERK1/2 pathway components (Hartmann et al., 2014). In the second case, S6K1 phosphorylates the Rictor subunit of mTORC2 (Julien et al., 2010) resulting in diminished output of mTORC2. The result is a decrease in AKT1 phosphorylation (Ibid). There is also some evidence that mTORC2 may also play a role in the degradation of IRS-1, via phosphorylation, of the Serine residue at position 83, and thus, increasing the stability of the ubiquitin ligase Fbw8 (Kim et al., 2012).

Rapamycin is known to affect AKT1 differently depending upon the concentration of the mTORC1 inhibitor used (Sun et al., 2005; Chen et al., 2010; Saores et al., 2013). This leads to the conclusion that differing concentrations of Rapamycin are capable of affecting these feedback loops in different ways.

Given the qualitative nature of the model, inhibitors behave in an all-or-nothing fashion with complete inhibition being found when they are present and complete activity of the target when absent. As Rapamycin is capable of inhibiting mTORC2 (Sarbassov et al., 2006; Lamming et al., 2012; Schreiber et al., 2015), Rapamycin is viewed in the model as inhibiting this component so as to replicate the effects of exposure to prolonged or high concentrations of Rapamycin.
A by-product of testing the effects of Rapamycin treatment on the activity of ERK1/2 and p38α in RAW 264.7 murine macrophage cells was the investigation of AKT1 phosphorylation. The Human Phospho-MAPK Array data revealed that AKT1 and AKT2 displayed significant inhibition in the presence of Rapamycin compared with the 1% DMSO control. However, only AKT Pan remained significantly inhibited following the Holm-Sidak Correction. To further validate this prediction, immunoblotting was carried out against AKT1 phosphorylation. During this investigation, only very low levels of AKT1 phosphorylated on the Ser-473 site was found, as shown in Figure 3.3.4a. This very low signal resulted in accurate quantification becoming very difficult and raises the possibility that sensitivity of the differing antibodies (between the Human Phospho-MAPK Array and Western Blotting) may have been a contributing factor here. Indeed the low signal may explain the relatively large Standard Error of the Mean values, as shown by the large error bars in Figure 3.3.4a. Additionally, there is some evidence that the activation of AKT1 is governed, at least partly, by the cell cycle (Liu et al., 2014). Given that the RAW 264.7 cells were growing asynchronously, it is possible that the reduction in AKT1 phosphorylation predicted by the model does occur despite not being observed in the Western Blotting work shown in Figure 3.3.4a. Moreover, given the large change in AKT Pan phosphorylation, shown in Figure 3.3.2a it is not unreasonable to conclude that following a one hour treatment with 10µM Rapamycin, Ser-473 phosphorylation is discernibly reduced in a manner that is consistent with Rapamycin may be capable of inhibiting the Rictor subunit of mTORC2. Based solely upon the results of the immunoblotting experiments presented in Section 3.3.4 – Confirmation of the Effect Rapamycin has Upon AKT1 Phosphorylation, it is however, not possible to say if this trend continues or if the situation is reversed.
From a modelling perspective both situations (of Rapamycin bringing about either an increase or decrease in AKT1 phosphorylation) are easily explainable. If AKT1 phosphorylation is to increase, the removal of Rapamycin-mediated mTORC2 inhibition is required. Similarly, if a decrease is observed, the inhibition can remain. Although not shown, both cases have been observed in silico during model simulations.

Section 3.4.3 – Problems Encountered & Troubleshooting

The work outlined in this chapter represented one of the most demanding points of the project in terms of the length of time taken.

Initially, the intensity of bands obtained from Western Blotting was very low. Several reasons were suggested that may account for this: The method of cell lysis being such that samples were degrading during the harvesting process or the cells were not sufficiently stimulated to see background effects. Thirdly, the method of development may not be sensitive enough to detect the relatively low signal of phosphorylated eIF4E. LPS treatment of the RAW 264.7 cells was attempted which resulted in significant activation of multiple signalling pathways, as measured by both Western Blotting and Human Phospho-MAPK Arrays. However, the stimulation of RAW 264.7 cells with LPS was found to be obscuring effects that Rapamycin was having upon the signalling pathways. The results presented in Section 3.3.2 – Evaluation of the Human Phospho-MAPK Arrays show the effects that Rapamycin is capable of exerting upon the cell. However during initial testing with the array kits, far fewer regions of difference were found between the 1% DMSO and 10μM Rapamycin samples when both were exposed to 10ng per ml LPS (from Escherichia coli Serotype O55:B5, TLRgrade™ Enzo (Life Sciences, Inc., Farmingdale, NY (USA))).
It would be interesting to examine the effects of Toll-Like Receptor (TLR) ligands other than LPS. Other TLR ligands, such as Polynosinic-Polycytidylic Acid (Poly(I:C)), are relevant to the cellular response to viral infection. Poly(I:C) is a synthetic double-stranded RNA analogue (Field et al., 1967; Levy et al., 1969). Consequently, the use of this TLR3 ligand would have produced an experimental system in which signalling pathways were activated in a manner that is similar, but not identical, to a virus-infected cell (Li et al., 2005b). Although chemical stimulation of signalling pathways was attempted to overcome the problem of low signal intensity, it is worth noting that at this stage of the project, it was necessary to produce a model that accurately reproduced the cellular regulation of translation and not a system in which viral translation is modelled.

This problem also extended to the evaluation of AKT1 phosphorylation with immunoblotting. The very low signal generated is likely due to very low levels of AKT1 phosphorylation in cells without stimulation. This issue could have been intensified by potentially considerable differences in antibody sensitivities. To address these two issues, stimulating RAW 264.7 cells with either a growth factor or TLR ligand could be of benefit if the experiments were going to be repeated. Furthermore, the role of cell cycle progression in AKT1 phosphorylation should be investigated, possibly through the synchronisation of cell growth.

The second potential problem was addressed by changing the lysis buffer used in harvesting the cells. Several buffers were trialled and all produced a marked increase in the signal found after exposure to chemiluminescent substrate. These methods included lysing the RAW 264.7 cells directly into sample buffer, as described in Simmonds et al. (2009). It was, however, decided that using the Lysis Buffer 6 supplied with the Human Phospho-MAPK
Array kits would give some degree of uniformity to these experiments and thus ensure that the cells were handled in the same way across experiments.

The method of developing the initial blots was limited to the FluorChem®Q apparatus (Alpha Innotech, ProteinSimple Inc., San Jose, CA (USA)). Whilst this system was capable of detecting bands of high intensity, such as loading controls, it was not capable of detecting low intensity bands. Even replacing the SuperSignal West Pico Chemiluminescent Substrate with SuperSignal West Femto Chemiluminescent Substrate (Pierce®, Thermo Fisher Scientific Inc., Rockford, IL (USA)) did not alleviate this problem and resulted in even higher levels of background signal.

It was possible to overcome these problems with a combination of a traditional radiographic film based development process and switching to a new lysis buffer. This suggests that a combination of signal degradation and insufficient sensitivity of the development process were to blame for these initial setbacks. This being said, one point is still unresolved with regards to the method used to quantify radiographic film using pixel counting (as used in Section 3.3.2 – Evaluation of the Human Phospho-MAPK Arrays, Section 3.3.3 – Alterations to eIF4E Phosphorylation Levels and Section 3.3.4 – Confirmation of the Effect Rapamycin has Upon AKT1 Phosphorylation): the fact that exposure of radiographic film to the fluorescent signal of chemiluminescence does not produce a linear response (Gassmann et al., 2009). Consequently, it may not be the most accurate method of quantification, although this method was recommended for use with the Human Phospho-MAPK Arrays by the manufacturer.

One question is still unresolved with regards to Rapamycin leading to an increase in eIF4E phosphorylation. This problem lies with the identification of the MAPK-Interacting Kinase...
responsible for this increase. Attempts have been made by other members of the research group, as part of the Royall et al. (2015) study, looking at using small interfering RNA (siRNA) to knock-down MNK1 and MNK2 in RAW 264.7 cells. However, this work was not taken forward due to a lack of suitable internal controls and in some cases, knock-down of MNK1 with siRNA targeted against MNK2. Similarly, there is a lack of specific MNK1 or MNK2 antibody, thus preventing direct visualisation of specific MNK phosphorylation during Rapamycin treatment. The potential benefit of an additional approach, Clustered, Regularly Interspaced, Short Palindromic Repeats (CRISPR)/CRISPR-Associated Protein (Cas) (Cong et al., 2013; Mali et al., 2013) is discussed in Section 3.4.4 – Potential Future Work.

Section 3.4.4 – Potential Future Work

eIF4E phosphorylation is now being viewed as playing a part in mediating changes in gene expression. mRNA transcripts involved in cellular stress appear to require the phosphorylation of the 5’ cap-binding protein in order for efficient translation initiation (Herdy et al., 2012). Given this and that mTORC1 plays a vital role in regulating cell growth, it would be of interest to determine if eIF4E phosphorylation occurring during mTORC1 inhibition is accompanied by global changes to the transcriptome of the cell. Moreover, establishing if other MNK substrates, such as hnRNP A1, exhibit a Rapamycin-induced increase in phosphorylation may help to demonstrate if mTORC1 inhibition leads to a global stress response.

As AKT1 phosphorylation has been shown to be negatively regulated in a manner implying mTORC2 inhibition following a one hour treatment with 10μM Rapamycin, establishing the sensitivity of mTORC2 to Rapamycin in RAW 264.7 cells would be of importance. Schreiber and colleagues (2015) have shown that a high ratio of relative FKBP12 to FKBP51 expression
is an important determinant in the degree of mTORC2 inhibition in response to Rapamycin treatment. Determining this ratio in RAW 264.7 cells could, potentially, be of interest.

Determining if MNK1 or MNK2 is responsible for mediating the Rapamycin-induced increase in eIF4E phosphorylation, although technically demanding, would be important. If this area were to be pursued, the use of the CRISPR/Cas technique could be considered over the development of a transgenic mouse. Given the ease and efficiency of CRISPR/Cas, the technique could be of benefit and has proven useful in knocking out genes on a large scale in mammalian cells (Shalem et al., 2013; Walsh & Hochedlinger, 2013; Auer et al., 2014; Wang et al., 2014). However, this approach does suffer from the limitation that it requires a Protospacer Adjacent Motif (PAM), to which the Cas protein can bind (Ibid), surrounding, what in this case would be the MNK1 or MNK2 genes. A lack of this PAM site, or insufficient sequence homology between the Cas protein and this sequence, would result in reduced efficiency (Fu et al., 2013; Hsu et al., 2013 (As cited by Walsh & Hochedlinger, 2013)). Although this approach has been successfully used in vivo to create knockout mice (Wang et al., 2013; Fujii et al., 2013), to date no such studies have used this technique on RAW 264.7 cells. Walsh and Hochedlinger (2013) noted that additional research, such as that by Hou et al. (2013), in identifying novel Cas proteins may mean that this approach will become more applicable here.

The use of transgenic mice is costly in the terms of financial implications, time and the high failure (Hall et al., 2009). There are also ethical considerations. The number of animal experiments required in any project should be minimised. Furthermore, experimental validation of predictions in this project is aimed at improving a model which was to be used
to mimic viral infection, as such using primary cells taken from a mouse would, most probably, behave in a different manner to the transformed RAW 264.7 cells.

Section 3.4.5 – Concluding Remarks

Despite experimental challenges, the effect of Rapamycin on eIF4E phosphorylation, predicted by the model, has been shown. Although there were also problems regarding the effect this macrolide has upon AKT1 phosphorylation, there is sufficient evidence to conclude that a considerable reduction does occur following one hour of treatment. From this perspective, given the qualitative nature of the model, this prediction has also been validated.

By using more of the data from the array in the comparison to the binary model behaviour, it has been shown that the model is in the majority of cases behaving in a manner reminiscent of the model system, the murine macrophage RAW 264.7 cell. In doing so the validity of the model has increased and, although the comparison was of a relatively small number of data points, without further refinement or adjustment, the final model has produced a positive MCC value.

The experimental validation of these two predictions was a major step forward in this project. Although these two predictions are well documented in the literature, in experimentally corroborating these hypothesised behaviours, it has been shown that the model is proficient in predicting complex and non-trivial patterns of behaviour. In keeping with this, at this stage the rationale for conducting these experiments was not to necessarily provide solid mechanistic data regarding the link between the phosphorylation of AKT1 and eIF4E and Rapamycin-mediated inhibition of mTORC1, but rather to provide clear evidence that the behaviour of the model is predictable when certain, specific perturbations are
applied. In doing so, it has been shown that the model was ready to be used in novel ways, such as the investigation of the interactions between a virus and the host signalling network regulating translation initiation.
Chapter 4 – Modelling the Effects of Viral Infection
Section 4.1 – Introduction

Viruses are pathogens which commandeer cellular machinery to prioritise the synthesis of viral proteins. The validated model presented in this work gives a unique opportunity to study these host-pathogen interactions. Within Chapter 2 – Model Creation & Refining and Chapter 3 – Experimental Validation of Model Predictions, the model of translational regulation was created and experimentally validated. The predictive power of the model was raised significantly, to MCC = 0.4558. This improved predictive power and the experimental validation of two predicted behaviours meant than the model was behaving in a reliable way. In order to demonstrate the value of model, in a broad context, the incorporation of viral effects into the model was carried out.

Controlling and combatting viral infection is a major challenge facing modern science. Despite considerable advances in the field of virology, there is a limited spectrum of viable treatments for viral infections. Moreover, some of these diseases are causes of considerable morbidity and mortality and represent a major burden on the National Health Service and other healthcare systems and public health services worldwide. With the use of systems biology techniques to study viral infection, it is hoped that new cellular targets for treating viral infection will emerge.

Whilst the main focus of this project has been to create a model of mammalian translation, the possibility of using this model to mimic the effects viral infection has on the translational machinery and upstream regulatory signalling pathways was investigated. Three viruses have been selected on the basis that they perturb different cellular pathways and have different effects on the translational machinery. JUNV, ANDV and Murine Norovirus (MNV) are all considered important human pathogens or models of such pathogens. Although two
of these viruses may not infect RAW 264.7 cells, they have been included here as they produce different effects on the cellular signalling network and manipulate the host translational machinery in ways that MNV does not.

Section 4.2 – Model Changes & Simulation Methods

For each virus it was necessary to run simulations with two models. The models were identical in every respect except that one was used to represent an ‘infected’ system and one as the ‘mock’, uninfected control. QSSPN simulations were carried out on the ‘infected’ or ‘mock’ models. The ‘mock’ model was initially tested to ensure that the simulation times were adjusted to result in the frequency of each effector being equal to around 0.5. Binomial Confidence Intervals, calculated in the software R, were used to identify regions of significant difference between the ‘infected’ or ‘mock’ models. The final model presented at the end of Chapter 2 – Model Creation & Refining and experimentally validated in Chapter 3 – Experimental Validation of Model Predictions was used here. More model effectors, than used in previous chapters, were monitored as it was hoped that new effects or, in the case of MNV, testable predictions could be found.

Section 4.2.1 – Junin Virus Model

To mimic the large scale changes to the host translation factors, two features of Junin Virus infection were considered. Firstly, JUNV contains the same conserved amino acids believed to be found in the active site of the L endonuclease that are found in other Arenaviruses (Morin et al., 2010). Therefore, it is reasonable to assume that a cap-snatching mechanism is employed by JUNV. Secondly, like many of the other New World Arenaviruses, JUNV replication occurs with the Nucleoprotein acting as the cap-binding protein in place of eIF4E (Linero et al., 2013). The ability of the Arenaviridae to actively use eIF4E is controversial. As
noted by Knopp and colleagues (2015), the Old World Arenaviruses, such as Lassa Fever Virus, may differ from the New World Arenaviruses in the need for this cellular factor (Volpon et al., 2008; Volpon et al., 2010; Kranzusch & Whelan, 2011). Furthermore, Arenaviruses contain RNA that is not polyadenylated (Emonet et al., 2009). This would suggest that PABP is redundant in JUNV infection.

To date JUNV infection is known to require the activation of two fundamental signalling pathways. Although displaying an oscillatory nature, JUNV requires the activation of ERK1/2 (Rodríguez et al., 2014). Of most significance here is the fact that at 24 hours post infection the level of ERK1/2 activation is significantly increased. This period of activation co-insides with the finding of Linero and colleagues (2013) that viral infection and protein synthesis were not affected by a knock-down of eIF4E. Rodríguez and co-workers (2014) posit that the N protein mediates the increase in ERK1/2 activation. Given these events occur at the same time point it is possible to model these events simultaneously. The second pathway upon which JUNV infection depends is the PI3K pathway (Linero & Scolaro, 2009). The significance of this pathway at 24 hours post infection is not clear as there is an absence of data at this time point regarding the activation of AKT1 or another PI3K substrate. This being said the presence of a PI3K inhibitor after 24 hours post infection did markedly reduce viral replication (Ibid). No clear data regarding other signalling pathways is known in respect of this virus. Boswick et al. (2007) have shown that the related Pichinde Virus induces p38 activation at earlier time points but no work has been carried out with JUNV.

Given that the model is entirely qualitative and provides no temporal information, both the ERK1/2 and PI3K pathways were activated in the ‘infected’ model. The use of both cap snatching and substitution of eIF4E were also applied in the model.
The effects that JUNV is capable of inducing, in the cellular signalling pathways and translational machinery, were incorporated into the final version of the Petri Net model. Building on this model, 17 new molecular species were included in the model, to account for different translational intermediates and viral proteins that effect signalling pathway components. These new species are involved in 27 new molecular interactions, which include viral translation and the effects of viral components on cellular signalling pathways.

Section 4.2.2 – Murine Norovirus Model

The role of the various initiation factors in MNV infection is complex and not fully understood. The VP$_g$ protein of MNV is known to function in place of the 5’ cap structure (Chaudhry et al., 2006). However, the interactions between this protein and cellular initiation factors are still not fully understood. Whilst an interaction between VP$_g$ and eIF4E has been reported (Ibid) the significance is not clear. Chaudhry and co-workers (2006) found that a reduction of cellular eIF4E or the overexpression of 4E-BP1 did not preclude viral protein synthesis. It has since been proposed that the interaction between the viral protein and eIF4E only functions to help circumvent a change in gene expression to bring about an immune response (Chung et al., 2014). In terms of an integral interaction, Chung and co-workers (2014) demonstrated that VP$_g$ must interact with eIF4G. Indeed this finding is corroborated by the view that eIF4E depletion does not affect viral translation (Chaudhry et al., 2006).

Whilst eIF4E is not required for the translation of MNV proteins, there is a dependence on the phosphorylation of this initiation factor (Royall et al., 2015). This work established that the increase in eIF4E phosphorylation serves to bring about a change in gene expression including an increase in the expression of an inhibitor of Nuclear Factor κB (Ibid). Whilst
both ERK1/2 and p38 phosphorylation were shown to be increased during MNV infection, only the axis of p38α-MNK1 was required for the increase in eIF4E phosphorylation (Royall et al., 2015). Additionally, unpublished data (generated by N. Doyle and E. Royall (University of Surrey)) indicates that MNV-1 infection produces a marked increase in the phosphorylation of AKT1 on the residue Ser-473. This Human Phospho-MAPK array data also demonstrated that the signalling axes of mTORC1-S6K1 and JNK1 were unaffected by MNV-1 infection.

The MNV model, building on the final model, included 13 new molecular species involved in 20 novel molecular interactions. These interactions were involved in viral translation and the interactions between viral factors, and the cellular translational machinery and signalling factors. These interactions were documented in existing literature or taken from experimental data generated in-house.

Section 4.2.3 – Andes Virus Model

Whilst not explicitly shown to be the case for ANDV, the Hantaviruses are widely believed to use cellular 5’ methylguanosine caps to initiate viral translation (Mir et al., 2010; Cheng & Mir, 2012). However, the work of Heinemann and colleagues (2013) has suggested that selective host mRNA degradation occurs in ANDV infection and is mediated by the L protein. When taken together these works suggest that the RdRp L protein may function to selectively degrade cellular mRNA in ANDV infection whilst the N protein may then sequester the cleaved cap fragment to prevent further cell-mediated degradation (Cheng & Mir, 2012; Heinemann et al., 2013). Furthermore, the N-terminal domain of the L protein, which is thought to mediate the cleavage of mRNA, is well conserved amongst the Hantaviruses (Reguera et al., 2010).
There is additional evidence that suggests translation of at least one ANDV protein, the NSs protein is encoded by a Small segment of viral RNA (Vera-Otarola et al., 2010) that does not require the use of PABP. Rather, the untranslated region within the 3’ end of the Small transcript is used (Vera-Otarola et al., 2012). These authors suggest that an unknown cellular protein may mediate Small RNA translation (Ibid). It would be of interest to compare the translation of histone mRNA to the translation of the Small genomic segment. The 3’ UTR of histone mRNAs contains a stem-loop structure to which SLBP must interact for translation to occur (Sànchez & Marzluff, 2002). As noted by Vera-Otarola and co-workers (2010) no base pairing between the 5’ and 3’ structures could occur. It raises the possibility that viral translation could occur through a mechanism that involves circularisation of viral mRNA by SLBP interacting with the 40S subunit, possibly through eIF3 (as is the case for histone mRNA (Ling et al., 2002)) and potentially by the N protein replacing the eIF4F complex (Mir & Panganiban, 2008).

In addition to the effects that ANDV imposes on the translational machinery, infection with this Hantavirus has the potential to exert large effects on the cellular signalling network. ANDV is capable of inducing mTORC1 activation through the autophagic degradation of the viral protein Gn (formed during cleavage of the GPC protein (Löber et al., 2001), which serves to maintain the pool of amino acids within the lysosomes (McNulty et al., 2013). Whilst mTORC1 activation has also been demonstrated by Gavrilovskaya et al. (2013), there is some disagreement over the level of S6K1 activity. Gavrilovskaya and co-workers (2013) showed that ANDV infection, in conjunction with hypoxia, resulted in increased S6K1 phosphorylation whereas no such increase in either this or 4E-BP1 phosphorylation was noted by McNulty et al. (2013). An obvious difference in these works is the induction of
hypoxia, however no S6K1 phosphorylation was found in hypoxic cells alone (Gavrilovskaya et al., 2013). These works do however suggest that the result of mTORC1 activation by ANDV is to bring about induction of a variety of mTORC1-dependent pathways such as Hypoxia Inducible Factor 1α (HIF-1α).

Additionally, no change in the activation of the inflammatory regulator Nuclear Factor-κB was noted in cells infected with ANDV (Khaiboullina et al., 2013). ANDV infection was capable of increasing Nuclear Factor-κB activation in cells pre-treated with Tumour Necrosis Factor α (Ibid).

Building on the final version of the Petri Net model of mammalian translational regulation, the effects of ANDV infection were incorporated. To this end, 17 new molecular species were added to the model to account for various viral factors and translational intermediate stages, where viral factors or viral RNA were incorporated. To model viral protein synthesis and the effects ANDV has upon the cellular signalling network, 22 additional molecular interactions were added to the model.

Section 4.2.4 – Experimental Validation of the Murine Norovirus Model

Upon examination of the behaviour of the MNV model, the level of phosphorylated Cyclic AMP-Responsive Element-Binding Protein (CREB) was found to be increased during MNV infection. Using Human Phospho-MAPK Arrays, it was possible to test this behaviour.

The experimental work presented in this chapter was conducted by N. Doyle and E. Royall using the method detailed in Royall et al. (2015). RAW 264.7 cells were cultured as per the method given in Royall et al. (2015). These cells were then infected with MNV strain CW1 at a Multiplicity of Infection (MOI) of ten 50% Tissue Culture Infectious Dose units and were
incubated at 37°C with 5% Carbon Dioxide for either two or 12 hours. Uninfected control samples for each time point were also taken. Human Phospho-MAPK Arrays were carried out using 300μg of protein, as determined by a BCA Assay, from each lysate. These Human Phospho-MAPK Arrays were carried out according the manufacturers’ instruction and were developed on to radiographic film. Quantification of each phosphoprotein, also carried out by N. Doyle, was conducted in ImageJ. As detailed in Royall et al. (2015), the values for CREB phosphorylation were expressed as relative pixel density. Prior to this, these values were expressed as relative values of the array control spots CREB. The statistical analysis of the data was carried out using two-way ANOVA with Tukey’s post-hoc test in GraphPad Prism 6.

Section 4.3 – Results

Section 4.3.1 – Simulation of the Junin Virus Models

The effects of JUNV infection were incorporated into the final version of the binary model. QSSPN simulations and Reachfq analyses were performed and the results, shown in Figure 4.3.1a, reveal multiple differences between the ‘infected’ and ‘mock’ models. As expected, both ERK1/2 and PI3K activity was increased in the ‘infected’ model. Similarly, the production of viral protein was limited to the ‘infected’ model. A significant reduction in host protein synthesis was also observed (as seen in the reduced frequency of trajectories resulting in protein synthesis). The activity of many substrates for either ERK1/2, such as MAPK Phosphatase 7 (MKP-7) or RSK1, or PI3K, such as AKT1 or PDK1, also exhibited increases.

The most unexpected results concerned the nature of feedback mechanisms between the ERK1/2 and JNK1 signalling pathways. Viral infection appeared to negatively affect eIF4E
phosphorylation through both MNK1 and MNK2 activity. mTORC1 signalling also appears to be, for the most part, increased. Interestingly, transcription factors, such as HIF-1α and Elk-1, were also up-regulated and suggest that JUNV infection would promote large scale gene expression changes. The HIF-1α expression also suggests that the metabolic network, as discussed in Section 5.1.1 – Linking mTORC1 to the Regulation of Metabolism, of the infected cell would exhibit considerable metabolic changes. It would therefore, be of considerable future interest to evaluate the nature and degree of impact on the metabolic network by simulations of the integrated signalling network model and GSMN. Chapter 5 – Integration of the Translational Control Model with a Genome Scale Metabolic Network of a RAW 264.7 Cell describes the preliminary results obtained in this direction.
<table>
<thead>
<tr>
<th>Protein</th>
<th>Mock</th>
<th>Upper</th>
<th>Lower</th>
<th>Infected</th>
<th>Upper</th>
<th>Lower</th>
</tr>
</thead>
<tbody>
<tr>
<td>ppMEK1</td>
<td>0.521</td>
<td>0.572</td>
<td>0.469</td>
<td>0.496</td>
<td>0.548</td>
<td>0.444</td>
</tr>
<tr>
<td>ppMEK2</td>
<td>0.507</td>
<td>0.559</td>
<td>0.455</td>
<td>0.455</td>
<td>0.507</td>
<td>0.404</td>
</tr>
<tr>
<td>ppERK1</td>
<td>0.521</td>
<td>0.572</td>
<td>0.469</td>
<td>0.992</td>
<td>0.997</td>
<td>0.976</td>
</tr>
<tr>
<td>ppERK2</td>
<td>0.534</td>
<td>0.585</td>
<td>0.492</td>
<td>0.989</td>
<td>0.996</td>
<td>0.972</td>
</tr>
<tr>
<td>pRSK1</td>
<td>0.430</td>
<td>0.481</td>
<td>0.380</td>
<td>0.726</td>
<td>0.770</td>
<td>0.677</td>
</tr>
<tr>
<td>rpS6_Pi</td>
<td>0.550</td>
<td>0.601</td>
<td>0.498</td>
<td>0.718</td>
<td>0.762</td>
<td>0.669</td>
</tr>
<tr>
<td>pMNK1</td>
<td>0.540</td>
<td>0.591</td>
<td>0.488</td>
<td>0.147</td>
<td>0.188</td>
<td>0.114</td>
</tr>
<tr>
<td>pMNK2</td>
<td>0.448</td>
<td>0.500</td>
<td>0.397</td>
<td>0.149</td>
<td>0.190</td>
<td>0.115</td>
</tr>
<tr>
<td>elf4E_Pi</td>
<td>0.480</td>
<td>0.532</td>
<td>0.429</td>
<td>0.134</td>
<td>0.173</td>
<td>0.102</td>
</tr>
<tr>
<td>pJNK1</td>
<td>0.460</td>
<td>0.512</td>
<td>0.409</td>
<td>0.138</td>
<td>0.178</td>
<td>0.106</td>
</tr>
<tr>
<td>cJun_Pi</td>
<td>0.552</td>
<td>0.603</td>
<td>0.500</td>
<td>0.119</td>
<td>0.157</td>
<td>0.089</td>
</tr>
<tr>
<td>ppp38a</td>
<td>0.541</td>
<td>0.592</td>
<td>0.489</td>
<td>0.552</td>
<td>0.603</td>
<td>0.500</td>
</tr>
<tr>
<td>pERK5</td>
<td>0.448</td>
<td>0.500</td>
<td>0.397</td>
<td>0.259</td>
<td>0.307</td>
<td>0.216</td>
</tr>
<tr>
<td>PI3K_Local</td>
<td>0.450</td>
<td>0.502</td>
<td>0.399</td>
<td>0.944</td>
<td>0.964</td>
<td>0.915</td>
</tr>
<tr>
<td>PIP3</td>
<td>0.508</td>
<td>0.560</td>
<td>0.456</td>
<td>0.912</td>
<td>0.937</td>
<td>0.878</td>
</tr>
<tr>
<td>PDK1_Local</td>
<td>0.553</td>
<td>0.604</td>
<td>0.501</td>
<td>0.910</td>
<td>0.935</td>
<td>0.876</td>
</tr>
<tr>
<td>pPKCz</td>
<td>0.490</td>
<td>0.542</td>
<td>0.438</td>
<td>0.768</td>
<td>0.809</td>
<td>0.721</td>
</tr>
<tr>
<td>ppAKT1</td>
<td>0.562</td>
<td>0.613</td>
<td>0.510</td>
<td>0.849</td>
<td>0.883</td>
<td>0.808</td>
</tr>
<tr>
<td>ppS6K1</td>
<td>0.150</td>
<td>0.191</td>
<td>0.117</td>
<td>0.008</td>
<td>0.024</td>
<td>0.003</td>
</tr>
<tr>
<td>p4EBP1</td>
<td>0.523</td>
<td>0.574</td>
<td>0.471</td>
<td>0.762</td>
<td>0.803</td>
<td>0.715</td>
</tr>
<tr>
<td>mTORC1_Local_Pi</td>
<td>0.482</td>
<td>0.534</td>
<td>0.430</td>
<td>0.731</td>
<td>0.774</td>
<td>0.683</td>
</tr>
<tr>
<td>pTSC12</td>
<td>0.489</td>
<td>0.541</td>
<td>0.437</td>
<td>0.895</td>
<td>0.923</td>
<td>0.859</td>
</tr>
<tr>
<td>elf4B_Pi</td>
<td>0.511</td>
<td>0.563</td>
<td>0.459</td>
<td>0.668</td>
<td>0.715</td>
<td>0.617</td>
</tr>
<tr>
<td>Protein_Synthesis</td>
<td>0.515</td>
<td>0.567</td>
<td>0.463</td>
<td>0.324</td>
<td>0.374</td>
<td>0.277</td>
</tr>
<tr>
<td>pCREB</td>
<td>0.531</td>
<td>0.582</td>
<td>0.479</td>
<td>0.758</td>
<td>0.800</td>
<td>0.711</td>
</tr>
<tr>
<td>pMSK2</td>
<td>0.586</td>
<td>0.636</td>
<td>0.534</td>
<td>0.584</td>
<td>0.634</td>
<td>0.532</td>
</tr>
<tr>
<td>pMKK3</td>
<td>0.542</td>
<td>0.593</td>
<td>0.490</td>
<td>0.512</td>
<td>0.564</td>
<td>0.460</td>
</tr>
<tr>
<td>pMKK6</td>
<td>0.534</td>
<td>0.585</td>
<td>0.482</td>
<td>0.524</td>
<td>0.575</td>
<td>0.472</td>
</tr>
<tr>
<td>pGSK3</td>
<td>0.575</td>
<td>0.625</td>
<td>0.523</td>
<td>0.625</td>
<td>0.674</td>
<td>0.574</td>
</tr>
<tr>
<td>pMSK1</td>
<td>0.538</td>
<td>0.589</td>
<td>0.486</td>
<td>0.886</td>
<td>0.914</td>
<td>0.848</td>
</tr>
<tr>
<td>pMK2</td>
<td>0.533</td>
<td>0.584</td>
<td>0.481</td>
<td>0.510</td>
<td>0.562</td>
<td>0.458</td>
</tr>
<tr>
<td>pTTP</td>
<td>0.551</td>
<td>0.602</td>
<td>0.499</td>
<td>0.599</td>
<td>0.649</td>
<td>0.547</td>
</tr>
<tr>
<td>HIF1a</td>
<td>0.535</td>
<td>0.586</td>
<td>0.483</td>
<td>0.716</td>
<td>0.760</td>
<td>0.667</td>
</tr>
<tr>
<td>psTAT3</td>
<td>0.481</td>
<td>0.533</td>
<td>0.429</td>
<td>0.697</td>
<td>0.742</td>
<td>0.647</td>
</tr>
<tr>
<td>pElk1</td>
<td>0.529</td>
<td>0.580</td>
<td>0.477</td>
<td>0.895</td>
<td>0.923</td>
<td>0.859</td>
</tr>
<tr>
<td>pATF2</td>
<td>0.534</td>
<td>0.585</td>
<td>0.482</td>
<td>0.496</td>
<td>0.548</td>
<td>0.444</td>
</tr>
<tr>
<td>pTpl2</td>
<td>0.533</td>
<td>0.584</td>
<td>0.481</td>
<td>0.543</td>
<td>0.594</td>
<td>0.491</td>
</tr>
<tr>
<td>ppPKCa</td>
<td>0.537</td>
<td>0.588</td>
<td>0.485</td>
<td>0.540</td>
<td>0.591</td>
<td>0.488</td>
</tr>
<tr>
<td>pMKP7</td>
<td>0.486</td>
<td>0.538</td>
<td>0.434</td>
<td>0.597</td>
<td>0.999</td>
<td>0.984</td>
</tr>
<tr>
<td>elf4G_Pi</td>
<td>0.574</td>
<td>0.624</td>
<td>0.522</td>
<td>0.437</td>
<td>0.489</td>
<td>0.386</td>
</tr>
<tr>
<td>Viral_Protein</td>
<td>0.000</td>
<td>0.011</td>
<td>0.000</td>
<td>0.335</td>
<td>0.386</td>
<td>0.288</td>
</tr>
</tbody>
</table>

Figure 4.3.1a – Modelling the Effects of JUNV Infection. Comparison of the impact JUNV infection has upon the cellular signalling network and the translation initiation pathway. Significant differences between the models are found by comparing the 99.99% Binomial Confidence Intervals, as calculated in R, of the QSPPN simulation results for each effector. Significant differences are shown in purple and indicate no overlap in the confidence intervals.
Section 4.3.2 – Simulation of the Murine Norovirus Models

MNV was chosen as it is permissive in RAW 264.7 cells, the experimental system used in this project. The effects of MNV were incorporated into the final version of the binary model and simulations were run with QSSPN. Reachfq was used to calculate the frequency of 1,000 trajectories which resulted in the activation or production of each effector. 99.99% Binomial Confidence Intervals were calculated for each effector, as shown in Figure 4.3.2a. Significant differences were identified by no overlap being present in these 99.99% confidence intervals. The model incorporating MNV infection is given in Supplementary Information E4 (given as both a Snoopy model file (E4a) and an SBML file (.xml)) (E4b).

In order to accurately model MNV infection, the activity of several pathways was altered. The activity of the ERK1/2 and p38 MAPK pathways was increased in keeping with experimental findings. Similarly, the activation of AKT1 was also promoted. This Human Phospho-MAPK Array data also confirmed that neither JNK1 nor mTORC1 activity was increased. One of the main findings of this model was the raised eIF4E phosphorylation. The production of viral protein was only found in the ‘infected’ model. The ‘infected’ model also displayed no significant reduction in cellular protein synthesis. This model prediction is somewhat in agreement, the work of Royall et al. (2015). When using polysome profiling, this work noted that the translational profile of the cell changes in response to infection. This being said, the level of a housekeeping gene, Glyceraldehyde 3-Phosphate Dehydrogenase, was not affected by MNV infection (Ibid) and whilst not conclusive proof, it is indicative of no global reduction in host cell protein synthesis.

These simulations, the results of which are outlined in Figure 4.3.2a, have predicted numerous molecular targets that have not, to date, previously been identified as being
affected by MNV infection. Several transcription factors, namely CREB, Elk-1 and ATF-2, downstream of ERK1/2 and p38 were found to exhibit increased phosphorylation. An increase in the activation of various signalling components was also noted. The majority of these components (MNK1, MNK2, MSK1, MSK2, MK2 and MKP-7) are, again, downstream of either ERK1/2 or p38. This model also suggests that AU-Rich Element-containing mRNA is stabilised as a result of increased phosphorylation of TTP. Experimental work to validate the increase in CREB phosphorylation is detailed in Section 4.3.2.1 – Experimental Validation of the Hypotheses Resulting from Simulations of the Murine Norovirus Model.

The perturbations to the model also led to the level of phosphorylation of several initiation factors to be altered. eIF4B phosphorylation is strikingly increased. This affect can be linked to the increase in the ERK1/2-mediated phosphorylation of RSK1. Despite no change in PKCα phosphorylation, eIF4G phosphorylation suffers a slight, but significant, reduction. As already noted, eIF4E phosphorylation is increased. The work of Royall and co-workers (2015) determined that MNK1, via p38, is responsible for this increase. Whilst this picture is somewhat confirmed by the behaviour of the model, this work also notes an increase in MNK2 phosphorylation.
<table>
<thead>
<tr>
<th>Protein</th>
<th>Mock</th>
<th>Upper</th>
<th>Lower</th>
<th>Infected</th>
<th>Upper</th>
<th>Lower</th>
</tr>
</thead>
<tbody>
<tr>
<td>ppMEK1</td>
<td>0.538</td>
<td>0.589</td>
<td>0.486</td>
<td>0.504</td>
<td>0.556</td>
<td>0.452</td>
</tr>
<tr>
<td>ppMEK2</td>
<td>0.510</td>
<td>0.562</td>
<td>0.458</td>
<td>0.521</td>
<td>0.572</td>
<td>0.469</td>
</tr>
<tr>
<td>ppERK1</td>
<td>0.534</td>
<td>0.585</td>
<td>0.482</td>
<td>0.594</td>
<td>0.998</td>
<td>0.979</td>
</tr>
<tr>
<td>ppERK2</td>
<td>0.540</td>
<td>0.591</td>
<td>0.488</td>
<td>0.992</td>
<td>0.997</td>
<td>0.976</td>
</tr>
<tr>
<td>pRSK1</td>
<td>0.462</td>
<td>0.514</td>
<td>0.411</td>
<td>0.832</td>
<td>0.867</td>
<td>0.790</td>
</tr>
<tr>
<td>rpS6_Pi</td>
<td>0.471</td>
<td>0.523</td>
<td>0.420</td>
<td>0.744</td>
<td>0.787</td>
<td>0.696</td>
</tr>
<tr>
<td>pMNK1</td>
<td>0.00</td>
<td>0.011</td>
<td>0.000</td>
<td>0.931</td>
<td>0.953</td>
<td>0.900</td>
</tr>
<tr>
<td>pMNK2</td>
<td>0.515</td>
<td>0.567</td>
<td>0.463</td>
<td>0.775</td>
<td>0.815</td>
<td>0.729</td>
</tr>
<tr>
<td>eIf4E_Pi</td>
<td>0.511</td>
<td>0.563</td>
<td>0.459</td>
<td>0.827</td>
<td>0.863</td>
<td>0.784</td>
</tr>
<tr>
<td>pJNK1</td>
<td>0.468</td>
<td>0.520</td>
<td>0.417</td>
<td>0.452</td>
<td>0.504</td>
<td>0.401</td>
</tr>
<tr>
<td>cJun_Pi</td>
<td>0.462</td>
<td>0.514</td>
<td>0.411</td>
<td>0.450</td>
<td>0.502</td>
<td>0.399</td>
</tr>
<tr>
<td>ppp38a</td>
<td>0.444</td>
<td>0.496</td>
<td>0.393</td>
<td>0.882</td>
<td>0.912</td>
<td>0.844</td>
</tr>
<tr>
<td>pERK5</td>
<td>0.458</td>
<td>0.510</td>
<td>0.407</td>
<td>0.284</td>
<td>0.333</td>
<td>0.240</td>
</tr>
<tr>
<td>PI3K_Local</td>
<td>0.448</td>
<td>0.500</td>
<td>0.397</td>
<td>0.456</td>
<td>0.508</td>
<td>0.405</td>
</tr>
<tr>
<td>PIP3</td>
<td>0.441</td>
<td>0.493</td>
<td>0.390</td>
<td>0.484</td>
<td>0.536</td>
<td>0.432</td>
</tr>
<tr>
<td>PDK1_Local</td>
<td>0.551</td>
<td>0.602</td>
<td>0.499</td>
<td>0.569</td>
<td>0.620</td>
<td>0.517</td>
</tr>
<tr>
<td>pPKc</td>
<td>0.460</td>
<td>0.512</td>
<td>0.409</td>
<td>0.462</td>
<td>0.514</td>
<td>0.411</td>
</tr>
<tr>
<td>ppAKT1</td>
<td>0.461</td>
<td>0.513</td>
<td>0.410</td>
<td>0.616</td>
<td>0.665</td>
<td>0.564</td>
</tr>
<tr>
<td>ppS6K1</td>
<td>0.127</td>
<td>0.166</td>
<td>0.096</td>
<td>0.078</td>
<td>0.111</td>
<td>0.054</td>
</tr>
<tr>
<td>p4EBP1</td>
<td>0.479</td>
<td>0.531</td>
<td>0.428</td>
<td>0.435</td>
<td>0.487</td>
<td>0.384</td>
</tr>
<tr>
<td>mTORC1_Local_Pi</td>
<td>0.479</td>
<td>0.531</td>
<td>0.428</td>
<td>0.447</td>
<td>0.499</td>
<td>0.396</td>
</tr>
<tr>
<td>pTSC12</td>
<td>0.507</td>
<td>0.559</td>
<td>0.455</td>
<td>0.730</td>
<td>0.774</td>
<td>0.681</td>
</tr>
<tr>
<td>eIf4B_Pi</td>
<td>0.572</td>
<td>0.622</td>
<td>0.520</td>
<td>0.814</td>
<td>0.851</td>
<td>0.770</td>
</tr>
<tr>
<td>Protein_Synthesis</td>
<td>0.386</td>
<td>0.438</td>
<td>0.337</td>
<td>0.293</td>
<td>0.342</td>
<td>0.248</td>
</tr>
<tr>
<td>pCREB</td>
<td>0.507</td>
<td>0.559</td>
<td>0.455</td>
<td>0.915</td>
<td>0.940</td>
<td>0.881</td>
</tr>
<tr>
<td>pMSK2</td>
<td>0.480</td>
<td>0.532</td>
<td>0.429</td>
<td>0.803</td>
<td>0.841</td>
<td>0.758</td>
</tr>
<tr>
<td>pMKK3</td>
<td>0.461</td>
<td>0.513</td>
<td>0.410</td>
<td>0.437</td>
<td>0.489</td>
<td>0.386</td>
</tr>
<tr>
<td>pMKK6</td>
<td>0.442</td>
<td>0.494</td>
<td>0.391</td>
<td>0.435</td>
<td>0.487</td>
<td>0.384</td>
</tr>
<tr>
<td>pGSK3</td>
<td>0.444</td>
<td>0.496</td>
<td>0.393</td>
<td>0.462</td>
<td>0.514</td>
<td>0.411</td>
</tr>
<tr>
<td>pMK1</td>
<td>0.464</td>
<td>0.516</td>
<td>0.413</td>
<td>0.946</td>
<td>0.965</td>
<td>0.917</td>
</tr>
<tr>
<td>pMK2</td>
<td>0.577</td>
<td>0.627</td>
<td>0.525</td>
<td>0.859</td>
<td>0.891</td>
<td>0.819</td>
</tr>
<tr>
<td>pTTP</td>
<td>0.542</td>
<td>0.593</td>
<td>0.490</td>
<td>0.705</td>
<td>0.750</td>
<td>0.656</td>
</tr>
<tr>
<td>HIF1a</td>
<td>0.497</td>
<td>0.549</td>
<td>0.445</td>
<td>0.475</td>
<td>0.527</td>
<td>0.424</td>
</tr>
<tr>
<td>pSTAT3</td>
<td>0.477</td>
<td>0.529</td>
<td>0.426</td>
<td>0.447</td>
<td>0.499</td>
<td>0.396</td>
</tr>
<tr>
<td>pElk1</td>
<td>0.466</td>
<td>0.518</td>
<td>0.415</td>
<td>0.952</td>
<td>0.970</td>
<td>0.924</td>
</tr>
<tr>
<td>pATF2</td>
<td>0.472</td>
<td>0.524</td>
<td>0.421</td>
<td>0.786</td>
<td>0.825</td>
<td>0.740</td>
</tr>
<tr>
<td>pTpi2</td>
<td>0.526</td>
<td>0.577</td>
<td>0.474</td>
<td>0.510</td>
<td>0.562</td>
<td>0.458</td>
</tr>
<tr>
<td>ppPKca</td>
<td>0.557</td>
<td>0.627</td>
<td>0.525</td>
<td>0.610</td>
<td>0.659</td>
<td>0.558</td>
</tr>
<tr>
<td>pMKP7</td>
<td>0.483</td>
<td>0.535</td>
<td>0.431</td>
<td>0.998</td>
<td>1.000</td>
<td>0.986</td>
</tr>
<tr>
<td>eIf4G_Pi</td>
<td>0.498</td>
<td>0.550</td>
<td>0.446</td>
<td>0.357</td>
<td>0.408</td>
<td>0.309</td>
</tr>
<tr>
<td>Viral_Protein</td>
<td>0.000</td>
<td>0.011</td>
<td>0.000</td>
<td>0.377</td>
<td>0.428</td>
<td>0.328</td>
</tr>
</tbody>
</table>

Figure 4.3.2a – Modelling the Effects of MNV Infection. Outline of the effects MNV infection has upon the cellular signalling network and translation initiation pathway. QSSPN simulations, followed by Reachfq, generated information on the activation or production of various model effectors. Comparison of the Binomial Confidence Intervals revealed regions of significant difference, as shown in red.
Section 4.3.2.1 – Experimental Validation of the Hypotheses Resulting from Simulations of the Murine Norovirus Model

The MNV model demonstrates a clear, significant increase in CREB phosphorylation in response to MNV infection. Additionally, no significant differences were noted in the phosphorylation of Glycogen Synthase Kinase 3 (GSK3), MKK3 and MKK6. The data collected by N. Doyle and E. Royall using Human Phospho-MAPK Arrays enabled validation of this prediction. To this end, RAW 264.7 cells were infected with MNV strain CW1 at an MOI of 10. Uninfected cells were used as a control. Both infected and uninfected cells were incubated for two and 12 hours prior to lysis. Cells were lysed using Lysis Buffer 6 according to the manufacturers’ instructions.

As shown in Figure 4.3.2.1a, although no difference in the phosphorylation state of CREB was seen at 2 hours post infection (hpi), there is a significant difference between the uninfected, mock and infected samples at 12hpi ($p = 0.0019$). At 12hpi, MNV infection induces the phosphorylation of this transcription factor. Moreover, a comparison between the infected samples at 2hpi and 12hpi reveals a significant increase in CREB phosphorylation ($p < 0.0001$). No significant difference between the mock, control time points was seen.

In the cases of GSK-3, MKK3 and MKK6, the experimental results demonstrated no significant differences in the level of phosphorylated forms of any of these kinases at either 2hpi or 12hpi (data not shown). These findings indicate the model is capable of predicting a diverse array of signalling behaviours.
Only one point of difference was noted between the behaviour of the model and the experimental data. MSK2 phosphorylation was predicted to increase however, the experimental data failed to detect a significant increase. It is worth noting that a non-significant increase in MSK2 phosphorylation ($p = 0.1792$) was noted in the Human Phospho-MAPK Array. This increase may become significant if the time course were extended.
Figure 4.3.2.1a – Effect MNV Infection has Upon CREB Phosphorylation.

RAW 264.7 cells were infected for either two or 12 hours. These cells (MNV), along with uninfected samples (mock) incubated for the same length of time, were lysed. The protein concentration was determined using the BCA Assay. Human Phospho-MAPK Arrays using 300μg of protein from each lysate were carried out in accordance with the manufacturers’ instructions. Quantification was carried out in ImageJ. Statistical analysis, using GraphPad Prism 6, involved two-way ANOVA with Tukey’s post-hoc test of significance. The data presented has been normalized to the array control spots, as detailed in Royall et al. (2015), and therefore expressed as Relative Pixel Density. Error bars represent the Standard Error of the Mean. The data presented above was supplied by N. Locker and obtained by N. Doyle and E. Royall.
Section 4.3.3 – Simulation of the Andes Virus Models

The final virus investigated in this project was ANDV. In terms of the effects this virus has upon the signalling network, only very few documented affects were previously noted and consequently this virus has the smallest number of perturbations implemented into the final version of the binary model.

Reachfq analyses were performed on the QSSPN simulation results. A comparison of the Binomial Confidence Intervals calculated for each of the results, shown in Figure 4.3.3a, was used to find regions of significant difference (as shown in dark purple). The only pathway altered in this model was the mTORC1 pathway. As can be seen in Figure 4.3.3a, this pathway was significantly increased. Surprisingly, although viral protein was produced and this required a cap-snatching mechanism, no significant reduction in host protein synthesis was observed. It is worth noting that the work of Jääskeläinen and co-workers (2007) demonstrates that other Hantaviruses, namely Tula Virus and Puumala Virus, do not induce a strong inhibition of cellular protein synthesis. This work is somewhat supportive of the finding here. The polymerase encoded by the Large genome segment has been shown to not decrease the expression of cellular proteins in general, but it has been shown to specifically limit the production of thoses proteins only of high molecular weight (Heinemann et al., 2013). It could be argued, therefore, that the non-significant reduction of cellular protein synthesis, shown in Figure 4.3.3a, is in agreement with the view that no general suppression of protein synthesis is seen during ANDV infection.
Figure 4.3.3a - Modelling the Effects of ANDV Infection. These simulation results show the effects ANDV infection has upon the cellular signalling network and the translation initiation pathway. Binomial Confidence Intervals of Reachfq analyses were calculated for each model effector from the QSSPN simulations. Regions of significant difference, shown in dark purple, are found where no overlap in confidence intervals is present.
Section 4.4 – Discussion

The elaboration of the binary model to include viral effects was necessary to further validate the benefits of taking a systems approach to understanding complex biological systems.

Section 4.4.1 – Problems Encountered & Troubleshooting

Given that up until this point it has been possible to give a quantitative measure of the predictive power of the model, the single biggest problem encountered in this part of the work was the lack of a suitable benchmark dataset to establish the predictive power of the models. Data concerning the effects of these viruses is very limited and the action of inhibitors is not directly transferable from the datasets used to quantify the predictive power of the cellular translational control model.

The only solution to this would be to test the effects each of the inhibitors has when combined with infection. However, this is not feasible and so the solution implemented here is to evaluate the effects qualitatively. For example, Rodríguez et al. (2014) noted that ERK1/2 activity is promoted by JUNV infection and so this effect was incorporated into the model. By comparing the simulation results for the ‘mock’ and ‘infected’ models it is possible to qualitatively prove that the model is functioning in accordance with published data. Furthermore, even if efforts were made to quantify these effects, such as by comparing the literature and simulation results in a Confusion Matrix, the small number of data points available for each model would mean that the resulting MCC value is of little benefit.

As already noted in Section 3.4.4 – Problems Encountered & Troubleshooting, the use of pixel counting to quantify the Human Phospho-MAPK Arrays in validating MNV predictions, is not considered particularly accurate due to the non-linear response of radiographic film to
a fluorescent signal (Gassmann et al., 2009). However, it was used in this work as it was the method recommended by the array manufacturer.

Section 4.4.2 – Potential Future Work

Both the literature examined during the construction of the viral models and the simulation results indicate that viral infection can potentially affect the global metabolic flux distributions. This motivates future work which is aimed at utilising the full capabilities of the QSSPN software to integrate the signalling network model with a GSMN. Chapter 5 – Integration of the Translational Control Model with a Genome Scale Metabolic Network of a RAW 264.7 Cell describes the first version of such an integrated model. However, given the low predictive power of the integrated model, when evaluated against the Metabolic Benchmark dataset, more work is needed before this model can be used to generate hypotheses directing experimental work. However, examining the effects of viral infection upon the cellular metabolic network would have added significant value to this section of the project. Future use of this type of model and in this way should involve working towards understanding the impact of viral infection upon the cellular metabolic network. This would be achievable with this type of model if the predictive power were raised considerably.

Furthermore, experimental works looking at predictions generated by the MNV model are needed. Given that a number of these predictions involve looking at the post-translational modification, particularly phosphorylation, of well-recognized signalling components (such as MK2 and MSK1), it should be possible to use immunoblotting for many. A possibility also exists to examine the effects of the increases to transcription factor activity, such as CREB and ATF-2, through the use of the Polymerase Chain Reaction (PCR) technique targeting
genes under the control of these transcription factors. Conducting these works would give a broader understanding to the effects MNV infection has upon the cell.

Developing a method for quantifying the predictive power of the model should also be considered if modelling of this type is to be used when looking at host-pathogen interactions. Despite the difficulties noted in Section 4.4.1 – Problems Encountered & Troubleshooting, a method based upon the one used throughout this project may be viable. For this to be the case, increasing the number of data points used in the calculation of the MCC value must be achieved. The increasing availability of transcriptomics and proteomics datasets looking at viral infection, such as those generated by Aevermann et al. (2014), will hopefully begin to provide access to large-scale datasets that would be necessary for a quantitative measure of the predictive power of the model.

Section 4.4.3 – Concluding Remarks

The incorporation of the various effects viral infection can induce, in both the cellular signalling network and to the translational machinery, was carried out. Moreover all three viruses represent understudied, important human pathogens, or in the case of MNV, a surrogate of one. Whilst it is not being suggested that the viruses used here all infect RAW 264.7 cells, the three viruses chosen in this work were selected as they produce different effects to the cellular signalling network and the translational apparatus. With this in mind, the number of effectors monitored during simulations was substantially increased so as to be able to detect more widespread effects.

The work presented here differs from other published works looking at viral infection from a systems biology perspective. The model to which the viral effects were added was a large-scale reconstruction of the signalling network that surrounds and regulates translation
initiation in mammals. Other models, such as the work of Jensen and co-workers (2013), were very limited in scope. Whilst the approach of Jensen et al. (2013) does lead to experimentally validated findings, it does limit the future uses of the model. The main benefit of using a literature-based reconstruction model is that the model can be adapted with relative ease to simulate new viral or, more generally, pathological effects. Moreover, the qualitative nature of the model means that kinetics of viral perturbations are not necessary and do not require proper integration with the kinetic parameters of the host. Furthermore, the scope of this model means that it is possible to monitor the potential impact of infection upon a wider range of cellular effectors and consequently lead to the generation of more testable behaviours.

The model of JUNV infection demonstrates that large-scale aberrations to the cellular signalling network are possible even when only a few initial alterations are made. This demonstrates the high degree of cross-talk between signalling pathways, and the effects that this has upon the functionality of the model as a whole. Consequently, this highlights just how viral infection can be thought of as a ‘systems-level perturbation’ (Garmaroudi et al., 2010) to the host. The results of this set of simulations are suggestive of JUNV producing effects to numerous signalling pathways adjacent to ERK1/2 and PI3K, such as MSK1. Furthermore, these simulations suggest that the increase in ERK1/2 activity promotes alterations to the metabolic network through mTORC1.

Of the three viruses selected, perhaps the most disappointing model was that of ANDV infection. Both Gavrilovskaya et al. (2013) and McNulty et al. (2013) demonstrate mTORC1 activation and Cimica and co-workers (2014) have shown that Interferon signalling is affected by ANDV infection. Despite these findings, little else is known as to the cellular
effects of ANDV infection. It was hoped that this model would shed light on additional pathways activated by ANDV. However, the predictions generated were both few in number and for the most part, directly downstream of mTORC1. The reduction of eIF4G phosphorylation was also noted. This finding is suggestive of a reduction of mTORC2 activity, possibly as a result of the increase in S6K1 activity (Julien et al., 2010). However, with no reduction in either AKT1 or PKCα phosphorylation, which are both mTORC2 substrates in the model, it is not possible to state this with any certainty. This being said, this model had the fewest number of well documented perturbations to input into the model. However, the effectors with altered levels of activity are suggestive of cellular metabolic reprogramming. The implication of this model yielding only few predictions is redolent of both the need for more research into ANDV infection and to ensure that models are as expanded as possible to make certain that important events can be properly placed into context.

Given safety considerations, it is difficult to undertake work to validate hypotheses generated about either JUNV or ANDV infection. However, work with MNV is possible and may give valuable insight into the molecular biology of the important human pathogen, HuNV. As already stated in Section 4.4.2 – Potential Future Work, the effectors altered during MNV infection may lead to vast changes to the cell. Alterations to the level of activity of transcription factors may produce large shifts in gene expression. Both CREB (Zhang et al., 2005) and ATF-2 (Bailey & Europe-Finner, 2005) can regulate genes involved in a diverse array of important cellular processes. To date, with the exception of Royall et al. (2015), very little work has been embarked upon to examine the effects members of the Norovirus genus can induce in the cellular signalling network.
The work with a model of MNV infection has revealed that the virus may affect a large number of important signalling pathways. Whilst the findings that ERK1/2 and p38 were activated was expected (Royall et al., 2015), the activation of downstream effectors of these pathways, notably Elk-1, MK2 and MKP-7, is a novel hypothesis. PKCα activity has been linked positively, in the case of Respiratory Syncytial Virus entry (San-Juan-Vergara et al., 2004), or negatively, in the case of Vesicular Stomatitis Virus replication (Zhu et al., 2011), to the regulation of viral infection. However, to date, the involvement of PKCα activity has not been investigated in MNV infection. This work suggests that PKCα activity is not altered by MNV infection.

The ability to experimentally validate findings of the model was a significant milestone for this section of the project. The predicted increase in CREB phosphorylation, as shown in Section 4.3.2 – Simulations of the Murine Norovirus Models, in response to MNV infection was experimentally verified. The unpublished work by N. Doyle and E. Royall with Human Phospho-MAPK Arrays indicated that CREB phosphorylation was increased during MNV infection at 12hpi (as shown in Section 4.3.2.1 – Experimental Validation of the Hypotheses Resulting from Simulations of the Murine Norovirus Model). As detailed in Royall et al. (2015), RAW 264.7 cells were infected with MNV at an MOI of 10. Following 2hpi or 12hpi, the cells were harvested and the lysates used in Human Phospho-MAPK Arrays. Given that CREB phosphorylation was not specifically targeted for perturbation when the viral effects were added to the MNV model, this can be viewed as experimental validation of one of the predicted behaviours. This is true despite the behaviour of CREB not being unexpected given the position of this transcription factor downstream of both ERK1/2 (Impey et al.,

Although no work has investigated the effect of the MNV-induced increase in CREB phosphorylation, as noted in **Section 4.4.2 – Potential Future Work**, this could possibly be achieved through the use of PCR aimed at investigating if CREB target genes exhibit expression changes as a result of this. An increase in CREB phosphorylation on the Ser-133 site is not unique to MNV infection. Both Human Immunodeficiency Virus 1 (Gibellini *et al.*, 1998) and Human T-Cell Leukaemia Virus (HTLV-1) (Kim *et al.*, 2007b) are known to induce CREB phosphorylation on the Ser-133 residue. HTLV-1 appears to require this phosphorylation of CREB, through the viral protein Tax, so as to enhance transcription of the viral genome (Geiger *et al.*, 2008) and maybe functions to regulate apoptosis (Trevisan *et al.*, 2004; Trevisan *et al.*, 2006 (as cited by Geiger *et al.*, 2008)). Given that MNV (Bok *et al.*, 2009), and FCV (Roberts *et al.*, 2003; Sosnovtsev *et al.*, 2003), have been shown to induce apoptosis, any anti-apoptotic roles for phosphorylated CREB seem questionable but it remains to be seen if CREB is required for viral transcription. Bok and co-workers (2009) noted that viral RNA achieves a maximal level at 8hpi with the level being maintained at 12hpi. This is in keeping with the work presented here in which an increase in CREB phosphorylation was noted between 2hpi and 12hpi.

The factor resulting in an increase in p38 activity during MNV infection is not known. The results presented here and in the array data showing that neither MKK3 nor MKK6 exhibit a concomitant increase in phosphorylation, suggest that a viral protein maybe responsible for this increase. Many other viruses, including JUNV (Rodríguez *et al.*, 2014) and members of
the Flavivirus genus (Lin et al., 2006; Werme et al., 2008; Laurent-Rolle et al., 2010), use viral factors to alter the activation of cellular signalling pathways.

The behaviour of the other model components, GSK3, MKK3 and MKK6, were also experimentally validated. In all three cases, no significant increase in phosphorylation was predicted. The Human Phospho-MAPK Array data confirmed this behaviour. Whilst the model failed to predict that MSK2 phosphorylation would be unaffected by MNV infection, the validation of four predicted behaviours so far indicates that the MNV model has an 80% success rate and, whilst not directly comparable, is indicative of a positive MCC value.

The models presented in this chapter reveal that it is possible to mimic viral infection and predict additional behaviours. As has been shown here, a systems approach is capable of identifying cellular targets involved in viral infection which may not have been immediately obvious targets of study. On the whole, all viruses examined here produce a wide variety of cellular affects and in all cases generate findings that are not readily explainable, which is indicative of complex, non-trivial behaviour. As such the work presented here can be seen as providing a new method for modelling the complex network of host-pathogen interactions that accompany infection. Whilst many of the predicted behaviours are not yet experimentally validated, the background model of cellular translation has been validated (as shown in Chapter 3 – Experimental Validation of Model Predictions) and demonstrated as having a significant degree of predictive power (as shown in Chapter 2 – Model Creation & Refining). To conclude, the simulations presented here generated plausible hypotheses that can be used to direct future experimental research on the interactions of viruses with the host translational machinery and signalling network.
Chapter 5 – Integration of the Translational Control Model with a Genome Scale Metabolic Network of a RAW 264.7 Cell
Section 5.1 – Introduction

This chapter covers the work aimed at linking the model of the regulation of translation to a GSMN, Flux Balanced Analysis model of a RAW 264.7 cell. This model was developed by Bordbar et al. (2012) and is based upon transcriptomic, proteomic and metabolomic data. There is a growing body of literature that highlights the pronounced effects the host-pathogen interactions have upon the metabolic network of the cell (Wang et al., 2008b; Saric et al., 2010; Coelho et al., 2015; Milner et al., 2015), or of the pathogen (Raghunathan et al., 2010). This motivates the integration of the translational control model with the GSMN to study whether translation initiation is co-regulated with the biosynthesis of metabolic precursors of protein synthesis.

In previous chapters, the level of predictive power of the model was raised to an MCC = 0.4558 and several model behaviours have been validated. From here, attention was turned to carrying out preliminary work of investigating the dual regulation of translation and metabolism. Additionally, as the method of simulating the model detailed in Chapter 2 – Model Creation & Refining is computationally demanding, several attempts in this chapter, were made at developing a method for running less computationally demanding simulations, which yield results of comparable predictive power to the existing method.

Section 5.1.1 – Linking mTORC1 via S6K1 to the Regulation of Metabolism

An emerging role for mTORC1 is being discovered in the regulation of metabolism. mTORC1 regulates anabolic metabolism by promoting the synthesis of new proteins, nucleotides and lipids while at the same time inhibiting the cellular process of autophagy (Shimobayashi & Hall, 2014). This regulation of anabolic metabolism occurs via either a regulation of transcription or post-translational modification of transcription factors or enzymes central
to a particular metabolic pathway. As shown in Figure 5.1.1a, the majority of these effects have been shown to be dependent on the mTORC1 substrate, S6K1.

One of the core metabolic pathways regulated by mTORC1 is the glycolytic pathway. This regulation has also been shown to be required for carrying out normal cellular functions and immune regulation (Donnelly et al., 2014). The ability to regulate glycolysis is mediated by the transcription factors, c-Myc and Hypoxia Inducible Factor-1α (HIF-1α) (Düvel et al., 2010). Both of these transcription factors are capable of regulating multiple enzymes in the glycolytic pathway (Ibid). Moreover, despite being capable of regulating the expression of distinct enzymes, both c-Myc and HIF-1α can bind to similar nucleic acid sequences and, despite sometimes acting in an antagonistic fashion (Mazure et al., 2002; Koshiji et al., 2004), function to enhance the expression of many of the same glycolytic enzymes, such as Phosphofructokinase, Hexokinase and Triose Phosphate Isomerase (Gordan et al., 2007). Additionally, mTORC1 activity has been shown to not only up-regulate Glucose Transporter 1 expression but also promotes membrane localization of the glucose transporter (Makinoshima et al., 2015).

S6K1 mediates HIF-1α and c-Myc transcription and translation through the phosphorylation of Signal Transducer and Activator of Transcription 3 (STAT3) (Kim et al., 2009a; Thiem et al., 2013; Dodd et al., 2014; Li et al., 2014; Rad et al., 2015) and eIF4B (Tandon et al., 2011; Csibi et al., 2014; Dodd et al., 2014). Regulation of pyrimidine synthesis occurs through S6K1-mediated phosphorylation of the rate-limiting enzyme Carbamoyl-Phosphate Synthetase 2, Aspartate Transcarbamylase and Dihydroorotase (CAD) (Lindsey-Boltz et al., 2004; Ben-Sahra et al., 2013; Robitaille et al., 2013). The final pathway canonically regulated by mTORC1 is lipid and sterol biosynthesis. The transcription factors, Sterol-Regulatory Element Binding
Proteins (SREBP) 1 and 2, are regulated at the level of expression by the mTORC1 substrate, Retinoid X Receptor (Norrmén et al., 2014) whilst proteolytic activation of these transcription factors is mediated by S6K1 (Düvel et al., 2010).
Figure 5.1.1a – Overview of the Role Played by S6K1 in Regulating Multiple Metabolic Pathways. The mTORC1 substrate S6K1 is responsible for carrying out many of the regulatory roles ascribed to mTORC1 in the regulation of metabolism. In the phosphorylated form, S6K1 can bring about changes in gene expression, by either increasing the level of phosphorylated eukaryotic Initiation Factor 4B (eIF4B), mediating the proteolytic processing of the Sterol-Regulatory Element Binding Proteins (SREBP) 1/2 (SREBP1/2c), or by directly phosphorylating Carbamoyl-Phosphate Synthetase 2, Aspartate Transcarbamylase and Dihydroorotase (CAD).
Section 5.2 – Methods

Section 5.2.1 – Proposed New Approaches

The asynchronous stochastic method outlined in Chapter 2 – Model Creation & Refining uses a Monte Carlo method to analyse sequences of feasible molecular interactions, given the connectivity of the model. As described in Section 2.1.4.2 – Evaluation of the Predictive Power of the Model, the 1,000 trajectories sampled for the Wild Type and inhibitor models were interrogated using Reachfq and 99.99% Binomial Confidence Intervals calculated and regions of significant difference were identified. A benchmark dataset (the Metabolic Benchmark dataset) detailing published effects of each inhibitor, on the selected model components, was used to validate the behaviour of the model and, quantitatively, analyse the predictive power of the model. Both the MCC (Matthews, 1975) and the AC (Burset & Guigó, 1996) were again used to convert the qualitative measures of predictive power, given by the comparison of simulation behaviours to the benchmark datasets, to a quantitative measure. The equations used to calculate these measures of predictive power are given in Section 2.1.4.2 – Evaluation of the Predictive Power of the Model.

The asynchronous simulation is computationally more demanding, as large numbers of molecular interactions firing need to be generated and saved to disk space for further analysis. This creates a challenge for QSSPN simulations integrating GSMN models. In every instance where a Constraint Place state change, the FBA of the entire GSMN needs to be run multiple times to update the states of objective nodes. Since FBA is the most time consuming operation in the QSSPN algorithm, this increases the computational cost when compared with the un-integrated signalling network alone.
In order to attempt to reduce the computational demands of the simulations, additional method were employed. The synchronous execution of at least some molecular interactions would lead to a reduction in the number of iterations and thus, diminish the number of times the FBA would need to be performed. Consequently, the simulation output files would be smaller.

The first alternative method proposed was aimed at interpreting the model as a series of ODEs (as discussed in Section 2.1.4.3.1 – Case Study Identifying Differences in the Asynchronous Simulation Approach & the Synchronous Methods). This deterministic approach was selected to apply to the full model, due the computational benefits this method has over the asynchronous stochastic method. In this approach, the reaction kinetic parameters are set equal to one. If a molecular species is present, the species concentration variable becomes one and the reaction is enabled. Inhibitor concentrations are seen as taking a value of one if it is absent, and so enabling the molecular interaction, or zero if the species is marked with a molecular activity level. This would have the effect of reducing the solution to that particular ODE to zero. A further synchronous simulation approach was tested, in which all enabled molecular interactions to fire in each time-step. However, due to the deficiencies discussed in Section 2.1.4.3 – Synchronous Simulation Methods and Section 2.1.4.3.1 – Case Study Identifying Differences in the Asynchronous Simulation Approach & the Synchronous Methods this method was tested only with Rapamycin incorporated into the model to establish if this approach could produce the expected model predictions.
Both the asynchronous stochastic and ODE methods were implemented using version 3.0 of the software QSSPN and developed by A. Kierzek (as part of the QSSPN formalism). These methods were run using an identical model so as to ensure the results were comparable.

The QSSPN software was used here, as it allowed the integration of a qualitative model of the regulation of translation initiation with a GSMN of a RAW 264.7 murine macrophage. This GSMN has been shown to be capable of reproducing experimentally derived data (Bordbar et al., 2012), and is constraint-based. The integration is achieved by identifying transcription factors regulated by mTORC1 and linking these transcription factors to genes found within the GSMN. As noted by Fisher and colleagues (2013), the activation of the constraint-based GSMN is achieved only if the activation thresholds are reached in the qualitative model. 23 metabolic genes, from a range of metabolic pathways including the pentose phosphate pathway, glycolysis and lipid and sterol biosynthesis, were linked to mTORC1 activation in this way. The functioning of the integrated model was confirmed by the production of Biomass. Biomass, as a ‘metabolite’ within the flux balanced analysis model, is the sum of all the external metabolites produced by the model, as such Biomass production was viewed as being required for translation to occur.

Section 5.3 - Results

Section 5.3.1 – Current, Asynchronous, QSSPN Simulation Method

Initial integration of the models was aimed at linking the mTORC1 signalling pathway to the regulation of metabolic pathways found in the RAW 264.7 cell GSMN. This integration resulted in 23 genes required for multiple metabolic pathways, including glycolysis, the pentose phosphate pathway and lipid/sterol biosynthesis, being placed under the control of mTORC1. In order to incorporate the expression of 23 new metabolic genes and the effects
of the mTORC1 substrate, S6K1 on these genes, it was necessary to add 197 novel molecular interactions. The expression of these 23 genes required the addition of multiple new molecular species including several transcription factors and the stages of expressing these genes. In total, including the repetition of molecular species to account for involvement in multiple molecular interactions, 188 new Petri Net components, to model these species, were incorporated into the model.

Before the two new methods could be tested, a point of comparison was necessary. This point of comparison was obtained by running simulations using the method detailed in Chapter 2 – Model Creation & Refining. The novel methods required version 3.0 of QSSPN. Since the implementation of the synchronous approach altered the main simulation loop in the QSSPN code, validation of the new code was necessary. Accordingly, version 1.0 of QSSPN was also included in this new round of simulations to ensure that functionality of QSSPN had been conserved through to version 3.0. The dataset against which the model integrated with the RAW 264.7 cell GSMN was tested is shown in Supplementary Information 3. This dataset was termed the Metabolic Benchmark, composed of 186 literature articles, concerned the effects of chemical inhibitors on key model effectors. The bibliographic information concerning this Benchmark can be found in Supplementary Information 1a. Simulations show that the results obtained with the asynchronous method implemented in version 3.0 of the QSSPN software are identical to the results obtained with version 1.0.

The simulations run in this section were done using the method presented and used elsewhere in this project. These results were obtained using version 3.0 of QSSPN. These results (shown in Supplementary Information E7 – Frequency Data Generated by Reachfq
with Associated 99.99% Binomial Confidence Intervals (sheet entitled Asynchronous Method Results)) were then to be used as a benchmark against which the other two, alternative methods would be tested. The upper and lower 99.99% Binomial Confidence Intervals, calculated in the software R, were then calculated for each identified key model effector.

As Rapamycin has been shown to affect eIF4E phosphorylation via MNK2, elsewhere in this project, it is worth noting that the integration of the Petri Net model with the RAW 264.7 cell GSMN did not eliminate this prediction. The frequency of trajectories resulting in eIF4E phosphorylation was reduced when Rapamycin was introduced to the model (0.472 in the Wild Type model versus 0.007 in the Rapamycin model). Moreover, the frequency of trajectories leading to MNK2 phosphorylation was also altered, with an increase from 0.502 in the Wild Type model to 0.845, in the Rapamycin model.

The summary of the Confusion Matrices, shown in Figure 5.3.1ai, describe the comparison of the results of simulations with QSSPN version 3.0, with the literature-based Metabolic Benchmark dataset. For each inhibitor model, a summary of the predictive power is given. The overall measure of predictive power, given in Figure 5.3.1aaii, contains the pooled results of all the matrices. This then provides enough information to make a reasonable assessment of the overall predictive power of the model. The results of individual models are of interest when looking at where model improvements are needed.
i.

<table>
<thead>
<tr>
<th>Model</th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
<th>MCC Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIX02189</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>3</td>
<td>0.4183</td>
</tr>
<tr>
<td>CGP57380</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>0.8081</td>
</tr>
<tr>
<td>GSK2334470</td>
<td>2</td>
<td>11</td>
<td>1</td>
<td>2</td>
<td>0.4623</td>
</tr>
<tr>
<td>LY294002</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td>0.0537</td>
</tr>
<tr>
<td>PD184352</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>0.2402</td>
</tr>
<tr>
<td>PD98059</td>
<td>11</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0.6268</td>
</tr>
<tr>
<td>Rapamycin</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>0.3812</td>
</tr>
<tr>
<td>SB203580</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>-0.2630</td>
</tr>
<tr>
<td>SP600125</td>
<td>11</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>0.7100</td>
</tr>
<tr>
<td>XMD8-92</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0.6667</td>
</tr>
</tbody>
</table>

ii.

<table>
<thead>
<tr>
<th>Model</th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
<th>MCC Value</th>
<th>AC Value</th>
<th>α-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIX02189</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>3</td>
<td>0.4183</td>
<td>0.3789</td>
<td>4.4x10^-7</td>
</tr>
<tr>
<td>CGP57380</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>0.8081</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSK2334470</td>
<td>2</td>
<td>11</td>
<td>1</td>
<td>2</td>
<td>0.4623</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LY294002</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td>0.0537</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PD184352</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>0.2402</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PD98059</td>
<td>11</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0.6268</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rapamycin</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>0.3812</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SB203580</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>-0.2630</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP600125</td>
<td>11</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>0.7100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XMD8-92</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0.6667</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 5.3.1a – Predictive Power of the Model of Translational Regulation Integrated with a GSMN of RAW 264.7 Cell Metabolism and Simulated with the Current, Asynchronous Method.  (i)

Following the simulation and interrogation of the model with QSNN (Version 3.0) and Reac fiz, the frequency of trajectories resulting in the activation of each model effector, were calculated for the Wild Type and inhibitor models. By calculating the 99.99% Binomial Confidence Intervals for each effector, and comparing the Wild Type and inhibitor models, it was possible to identify regions of significant difference. Comparing the behaviour of the model to the Metabolic Benchmark dataset allowed the construction of Confusion Matrices with behaviour being characterised as being either True Positive (TP), True Negative (TN), False Positive (FP) or False Negative (FN). Calculating the Matthews’ Correlation Coefficient (MCC) converted this qualitative measure of predictive power to a quantitative one. (ii) The overall predictive power of the model was calculated by pooling the variables in the Confusion Matrices of the individual inhibitors. The Approximate Coefficient (AC) was also calculated. By conducting a Chi Square Test, using the resulting α-Value, it was possible to confirm that the predictive power of the model was significantly different from random.
In order to ensure the continued functioning of QSSPN, it was necessary to ensure agreement between versions 1.0 and 3.0. The data shown in Supplementary Information E7 – Frequency Data Generated by Reachfq with Associated 99.99% Binomial Confidence Intervals details the values obtained with version 3.0. Although the data is not shown here, those values are identical to those obtained with version 1.0 and so demonstrates that functionality of the QSSPN software is conserved. Integration of the model is known to be functioning as Biomass was produced in the Wild Type model, albeit at a low level with the frequency of trajectories reaching Biomass production being 0.044. Such a low frequency is not unexpected given the experimentally derived lower bound of the Biomass reaction within the RAW 264.7 cell GSMN model was set to 0.0281/h (Bordbar et al., 2012). This may be indicative of a need for a more comprehensive integration of the GSMN with the signalling pathways.

During this evaluation of the functioning of the model, the alteration to the level of phosphorylated eIF4E and a reduction to phosphorylated AKT1 during Rapamycin incorporation were maintained. This suggests that the functioning of the model was maintained during the integration with the RAW 264.7 cell GSMN. Despite the two experimentally validated predictions being maintained through the integration process, the MCC value is markedly, but not significantly (as determined by a Chi Square Test, $\alpha = 0.305$), lower than it was prior to the integration with the GSMN (MCC = 0.3759 versus MCC = 0.4558 (as calculated in Section 2.2.4 – Evaluation of the Predictive Power of the Final Model)). Given this reduction in predictive power was quite substantial, albeit not significant, it was decided that the method that produced the smallest decrease in MCC
value would be the best method to take forward, even if this means simulations are relatively expensive in a computational sense.

**Section 5.3.2 – Proposed New Methods**

**Section 5.3.2.1 – Ordinary Differential Equation Method**

The first proposed method was deterministic and treated the model as a series of ODEs. Consequently, only one trajectory was necessary. While ODEs are usually used to express quantitative models, they were applied here to a qualitative model. All continuous variables were allowed to vary between zero and one. All kinetic parameters are set equal to one. If a molecular species is not marked with any discrete activity levels, then this parameter is multiplied by zero. Conversely, if molecular species are marked with discrete activity levels, then this is multiplied by one.

Attempts with this method revealed that the predicted behaviours of alterations to the level of eIF4E phosphorylation and a reduction in AKT1 phosphorylation, were shown to occur, as shown in Figure 5.3.2.1a.
Table 5.3.2.1a – Summary of the Results of the Ordinary Differential Equation Simulation Method. This method evaluates the model as a series of ODEs in which the kinetic parameters are set equal to one and are multiplied by one if the molecular species are marked with discrete activity levels. This subjective, yet deterministic, method is evaluated in an equally qualitative manner by comparing the Wild Type model and an Inhibitor model. If an increase of more than 10% in the amount of each effector is observed in the simulation a ‘↑’ is given. Conversely, for a decrease of more than 10% a ‘↓’ designation is listed. Where no change occurs, the result is given by ‘-’. For marginal increases or decreases (less than 10%), combinations of these are given.
It is worth noting that the use of the word ‘subjective’, in Figure 5.3.2.1a, to describe the results of the ODE method refers only to the fact that as rates are multiplied by one if discrete acitivty levels are present in the molecular species serving as reactants for a particular molecular interaction. As a consequence of this use of pseudo-kinetic parameters the traces of molecular amounts must be regarded as subjective. Had kinetic parameters been available concerning the concentration of particular species and the rate constants for model reactions, the output of the ODE model would be absolute with the presence of large molecule numbers minimising deviations from the mean (Kleinstein & Singh, 2001).

The ODE method is a deterministic method that has the potential to greatly improve the existing methodologies of systems biology but also suffers from significant drawbacks. This method is attractive in terms of the contribution it makes towards the aim of making simulations computationally less costly. The need for only one trajectory both speeds up the running of simulations and reduces the computational power required. However, the deterministic properties of this method, whilst attractive, suffer greatly from a lack of realistic kinetic parameters. The results of this simulation suffer from a lack of sensitivity, as shown in Figure 5.3.2.1a. In this Figure, no observed change in the trajectory was denoted by ‘-’, a marginal increase or decrease (of less than 10%) is denoted by ‘↑/-’ or ‘↓/-’, respectively and larger increases or decreases, of greater than 10%, are given by ‘↑’ and ‘↓’, respectively. Only gross changes are visible using this method. The summary of the Confusion Matrices shown below, Figure 5.3.2.1b, demonstrate this in the marked reduction in the predictive power.


i.

<table>
<thead>
<tr>
<th>Model</th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
<th>MCC Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIX02189</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>3</td>
<td>0.4183</td>
</tr>
<tr>
<td>CGP57380</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0.5833</td>
</tr>
<tr>
<td>GSK2334470</td>
<td>2</td>
<td>11</td>
<td>1</td>
<td>2</td>
<td>0.4623</td>
</tr>
<tr>
<td>LY294002</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>-0.3110</td>
</tr>
<tr>
<td>PD184352</td>
<td>8</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>-0.2070</td>
</tr>
<tr>
<td>PD98059</td>
<td>10</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>0.3508</td>
</tr>
<tr>
<td>Rapamycin</td>
<td>11</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>-0.2141</td>
</tr>
<tr>
<td>SB203580</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>7</td>
<td>0.1207</td>
</tr>
<tr>
<td>SP600125</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>9</td>
<td>0.3508</td>
</tr>
<tr>
<td>XMD8-92</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0.4082</td>
</tr>
</tbody>
</table>

ii.

<table>
<thead>
<tr>
<th>Model</th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
<th>MCC Value</th>
<th>AC Value</th>
<th>α-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>53</td>
<td>37</td>
<td>12</td>
<td>43</td>
<td>0.2922</td>
<td>0.2926</td>
<td>1.38x10^-5</td>
</tr>
</tbody>
</table>

Figure 5.3.2.1b – Overview of the Predictive Power of the Model when Simulated with the Ordinary Differential Equation Method. (i) The interactions in the Petri Net model were converted to a series of Ordinary Differential Equations. The tool to do this was developed by A. Kierzek as part of the QSSPN project. By comparing the traces of each effector in the Wild Type model to those in each inhibitor model, it was possible to identify regions of difference. This behaviour of the inhibitors in the model were then compared to the Metabolic Benchmark dataset. This behaviour was characterised as either being a True Positive (TP), True Negative (TN), False Positive (FP) or a False Negative (FN). These Confusion Matrix variables were then converted to a quantitative measure of predictive power in the form of the Matthews’ Correlation Coefficient (MCC). (ii) By pooling these variables for each inhibitor, it was possible to calculate an overall MCC value. This overall measure of predictive power was confirmed by calculating the Approximate Coefficient (AC). Moreover, to demonstrate that the level of predictive power was significantly different from random, the α-Value of a Chi Square Test was calculated.
This marked reduction in the MCC values between the current, asynchronous method (MCC = 0.3759) and this ODE method (MCC = 0.2922), whilst appearing to be non-significant (as determined by a Chi Square Test, α = 0.492), was of great importance in deciding which method was to be taken forward for use in later simulations.

**Section 5.3.2.2 – Synchronous Method**

Although synchronous simulations demonstrated that Rapamycin was able to alter the level of phosphorylated eIF4E and reduce AKT1 phosphorylation, it did produce some rather questionable results from the model of translational regulation integrated with the GSMN. As expected from **Section 2.1.4.3.1 – Case Study Identifying Differences in the Asynchronous Simulation Approach & the Synchronous Methods**, MNK2 phosphorylation was abolished by this method.

Signalling pathways require a highly coordinated response to a signal in order for signal transduction to occur. Each interaction must occur in a tightly temporally regulated fashion (Kholodenko *et al.*, 2010). The synchronous approach assumes that all molecular interactions proceed with the same velocity. While it is easy to understand how the temporal nature of a tiered signalling pathway is essential, it may be less clear for cross-talk between signalling pathways. In a MAPK cascade for example, in order for the next kinase to be phosphorylated the upstream kinase must be activated (*Ibid*). Similarly, the temporal nature of feedback mechanisms are clear. During negative feedback, a downstream effector acts to inhibit the pathway and thus prevents further transduction. This allows oscillatory behaviour to be observed (Kholodenko, 2006). However, in order to understand how cross-talk between signalling pathways is temporally regulated, it is necessary to note that the nature of cross-talk is complex and may only occur if sustained activation of a pathway is
achieved (Housden & Perrimon, 2014). Moreover, even within a particular pathway the efficacy of the interaction between the kinase and differing downstream targets is not uniform (Loog & Morgan, 2005 (as cited by Jin et al., 2008b)).

To relate this back to the synchronous approach, if AKT1 has in excess of 70 known cellular targets (Martelli et al., 2012) it is likely that AKT1 is able to interact with some with greater efficiency than others. Consequently, these substrates will be phosphorylated in a much shorter timescale than others. The synchronous approach does not take this into account when firing all enabled molecular interaction.

While the asynchronous approach does not include exact reaction rates, it does allow for some molecular interactions to occur before others. Gillespie algorithm simulations are used to create an ensemble of 1,000 individual trajectories, in which different orders of the firing of molecular interactions are tested. In this ensemble, there are examples of trajectories in which reactions occur in the correct order for biological behaviour to be feasible. In this way, non-parametric, qualitative simulations with the Gillespie algorithm are capable of generating biological behaviour of interest in signalling cascades. The qualitative synchronous approach is not capable of doing this. The only way to simulate signalling networks in a synchronous way would be to fully parameterise the model and run ODE simulations. Thus, examples in this dissertation show that the asynchronous simulations with the Gillespie algorithm can be used to qualitatively model signalling networks, while a synchronous approach is not.
Section 5.4 – Discussion

Section 5.4.1 – Potential Future Work

Given the reduction in predictive power that accompanied the integration of the RAW 264.7 cell GSMN, it would seem logical to attempt to increase the predictive power of the model through considerable increases in the level of cross-talk between the signalling pathways. A more detailed understanding of the nature of the interaction between signalling pathways, other than mTORC1, and the regulation of metabolism would also be of benefit. As noted in Chapter 4 – Modelling the Effects of Viral Infection, the preliminary work presented in this chapter is of importance as it provides a model which, when the predictive power is suitably raised, will allow the effects of viral infection upon metabolism to be modelled.

Section 5.4.2 – Concluding Remarks

Section 5.4.2.1 – Proposed Novel Modelling Strategies

Swanson and colleagues (2003) noted that one of the key requirements in mathematical modelling is the need for as complete an understanding as possible of the biological system in question. The ODE method, whilst initially promising, failed to meet this requirement. Even with a seemingly complete knowledge of the reaction kinetic parameters required to accurately model a biological system, the fundamental problems of failing to take into account all variables and not having parameters that translate well to other systems are always present. When referring to the models proposed by Tracqui et al. (1995) and Swanson et al. (2002) of gliomas, Hatzikirou and co-workers (2005) remark that the parameters used are only likely to refer explicitly to one system and may not be representative of all glioma systems. This example serves to illustrate that even with truly mathematical models it is difficult to represent a biological system in a way that is both
realistic and meaningful. The treatment of the qualitative, stochastic translational control model as an ODE model resulted in an obvious, albeit non-significant, decrease in predictive power. The ODE method used in this study resulted in a considerable decrease in computational cost but was only brought about at the expense of accuracy and verisimilitude. This being said, of the new ODE and synchronous methods, only the ODE method was worthy of carrying forward for a complete examination of the model.

Section 5.4.2.2 – Integration of the GSMN with the Translational Regulation Model

In addition to the trial of new modelling approaches, the integration of a GSMN with the model of the regulation of translation initiation represented a new direction for the project. Given that protein synthesis requires around one third of the total ATP produced by oxidative phosphorylation (Buttgereit & Brand, 1995), the ability to model both allowed the dynamic relationship between the two processes to be simulated. The work shown here establishes the plausibility of connecting a GSMN to a stochastic model. Although a significant reduction in the predictive power of the translation model was noted, the predictive power did remain positive. A possible explanation for this is that additional signalling pathways must be linked to the expression of genes found within the Bordbar et al. (2012) GSMN. As shown in Figure 5.3.1a in Section 5.3.1 – Current, Asynchronous, QSSPN Simulation Method, there is low correlation between the majority of the inhibitor models and model effectors involved in the linking of the translational regulation model and the GSMN.
Chapter 6 – Conclusion
Section 6.1 – Introduction

This chapter summarises the findings of the project as a whole and describes future uses of the model presented here. More detailed information on the findings can be found in the individual chapters.

The process of synthesising nascent polypeptides is a major energetic burden to the cell, requiring around 30% of all ATP produced (Buttgereit & Brand, 1995). The understanding of translation in mammalian cells is far from complete with much information being inferred from work with yeast. This being said, whilst the process in both systems is analogous there are differences. An illustration of the differences can be seen in the structure and size of the yeast and mammalian ribosomes (Morgan et al., 2000). Moreover, translation is involved in the pathophysiology of a number of significant human and veterinary diseases. Graff and co-workers (2008) even posit that the targeting of specific initiation factors may offer therapeutic approaches for diseases, such as cancer.

The signalling pathways that regulate translation are complex. The interactions that exist between such pathways add a further level of complexity to the cellular environment. Siso-Nadal and co-workers (2009) noted that in order for signals to be transduced in concert, the signalling pathways must form an interconnected network. Indeed, the connectivity within this network is the source of the robust nature of many of the dynamic behaviours that cells exhibit (Ibid) and functions to allow the integration of multiple signals (Hartwell et al., 1999). A consequence of this integration is to allow cells to demonstrate complex and coordinated patterns of behaviour (Ibid). With this in mind, any attempt to model a process, such as translation, which affects the functioning and indeed, viability of the cell must take into account the interconnected nature of the signalling pathways.
In Section 1.5.1 – Aims & Objectives, the aims of the project were listed. In all cases, these aims were achieved with various degrees of success and work investigating them was carried out. The main aim of this project was not to just computationally simulate the regulation of mammalian translation initiation by a large-scale reconstruction of a signalling network, but also to carry out these simulations alongside experimental validation of the behaviour of the model. In keeping with this view, it was possible to experimentally validate six predicted behaviours.

The overarching aim of this work was to produce a stochastic, qualitative model of the regulation of translation initiation by a signalling network. Over the course of this project, several accomplishments were necessary in order to achieve this aim:

- Analyse available literature of over 1,100 articles on translational regulation and construct large-scale molecular interactions network model
- Develop a literature-based benchmark dataset and a method to quantitatively analyse the predictive power of the model
- Experimental validation of the prediction involving an increase in the phosphorylation of eIF4E in response to Rapamycin
- Experimental validation of the prediction involving a decrease in the phosphorylation AKT1 in response to Rapamycin
- Provide a mechanism to explain the increase in eIF4E phosphorylation
- Incorporate the effects of viral infection into this model
- Experimental validation of the prediction involving an increase in CREB phosphorylation during MNV infection and three other behaviours
• Quantification of the model through the incorporation of a model of ERK1/2 activation

• Attempt to link metabolism to the regulation of translation

Section 6.2 – Model Construction & Predictive Power Evaluation

In order to construct the model it was necessary to collate information published in over 1,158 peer-reviewed articles. This model was then converted to the binary model formalism, which represents the activation conditions of each molecular interaction, without introducing unrealistic oscillations. The final model accounted for 584 interactions between 511 molecular species. This model differs from currently existing ones, such as that of Firczuk et al. (2013), in several regards. The yeast model presented by Firczuk and co-workers (2013) is considerably smaller than the one created here and does not include the regulatory network controlling translation initiation. This difference is important as it means that the model of regulating translation covers a much larger range of pathways. Consequently, this model has a much wider range of future implications.

The future development of this model should focus upon linking translational regulation to important and emerging pathways and processes. Furthermore, any continuation of this work should also look to incorporate new pathways that are found to regulate translation, for example the phosphorylation of eIF4B by Pim kinases (Yang et al., 2013b; Cen et al., 2014).

Given that cellular stress, and the formation of Stress Granules, have been shown to bring about considerable alterations to the levels of translation (Hoffman et al., 2012; Ruggieri et
(Qi et al., 2012), it seems that the incorporation of the process of forming Stress Granules should be considered. If the translational control model were to incorporate Stress Granule formation, it would better allow the phosphorylation of eIF2α to be modelled. This point of translational regulation has been one of the hardest to model as the exact nature of the input signal was hard to define and the upstream signalling pathways are very complex and not necessarily directly linked to the regulation of translation. Moreover, the biological relevance of eIF2α phosphorylation in RAW 264.7 cells is hard to define. Overexpression of Monocyte Chemotactic Protein-Induced Protein 1 has been shown to attenuate the phosphorylation of this initiation factor (Qi et al., 2011). Similarly, given the model is amenable to the simulation of viral infection, it would be necessary to also model strategies used by RNA viruses to limit the activation of PKR and the interferon response (Överby et al., 2010; Uchida et al., 2014).

The cell cycle represents the second process that would be of interest to link to translation. The process of translation has been noted to be a key regulatory point in the progression of the cell cycle in yeast (Daga & Jimenez, 1999), human (Göpfert et al., 2003) and Xenopus (Groisman et al., 2002) cells.

A further application for extending this work may come from combining the model of the regulation of translation initiation with models of elongation. Attempts at this may prove difficult given the qualitative nature of the model presented in this work. However, the stochastic framework used in the implementation of the Totally Asymmetric Exclusion Process models of elongation, such as that of Reuveni and co-workers (2011), may lead to a better understanding of how translation is regulated by a combination of extrinsic (in the
case of the signalling pathways) and intrinsic (for example at the level of the ribosome) factors.

The conversion to the binary model formalism eliminated the random fluctuations between the on and off states of effectors, such as kinases, in the model. Moreover, a binary model, in which effectors are either active or inactive, removes the pseudo-kinetic parameters that marking molecular species with multiple discrete activity levels introduces and, therefore, the model is governed solely by the known connectivity of the molecular network. Furthermore, the functionality of the model is reduced to simulating the conversion of sufficient model effectors that would bring about a biological effect. This binary model formalism represents a powerful, novel method of simulating large-scale biological systems.

One of the major outcomes of this work was the creation of a benchmark dataset and a method for measuring the predictive power of the model. This method concerns a comparison of the behaviour of the model with known effects. Existing literature has been examined to compile data on the effect of 12 chemical inhibitors on 23 molecular effectors. Using the Human Phospho-MAPK Arrays, it was possible to examine how 13 of these effectors respond to Rapamycin in RAW 264.7 cells. From this it was possible to provide a quantitative measure of predictive power by way of the MCC and the AC. Beck and co-workers (1997) argued that model validity is a determination as to whether the system can carry out a particular function in a consistent manner with the number of adverse outcomes being minimised. To this end, this model validation approach used here provides insight into how well existing data can be reproduced by the model and gives an indication as to how predictive the behaviour of the model is.
Based on the findings of this work that the MCC value is sensitive to small changes in the model, it is recommended that this method should be taken forward. However, this in silico predictive power analysis has a potential shortcoming. The method is only amenable to large systems or systems where sufficient data about model components is available. For such models, the number of data points is limited and therefore the resulting MCC value is susceptible to large changes being introduced after relatively small changes to model connectivity. It was shown in Section 2.2.1.1 – Updated Benchmark Dataset that the use of data only relating to small data subsets, such as specific cell types, can also lead to a marked reduction in data points. For this reason, the inclusion of data from all cell types was used. Although it could be argued that this would serve to reduce the overall value of the dataset against which to compare the model, the large size of this dataset means that variations are potentially reduced.

Although Fisher and co-workers (2013) employed the use of MCC values to measure the predictive power of the hepatocyte model, the use of this correlation coefficient in this project is novel in one key regard. In order to evaluate the functioning of the translational regulation model, no dataset against which to compare model behaviour was available. Fisher et al. (2013) were able to compare the behaviour of the hepatocyte model to the work of Song et al. (2009). Here 134 literature articles have been used to create a benchmark dataset of the effects of 12 chemical inhibitors on 23 model effectors. Apart from being applied for model validation, this dataset constitutes a comprehensive literature review that can be used by researchers when planning future work on the effects various inhibitors have upon key model signalling effectors.
The standardisation of nomenclature used in the model, with links to databases through the assignment of identifiers, adds to the quality of the project by providing an open and transparent development process. One of the major obstacles within the field of systems biology is the arbitrary naming of molecular species. This prevents the integration of models developed by different groups. Whilst the Systems Biology Markup Language has helped to standardise biochemical models (Hucka et al., 2003 (as cited by Bornstein et al., 2008)), the links to established databases, such as Uniprot and Entrez, will help to make the model more accessible to those outside of systems biology.

Section 6.3 – Experimental Validation

In an oversimplified way, systems biology aims to reconstruct biological systems and, using these models, to generate predictions (Ideker et al., 2001). The usefulness of the model constructed here was, therefore, demonstrated by the ability to generate seven predicted behaviours that were linked to translational machinery directly and also the regulatory signalling pathways. Furthermore, these behaviours were amenable to experimental validation. Moreover, as stated above, the model not only consistently reproduced the behaviour of inhibitors in multiple cell types but also displayed good agreement with experimental data collected in RAW 264.7 cells treated with Rapamycin using the Human Phospho-MAPK Arrays. Indeed this ability ‘to bring the theoretical predictions and experimental data into close apposition’ is one of the fundamental principles of systems biology (Ideker et al., 2001). Consequently, the experimental validation of six of the seven predicted behaviours overcame a major obstacle for this work.

The ability to demonstrate that six \textit{in silico} behaviours occur in a cell line adds significant value to the model as a whole. The work undertaken in this project has produced a model
of cellular translation that it has been experimentally validated. However, in regard to the 
behaviour of the MNV model, despite several predicted behaviours being experimentally 
validated, it was not possible to validate the behaviour of MSK2, in this model. This 
indicates that further refinement of the model upstream of MSK2 is required. 
Consequently, of the seven predicted behaviours that were experimentally tested, over the 
course of this project, it was possible to confirm six of these. This corresponds to a success 
rate of 85.7%.

The exact nature of eIF4E phosphorylation has only recently become known. Herdy et al. 
(2012) noted that eIF4E phosphorylation allows for a change in gene expression. This view 
is in keeping with the work presented here. Experimentally, Rapamycin-mediated mTORC1 
inhibition was shown in Chapter 3 – Experimental Validation of Model Predictions, to cause 
an increase in eIF4E phosphorylation. Given the role of mTORC1 in mediating cell growth 
(Dowling et al., 2010) and the finding that cellular stress signalling is inhibited by the 
mTORC1 substrate, S6K1 (Shin et al., 2013), it seems plausible that mTORC1 inhibition 
should bring about a stressed state in the cell and necessitate a change in gene expression.

Whilst it is a little harder to conclusively state the effect Rapamycin has upon the 
phosphorylation of AKT1, the work presented here does indicate that inhibition of this 
protein kinase occurs, possibly through mTORC2 inhibition, following one hour of treatment 
of cells with this macrolide. The work presented in Chapter 3 – Experimental Validation of 
Model Predictions, using Human Phospho-MAPK Arrays, clearly demonstrates that the level 
of phosphorylated AKT Pan is significantly reduced. This is in keeping with the work of Chen 
and co-workers (2010). These authors observed that high concentrations of Rapamycin 
served to reduce the phosphorylation of AKT1. Only with respect to treatment length does
the work demonstrated here offer some disagreement with Sarbassov et al. (2006). Both of these studies use treatment lengths of 24 hours. However, Sarbassov and co-workers (2006) noted that the effect of diminishing AKT1 phosphorylation was cell type specific. This specific effect correlates well with the work by Schreiber et al. (2015) in which the ratio of FKBP12 and FKBP51 was found to be important for determining the extent of Rapamycin-mediated inhibition of mTORC2. As noted in Chapter 3 – Experimental Validation of Model Predictions, determining this ratio in RAW 264.7 cells may prove useful. Despite the fact that immunoblotting with lysates of cells treated with Rapamycin for three hours failed to demonstrate that AKT1 phosphorylation was reduced, it is still possible to argue that this hypothesis has been, to some extent, experimentally demonstrated. Possible reasons for the inability to corroborate the Human Phospho-MAPK Array result is given in Section 3.4.2 – Validation of the AKT1 Phosphorylation Prediction.

Section 6.4 – Mechanism for Rapamycin-Induced Increases in eIF4E Phosphorylation

Whilst experimentally, Rapamycin has been shown to increase the level of eIF4E phosphorylation, the first version of the binary model predicted a decrease (as noted in Section 2.2.2 – Re-Evaluation of the Binary Model). This discrepancy motivated further work that has led to the improvement of the model and a better understanding of the nature of the relationship between mTORC1 and eIF4E. In order to identify a potential mechanism by which eIF4E phosphorylation can be increased by Rapamycin, a small model was created. This model only concerned the molecular interactions directly linked to mTORC1 activation and eIF4E phosphorylation. Interrogation of this small model revealed
the three points which would allow mTORC1 inhibition to positively affect eIF4E phosphorylation. These points, as noted in Section 2.2.3 – Mechanism by Which mTORC1 Inhibition Affects eIF4E Phosphorylation, detail factors which when taken together all contribute to eIF4E phosphorylation. These involve the relative levels of eIF4E and 4E-BP1, an alternative kinase for PKCα and potentially alterations to the activity of a protein phosphatase. Work to correct the direction of change in eIF4E phosphorylation was able to provide in silico evidence for two of these points. In the model eIF4G phosphorylation was disconnected from eIF4E phosphorylation and 4E-BP1 was maintained at a level lower than that of eIF4E. Together both changes permitted an increase in eIF4E phosphorylation that was dependent upon MNK2 and independent of MNK1. This is in agreement with the work by Stead and Proud (2013) in which MNK2 was implicated in mediating a Rapamycin-induced increase in eIF4E phosphorylation. Moreover, the simulations presented here suggest that this increase is not dependent upon either ERK1/2 or p38, again in agreement with earlier work (Stead & Proud, 2013; Eckerdt et al., 2014). Although numerous other studies have also noted that Rapamycin is capable of increasing eIF4E phosphorylation (Sun et al., 2005; Wang et al., 2007c; Grosso et al., 2011; Marzec et al., 2011), no real consensus was found between these studies as to how mTORC1 is related to the post-translational modification of this initiation factor.

Section 6.5 – Models of Viral Infection

In order to demonstrate the potential of the translational control model, perturbations normally associated with viral infection were implemented into the model. Additionally, to show the flexibility of the modelling approach, a range of viral effects were chosen. Two of the viruses (JUNV and ANDV) are severe, emerging human pathogens about which little is
known. The third virus, MNV, is a surrogate of the medically important HuNV. Although not permissive in RAW 264.7 cells, ANDV and JUNV are capable of inducing changes in the activation of different cellular signalling pathways. The changes that were made to the three models were all taken from existing literature or, in the case of MNV, from experimental data generated in-house.

QSSPN simulations and Reachfq analysis revealed that MNV infection induces considerable changes to the signalling network of the cell. As shown in Figure 6.5a, the outcome of these simulations led to the conclusion that MNV not only mediates changes in gene expression through the promotion of eIF4E phosphorylation (Royall et al., 2015), but this also occurs through alterations to the phosphorylation states of the transcription factors, CREB, ATF-2 and Elk-1. The experimental increase in CREB phosphorylation correlates well to the increase seen in the simulations presented in Chapter 4 – Modelling the Effects of Viral Infection. Despite this increase not being an unexpected finding, seeing as it is downstream of both p38 (Di Petro et al., 2007; Di Giacomo et al., 2009) and ERK1/2 (Impey et al., 1998; Xing et al., 1998), it was not a perturbation that was specifically induced in the model and thus, represents the experimental validation of one viral prediction.
Figure 6.5a – Schematic of the Effects of MNV-Induced ERK1/2 and p38 Activation. MNV induces the activation of ERK1/2 and p38α. Both the model and Human Phospho-MAPK Array data suggest that p38α activation is independent of MKK3 and MKK6. The work of Royall et al. (2015) noted that the increase in eIF4E phosphorylation is dependent upon p38α and MNK1. Experimentally, an increase in the phosphorylation of CREB has been demonstrated. Given that no significant change in MSK2 phosphorylation was noted in the array data, it is likely that the increase in CREB phosphorylation requires ERK1/2- (Deak et al., 1998; Sindreu et al., 2007; Brami-Cherrier et al., 2009) or p38α- (Aggeli et al., 2006; McCoy et al., 2007) dependent MSK1 phosphorylation (Ananieva et al., 2008). The stability of AU-Rich Element-containing mRNA is also hypothesised as being affected during MNV infection with TTP phosphorylation being increased in a MK2-dependent manner. Hypothesised links are denoted by broken blue arrows. The solid blue arrow is the proposed mechanism by which the increase in CREB phosphorylation may occur during MNV infection. Red arrows denote links where the model and array data are not in agreement or, in the case of broken, faded red arrows, where link are found in the literature but not believed to be important based upon the array data. Green arrows represent those reactions perturbed in the model.
In terms of the other viral models, the effects being simulated are of considerable importance to public health in South America as the pathogens triggering them cause severe and often fatal human infection. The incorporation of JUNV infection into the model produced considerable effects to the cellular signalling network. The cause of which appears to be connected to the feedback loops linking ERK1/2 to JNK1. Despite knowing that ERK1/2 activation is required for JUNV replication (Rodríguez et al., 2014), it is unclear whether this feedback loop is of importance during JUNV infection.

Whilst the model of ANDV infection produced only limited testable predictions, the results do indicate that metabolic reprogramming of the cell could occur through mTORC1 activation, as indicated by the increase in HIF-1α expression (shown in Section 4.3.3 – Simulation of the Andes Virus Models). The model presented here indicates that S6K1 and 4E-BP1 phosphorylation occurs during ANDV infection. This is in contrast with the work of McNulty et al. (2013) but somewhat in agreement with Gavrilovskaya et al. (2013). It is worth noting, however, that the latter publication used cells under hypoxic conditions but does show that S6K1 phosphorylation accompanies ANDV infection. The use of metabolomics approaches during ANDV infection may help to elucidate whether the activation of mTORC1 is linked to metabolism.

The work presented in Chapter 4 – Modelling the Effects of Viral Infection clearly demonstrates that the translational regulation model is amenable to modelling other systems. Further work must be undertaken to improve the functioning of the model integrated with the metabolic network. This is of considerable importance given the growing body of literature identifying the metabolic network as a target for viral infection (Vastag et al., 2011; Grady et al., 2013).
Despite the successful modelling of viral infection, this section of work suffers from a lack of suitable methods to quantify the predictive power, as noted in Section 4.4.1 – Problems Encountered & Troubleshooting. The method used in this work concerned validation of effects by comparing expected behaviours to the reality of simulations. Given the small number of known effects for these viruses, the use of the method used elsewhere in this project was not deemed appropriate. Future work looking at viral infection with such models should be aimed at developing a method, perhaps using the MCC values, for the quantification of predictive power. This method will become appropriate as more data is collected and particularly so if more data concerning the effects chemical inhibitors can have on the infection process.

Section 6.6 – Integration of the Translational Regulation Model with a GSMN

Whilst the complete integration of the ERK1/2 model did not provide any additional benefit to the model, linking mTORC1 signalling to the regulation of metabolism was tested. The use of FBA models is recognised as allowing the experimentally suggested alterations in metabolism to be recognised (Quek et al., 2014). Moreover, QSSPN permits the integration of such models of metabolism with Petri Net models (Fisher et al., 2013). Given the intimate relationship between cell growth and anabolic metabolism, linking mTORC1 to the control of metabolic gene expression seemed a logical way to bring about this integration. Several models of metabolism, derived from both the human (Thiele et al., 2013) and mouse (Sigurdsson et al., 2010) genomes, are available. Whilst of value, neither of these models was specific enough to use in a model that would ultimately be tested in RAW 264.7 cells.
23 genes found in the Bordbar et al. (2012) GSMN of a RAW 264.7 cell were placed, based upon existing literature, under control of mTORC1.

The linking of the GSMN of a RAW 264.7 cell allowed the relationship between translation and metabolism to be modelled. The production of Biomass in this GSMN was seen as a requirement for protein synthesis. Although the frequency of Biomass production was low, the lower bound of the Biomass reaction, as detailed in Bordbar et al. (2012), was set to the low value of 0.0281/h. Although this reaction rate is not directly comparable with the Reachfq analysis, this low frequency of Biomass production in the Wild Type model (with a value of 0.044) can in a way be seen as in keeping with the published data. Despite the integration being accomplished successfully, the result of this incorporation was a marked, albeit not significant, reduction in predictive power (MCC = 0.3376 versus MCC = 0.4558). Whilst no definitive reason can be given for this reduction, it could be argued that a more comprehensive integration of the Petri Net model, through additional signalling pathways, with the GSMN could serve to correct this. This argument is valid, as the comparison between the literature dataset and the simulation results demonstrated a somewhat less than ideal correlation for those effectors involved in linking of mTORC1 to the GSMN. This first attempt at linking metabolism to the control of translation highlighted the need for a more complete understanding of the dual regulation of translation and metabolism by the cellular signalling network.

**Section 6.7 – Novel Modelling Strategies**

The main aim of the work linking metabolism to the translational regulation model was to improve the overall quality of the model by allowing the interplay between metabolism and translation to be modelled. However a second aspect to the work involved the testing of
two novel modelling strategies. As detailed in Section 5.2.1 – Proposed New Approaches, these approaches aimed to reduce the computational cost of running simulations with little or no detrimental effects on predictive power. These methods, termed the synchronous simulation and the Ordinary Differential Equation Methods, were run in parallel with the asynchronous simulation method (outlined in Chapter 2 – Model Creation & Refining). The performance of the two novel methods further validated the method that had been used throughout the project and led to a better understanding of the need for the use of exact stochastic simulation methods when modelling biological systems.

An initial test of these approaches was carried out to ensure that the two experimentally tested behaviours of the model were conserved. The synchronous method was unable to model the Rapamycin-mediated increase in eIF4E phosphorylation. It was discovered that this approach introduces the false assumption that within a cell, all biochemical reactions that can occur will occur, in the same time interval. The removal of this temporal control of reactions, as noted in Section 5.3.2.2 – Synchronous Method, limits any further use of this method. It highlights the importance of the process of selecting the next reaction, which is based on probability and the Propensity Function (as discussed in Section 1.4.1 – Exact Stochastic Simulation with the Gillespie Algorithm).

The other method, the Ordinary Differential Equation Method, relies on converting the model to a series of qualitative ODEs, in which reaction kinetics are equal to one. Although it was possible to quantitatively analyse the predictive power of the model using the MCC and AC formulae, this method produced results that were entirely subjective and, consequently, the identification of regions of difference was open to observer bias. Perhaps
because of this subjective nature, the predictive power of the model analysed in this way was significantly reduced from MCC = 0.3376 to MCC = 0.2785.

Section 6.8 – General Conclusion

In summary, the overall aim of this project was to create a model simulating the regulation of mammalian translation initiation by the network of cellular signalling pathways. This work has produced such a model and validated the predictive power of this model against a unique benchmark dataset based on a comprehensive review of published literature. Statistically significant predictive power has been achieved and the model has produced experimentally testable hypotheses. Experimental work directed by these findings provided mechanistic insight into cross-talk between the mTORC1 and eIF4E phosphorylation. Moreover, this mechanism has provided some insight into the cellular stoichiometry of eIF4E and the interacting partner, 4E-BP1. This model of cellular translation has also been linked to the GSMN of a RAW 264.7 cell and given much needed insight into how such integration, with a large-scale model, can be achieved. Finally, the overall potential for future uses of this model has been demonstrated by successfully modelling the effects several medically important viruses can have upon the cellular signalling network and the translational machinery. Although much future work has been suggested throughout, the main output of this project has been to produce a model that can function within other larger models. The potential applications for this model are not limited to just viral infections. Areas such as cancer or cardiovascular biology may also benefit from a model of translation which, with adjustment, could be used to help guide research in these fields.


Amable, L., Grankvist, N., Largen, J.W., Ortsäter, H., Sjöholm, Å. & Honkanen, R.E. (2011) Disruption of serine/threonine Protein Phosphatase 5 (PP5:PPP5c) in Mice Reveals a Novel Role for PP5 in the Regulation of Ultraviolet Light-Induced Phosphorylation


Goedert, M., Cuenda, A., Craxton, M., Jakes, R. & Cohen, P. (1997) Activation of the Novel Stress-Activated Protein Kinase SAPK4 by Cytokines and Cellular Stresses is Mediated by SKK3 (MKK6); Comparison of its Substrate Specificity with that of Other SAP Kinases. The EMBO journal, 16, 3563-3571.


Phosphatidylinositol 3-Kinase- and Mammalian Target of Rapamycin-Dependent Pathways. The Journal of Biological Chemistry, 278, 28644-28650.


**Inoki, K., Li, Y., Xu, T. & Guan, L.-K. (2003)** Rheb GTPase is a Direct Target of TSC2 GAP Activity and Regulates mTOR Signaling. *Genes & Development*, 17, 1829-1834.


Peterson, R.T., Desai, B.N., Hardwick, J.S. & Schreiber, S.L. (1999) Protein Phosphatase 2A Interacts with the 70-kDa S6 Kinase and is Activated by Inhibition of FKBP12–Rapamycin


Shaw, L.M. (2001) Identification of Insulin Receptor Substrate 1 (IRS-1) and IRS-2 as Signaling Intermediates in the α6β4 Integrin-Dependent Activation of Phosphoinositide 3-OH Kinase and Promotion of Invasion. *Molecular and Cellular Biology, 21,* 5082-5093.


Inflammatory Stimuli is Faithfully Detected by Specific Antibodies. *Molecular Cell Biology Research Communications, 3,* 205-211.


Wang, X., Yue, P., Chan, C.B., Ye, K., Ueda, T., Watanabe-Fukunaga, R., Fukunaga, R., Fu, H., Khuri, F.R. & Sun, S.Y. (2007c) Inhibition of Mammalian Target of Rapamycin Induces


Appendices
The follow Supplementary Information is provided:

**Supplementary Information 1 - References for the Construction of the Literature Datasets**

This information was used when evaluating the effect of inhibitors within the model. This dataset was used throughout this work (Chapter 2 – Model Creation & Refining and Chapter 5 – Integration of the Translational Control Model with a Genome Scale Metabolic Network of a RAW 264.7 Cell)

**Supplementary Information 2 - Dataset against Which the Binary Model was Tested (Final Benchmark Dataset)**

The information contained within this table should be used in conjunction with the information contained with Supplementary Information 1. The information contained within this dataset was used to analyse the final version of the binary model.

**Supplementary Information 3 - Dataset Against Which the Model Integrated with Metabolism was Tested (Metabolic Benchmark Dataset)**

The information contained within this table should be used in conjunction with the information contained with Supplementary Information 1. The overall dataset is similar to that presented in Supplementary Information 3. To prevent repetition, only the effects linked to the extra effectors are shown.

**Supplementary Information 4 – Ingredients List for Solutions Produced In-House**
Supplementary Information E1 – Standardization of the Nomenclature of Model Components

This file works to standardize the naming of model components by linking all proteins, genes and miRNAs within the model to the established databases, Uniprot, Entrez and miRBase, respectively.

Supplementary Information E2 - List of Model Transitions & References in the Binary Model

This provides complete documentation of all model reactions with associated references.

Supplementary Information E3 – Final Version of the Binary Model

The .spept (E3a) and .xml (E3b) files of the Final version of the model. This model represents the Wild Type model against which the behaviour is tested. The predictive power of this model is given in Section 2.2.4 – Evaluation of the Predictive Power of the Final Model.

Supplementary Information E4 – MNV Model

The .spept (E4a) and .xml (E4b) files contain the model used as the mock infected MNV model against which the effects of MNV infection were tested. The results of this model are given in Section 4.3.2 – Simulation of the Murine Norovirus Models.
**Supplementary Information E5 – LPS Benchmark Dataset**

This .xlsx file contains the information relating to the LPS Benchmark and the effect of cell type on the level of predictive power.

**Supplementary Information E6 – Revised Original Benchmark Dataset**

This .xlsx file contains an updated version of the Initial Benchmark. The references contained within this benchmark do not use LPS stimulation.

**Supplementary Information E7 – Frequency Data Generated by Reachfq with Associated 99.99% Binomial Confidence Intervals**

This .xlsx file contains all frequency data generated by Reachfq and the associated 99.99% Binomial Confidence Intervals for model refinement and predictive power analysis.

The supplementary information E1-E7 can be found on the flash drive that accompanies this work. To open the .spept files for Supplementary Information E3 and E4, please ensure that Snoopy is installed on your computer. To download Snoopy, copy the following into your internet browser:

http://www-dssz.informatik.tu-cottbus.de/DSSZ/Software/Snoopy#downloads
Supplementary Information 1 -
References for the Construction of
the Literature Datasets
Supplementary Information 1a - Benchmark Dataset

References

Shown below is the reference list used in the construction of the benchmark datasets concerning the effects the inhibitors have in the absence of Lipopolysaccharide stimulation.

Note: Those references that are faded are those that were initially used but later found to have used LPS-stimulation

BIX02189 Mock Dataset


CGP57380 Mock Dataset


**GSK2334470 Mock Dataset**


**Ku0063794 Mock Dataset**


**LY294002 Mock Dataset**


PD184352 Mock Dataset


PD98059 Mock Dataset


Rapamycin Mock Dataset


**SB203580 Mock Dataset**


**SP600125 Mock Dataset**


Term Memory Formation and Retrieval of an Inhibitory Avoidance Task. *The European Journal of Neuroscience*, 17, 897-902.


Terminal Kinase Switches Smad3 Signaling from Oncogenesis to Tumor Suppression in Rat Hepatocellular Carcinoma. *Hepatology, 49*, 1944-1953.


**Torin-1 Mock Dataset**

The Torin-1 dataset was not used in this project but is included as a tool to be used in future work.


Li, J., Liu, J., Song, J., Wang, X., Weiss, H.L., Townsend Jr., C.M., Gao, T. & Evers, B.M. (2011) mTORC1 Inhibition Increases Neureotensin Secretion and Gene Expression through


XMD8-92 Mock Dataset


Supplementary Information 1b - Benchmark Dataset
References for the Binary Model with Lipopolysaccharide (LPS Benchmark)

The references shown below were used in the construction of the benchmark dataset concerning the effects each inhibitor has on Lipopolysaccharide stimulated cells.

Note: Those that are faded denote studies conducted with Macrophage cell lines. These are used in Section 2.2.2.1 – Updated Benchmark Dataset as the basis of the benchmark dataset

CGP57380 & LPS Mock Dataset


**LY294002 & LPS Mock Dataset**


**PD98059 & LPS Mock Dataset**


**SP600125 & LPS Mock Dataset**


391


Rapamycin & LPS Mock Dataset


5Z-7-Oxozeaenol & LPS Mock Dataset


Supplementary Information 2 – Dataset Against Which the Binary Model was Tested (Final Benchmark Dataset)
<table>
<thead>
<tr>
<th>Protein</th>
<th>Effect Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ppMEK1</td>
<td>No Effect (PMID: 18834865)</td>
</tr>
<tr>
<td>ppMEK2</td>
<td>No Effect (PMID: 18834865)</td>
</tr>
<tr>
<td>ppERK1</td>
<td>Partial Inhibition with Cadmium Chloride (PMID: 22521884); No Effect (PMID: 18834865)</td>
</tr>
<tr>
<td>ppERK2</td>
<td>Partial Inhibition with Cadmium Chloride (PMID: 22521884); No Effect (PMID: 18834865)</td>
</tr>
<tr>
<td>RSK-Pi</td>
<td>Decrease for RSK3 (PMID: 22997248) &amp; RSK2 &amp; RSK4 (PMID: 18834865)</td>
</tr>
<tr>
<td>rpS6-Pi</td>
<td>Reduction Implied But Not Directly Shown</td>
</tr>
<tr>
<td>pMNK1</td>
<td>Partial Inhibition (PMID: 18834865)</td>
</tr>
<tr>
<td>pMNK2</td>
<td>No Effect (PMID: 18834865)</td>
</tr>
<tr>
<td>eIF4E-Pi</td>
<td>Not Tested</td>
</tr>
<tr>
<td>pJNK1</td>
<td>No Effect (PMID: 18834865)</td>
</tr>
<tr>
<td>cJun-Pi</td>
<td>Not Tested But Potential Indirect Inhibition (PMID: 20075332; PMID: 12574153)</td>
</tr>
<tr>
<td>ppp38a</td>
<td>Partial Inhibition (PMID: 18834865)</td>
</tr>
<tr>
<td>pERK5</td>
<td>Inhibition Demonstrated (PMID: 18834865; PMID: 22254155; PMID: 24491810)</td>
</tr>
<tr>
<td>PI3K-Act</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PIP3</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PDK1-Act</td>
<td>No Effect (PMID: 18834865)</td>
</tr>
<tr>
<td>pPKCz</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppAKT</td>
<td>Partial Inhibition (PMID: 18834865)</td>
</tr>
<tr>
<td>ppS6K</td>
<td>Partial Increase (PMID: 18834865)</td>
</tr>
<tr>
<td>elf4B-Pi</td>
<td>Not Tested</td>
</tr>
<tr>
<td>p4E-BP1</td>
<td>Not Tested</td>
</tr>
<tr>
<td>mTORC1-Act</td>
<td>Not Tested</td>
</tr>
<tr>
<td>pTSC2</td>
<td>Not Tested</td>
</tr>
<tr>
<td>Protein</td>
<td>Effect</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>ppMEK1</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppMEK2</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppERK1</td>
<td>No Effect (PMID: 11463832) - Disputed (Partial Increase) (PMID: 18694961)</td>
</tr>
<tr>
<td>ppERK2</td>
<td>No Effect (PMID: 11463832) - Disputed (Partial Increase) (PMID: 18694961)</td>
</tr>
<tr>
<td>RSK-Pi</td>
<td>Possible Inhibition (PMID: 18694961) - Maybe Specific to CML</td>
</tr>
<tr>
<td>rpS6-Pi</td>
<td>Possible Inhibition (PMID: 18694961) - Disputed (PMID: 20664001)</td>
</tr>
<tr>
<td>pMNK1</td>
<td>Inhibition Demonstrated (PMID: 23269249)</td>
</tr>
<tr>
<td>pMNK2</td>
<td>Inhibition Demonstrated (PMID: 23269249)</td>
</tr>
<tr>
<td>elf4E-Pi</td>
<td>Inhibition Demonstrated (PMID: 23269249; PMID: 18081851)</td>
</tr>
<tr>
<td>pJNK1</td>
<td>No Effect (PMID: 11463832)</td>
</tr>
<tr>
<td>cJun-Pi</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppp38a</td>
<td>No Effect (PMID: 11463832)</td>
</tr>
<tr>
<td>pERK5</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PI3K-Act</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PIP3</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PDK1-Act</td>
<td>Not Tested</td>
</tr>
<tr>
<td>pPKCz</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppAKT</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppS6K</td>
<td>No Effect (PMID: 20664001)</td>
</tr>
<tr>
<td>elf4B-Pi</td>
<td>Not Tested</td>
</tr>
<tr>
<td>p4E-BP1</td>
<td>Partial Inhibition (PMID: 18694961)</td>
</tr>
<tr>
<td>mTORC1-Act</td>
<td>No Effect (PMID: 18694961)</td>
</tr>
<tr>
<td>pTSC2</td>
<td>Not Tested</td>
</tr>
<tr>
<td>GSK2334470</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>ppERK1</td>
<td>No Effect (PMID: 21087210)</td>
</tr>
<tr>
<td>ppERK2</td>
<td>No Effect (PMID: 21087210)</td>
</tr>
<tr>
<td>ppMEK1</td>
<td>No Effect (PMID: 21087210)</td>
</tr>
<tr>
<td>ppMEK2</td>
<td>No Effect (PMID: 21087210)</td>
</tr>
<tr>
<td>pRSK1</td>
<td>Partial Inhibition (PMID: 21087210)</td>
</tr>
<tr>
<td>rpS6-Pi</td>
<td>Inhibition Demonstrated (PMID: 21087210)</td>
</tr>
<tr>
<td>pMNK1</td>
<td>No Effect (PMID: 21087210)</td>
</tr>
<tr>
<td>pMNK2</td>
<td>No Effect (PMID: 21087210)</td>
</tr>
<tr>
<td>eIF4E-Pi</td>
<td>Not Tested</td>
</tr>
<tr>
<td>pJNK1</td>
<td>No Effect (JNK1 &amp; JNK3); Partial Activation (JNK2) (PMID: 21087210)</td>
</tr>
<tr>
<td>cJun-Pi</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppp38a</td>
<td>No Effect (PMID: 21087210)</td>
</tr>
<tr>
<td>pERKS</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PI3K-Act</td>
<td>No Effect (PMID: 21087210)</td>
</tr>
<tr>
<td>PIP3</td>
<td>No Effect (PMID: 21087210)</td>
</tr>
<tr>
<td>pPDK1</td>
<td>Inhibition Demonstrated (PMID: 21087210)</td>
</tr>
<tr>
<td>pPKCz</td>
<td>Partial Inhibition (PMID: 21087210)</td>
</tr>
<tr>
<td>ppAKT</td>
<td>No Effect (PMID: 21087210)</td>
</tr>
<tr>
<td>ppS6K1</td>
<td>No Effect (PMID: 21087210)</td>
</tr>
<tr>
<td>eIF4B-Pi</td>
<td>Not Tested</td>
</tr>
<tr>
<td>p4E-BP1</td>
<td>Not Tested</td>
</tr>
<tr>
<td>mTORC1-Act</td>
<td>Not Tested</td>
</tr>
<tr>
<td>pTSC2</td>
<td>Not Tested</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td><strong>LY294002</strong></td>
<td></td>
</tr>
<tr>
<td><strong>ppMEK1</strong></td>
<td>Partial Inhibition (PMID: 17850214)</td>
</tr>
<tr>
<td><strong>ppMEK2</strong></td>
<td>Partial Inhibition (PMID: 17850214)</td>
</tr>
<tr>
<td><strong>ppERK1</strong></td>
<td>Partial Inhibition (PMID: 17850214; PMID: 11157888); Disputed (No Effect) (PMID: 10480882; PMID: 12242656)</td>
</tr>
<tr>
<td><strong>ppERK2</strong></td>
<td>Partial Inhibition (PMID: 17850214; PMID: 11157888); Disputed (No Effect) (PMID: 10480882; PMID: 12242656)</td>
</tr>
<tr>
<td><strong>pRSK1</strong></td>
<td>No Effect (PMID: 17850214)</td>
</tr>
<tr>
<td><strong>rpS6-Pi</strong></td>
<td>Inhibition Demonstrated (PMID: 8895571; PMID: 11713299)</td>
</tr>
<tr>
<td><strong>pMNK1</strong></td>
<td>Partial Inhibition (PMID: 17850214)</td>
</tr>
<tr>
<td><strong>pMNK2</strong></td>
<td>Not Tested</td>
</tr>
<tr>
<td><strong>eIF4E-Pi</strong></td>
<td>Blocks Rapamycin Induced Increase (PMID: 16103051)</td>
</tr>
<tr>
<td><strong>pJNK1</strong></td>
<td>No Effect (JNK1 &amp; JNK2) Partial Inhibition (JNK3) (PMID: 17850214)</td>
</tr>
<tr>
<td><strong>cJun-Pi</strong></td>
<td>Not Tested</td>
</tr>
<tr>
<td><strong>pp38a</strong></td>
<td>No Effect (PMID: 11040049; PMID: 11367542; PMID: 17850214)</td>
</tr>
<tr>
<td><strong>pERKS</strong></td>
<td>No Effect Implied (PMID: 14670836)</td>
</tr>
<tr>
<td><strong>PI3K-Act</strong></td>
<td>Inhibition Demonstrated (PMID: 8106507)</td>
</tr>
<tr>
<td><strong>PIP3</strong></td>
<td>Inhibition Implied (PMID: 8106507)</td>
</tr>
<tr>
<td><strong>PDK1-Act</strong></td>
<td>Inhibition Demonstrated (PMID: 17850214; PMID: 9748166)</td>
</tr>
<tr>
<td><strong>pPKCz</strong></td>
<td>Inhibition Demonstrated (PMID: 9748166) Disputed (Activation) (PMID: 17850214)</td>
</tr>
<tr>
<td><strong>ppAKT</strong></td>
<td>Inhibition Demonstrated (PMID: 11801244; PMID: 16551362; PMID: 11755539)</td>
</tr>
<tr>
<td><strong>ppS6K</strong></td>
<td>Inhibition Demonstrated (PMID: 15677443; PMID: 10480882; PMID: 11755539; PMID: 11145615; PMID: 12242656; PMID: 9287347)</td>
</tr>
<tr>
<td><strong>eIF4B-Pi</strong></td>
<td>Not Tested</td>
</tr>
<tr>
<td><strong>p4E-BP1</strong></td>
<td>Inhibition Demonstrated (PMID: 11231341)</td>
</tr>
<tr>
<td><strong>mTORC1-Act</strong></td>
<td>Inhibition Demonstrated (PMID: 8895571; PMID: 19346248)</td>
</tr>
<tr>
<td><strong>pTSC2</strong></td>
<td>Not Tested</td>
</tr>
<tr>
<td>Protein</td>
<td>PD184352</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>ppMEK1</td>
<td>Inhibition Demonstrated (PMID: 10998351; PMID: 17850214; PMID: 12534346)</td>
</tr>
<tr>
<td>ppMEK2</td>
<td>Inhibition Demonstrated (PMID: 10998351; PMID: 17850214; PMID: 12534346)</td>
</tr>
<tr>
<td>ppERK1</td>
<td>Partial Inhibition (PMID: 12069688; PMID: 17850214)</td>
</tr>
<tr>
<td>ppERK2</td>
<td>Partial Inhibition (PMID: 12069688; PMID: 17850214)</td>
</tr>
<tr>
<td>pRSK1</td>
<td>Not Tested</td>
</tr>
<tr>
<td>rpS6-Pi</td>
<td>Partial Inhibition (PMID: 19376132)</td>
</tr>
<tr>
<td>pMNK1</td>
<td>Not Tested But Implied (PMID: 9155017)</td>
</tr>
<tr>
<td>pMNK2</td>
<td>Not Tested But Implied (PMID: 9155017)</td>
</tr>
<tr>
<td>eIF4E-Pi</td>
<td>Not Tested But Implied (PMID: 9155017)</td>
</tr>
<tr>
<td>pJNK1</td>
<td>Increase Demonstrated (PMID: 18347148)</td>
</tr>
<tr>
<td>cJun-Pi</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppp38a</td>
<td>No Effect (PMID: 22076433) But Disputed (PMID: 11431348) - Activation Reported (PMID: 18347148)</td>
</tr>
<tr>
<td>pERKS</td>
<td>No Effect (PMID: 12069688; PMID: 18358237)</td>
</tr>
<tr>
<td>P13K-Act</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PIP3</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PDK1-Act</td>
<td>Partial Inhibition (PMID: 10998351)</td>
</tr>
<tr>
<td>pPKcz</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppAKT</td>
<td>Partial Activation (PMID: 19147570; PMID: 22668349)</td>
</tr>
<tr>
<td>ppS6K</td>
<td>Partial Inhibition (PMID: 11431469)</td>
</tr>
<tr>
<td>eIF4B-Pi</td>
<td>Not Tested</td>
</tr>
<tr>
<td>p4E-BP1</td>
<td>Inhibition Demonstrated (PMID: 11799119)</td>
</tr>
<tr>
<td>mTORC1-Act</td>
<td>Inhibition Demonstrated (PMID: 21071439)</td>
</tr>
<tr>
<td>pTSC2</td>
<td>Partial Inhibition (PMID: 15292274)</td>
</tr>
<tr>
<td>Protein</td>
<td>Effect</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>ppMEK1</td>
<td>Partial Inhibition</td>
</tr>
<tr>
<td>ppMEK2</td>
<td>Partial Inhibition</td>
</tr>
<tr>
<td>ppERK1</td>
<td>Partial Inhibition</td>
</tr>
<tr>
<td>ppERK2</td>
<td>Partial Inhibition</td>
</tr>
<tr>
<td>pRSK1</td>
<td>Inhibition Demonstrated</td>
</tr>
<tr>
<td>rpS6-Pi</td>
<td>Inhibition Demonstrated</td>
</tr>
<tr>
<td>pMNK1</td>
<td>Inhibition Demonstrated</td>
</tr>
<tr>
<td>pMNK2</td>
<td>Inhibition Demonstrated</td>
</tr>
<tr>
<td>eIF4E-Pi</td>
<td>Inhibition Demonstrated</td>
</tr>
<tr>
<td>pJNK1</td>
<td>Activation Demonstrated</td>
</tr>
<tr>
<td>cJun-Pi</td>
<td>Partial Inhibition</td>
</tr>
<tr>
<td>ppp38a</td>
<td>No Effect</td>
</tr>
<tr>
<td>pERKS</td>
<td>Inhibition Demonstrated</td>
</tr>
<tr>
<td>PI3K-Act</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PIP3</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PDK1-Act</td>
<td>No Effect</td>
</tr>
<tr>
<td>pPKCz</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppAKT</td>
<td>No Effect</td>
</tr>
<tr>
<td>ppS6K</td>
<td>S6K2 Inhibited Degree Disputed</td>
</tr>
<tr>
<td>elf4B-Pi</td>
<td>Not Tested</td>
</tr>
<tr>
<td>p4E-BP1</td>
<td>Inhibition Demonstrated</td>
</tr>
<tr>
<td>mTORC1-Act</td>
<td>No Effect</td>
</tr>
<tr>
<td>pTSC2</td>
<td>Not Tested</td>
</tr>
<tr>
<td><strong>Rapamycin</strong></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td><strong>ppERK1</strong></td>
<td>Inhibition at Higher Doses (Activation at Lower Doses (PMID: 19151764)) (PMID: 20512842); No Effect (PMID: 23437362; PMID: 16160607; PMID: 11567647; PMID: 15767555; PMID: 12820963; PMID: 22343118)</td>
</tr>
<tr>
<td><strong>ppERK2</strong></td>
<td>Inhibition at Higher Doses (Activation at Lower Doses (PMID: 19151764)) (PMID: 20512842); No Effect (PMID: 23437362; PMID: 16160607; PMID: 11567647; PMID: 15767555; PMID: 12820963; PMID: 22343118)</td>
</tr>
<tr>
<td><strong>ppMEK1</strong></td>
<td>No Effect (PMID: 15767555)</td>
</tr>
<tr>
<td><strong>ppMEK2</strong></td>
<td>No Effect (PMID: 15767555)</td>
</tr>
<tr>
<td><strong>pRSK1</strong></td>
<td>Activation at Low Doses (PMID: 19151764)</td>
</tr>
<tr>
<td><strong>rpS6-Pi</strong></td>
<td>Inhibition Demonstrated (PMID: 23437362; PMID: 21075311; PMID: 16763566)</td>
</tr>
<tr>
<td><strong>pMNK1</strong></td>
<td>No Effect Indicated (PMID: 21406405)</td>
</tr>
<tr>
<td><strong>pMNK2</strong></td>
<td>Increase Demonstrated (PMID: 23831578)</td>
</tr>
<tr>
<td><strong>eIF4E-Pi</strong></td>
<td>Increase Demonstrated (PMID: 23831578; PMID: 16103051)</td>
</tr>
<tr>
<td><strong>pJNK1</strong></td>
<td>Activation Demonstrated (PMID: 21324487; PMID: 15767555); Inhibition Demonstrated (PMID: 9016789; PMID: 16160607)</td>
</tr>
<tr>
<td><strong>cJun-Pi</strong></td>
<td>Activation Demonstrated (PMID: 12820963)</td>
</tr>
<tr>
<td><strong>pp38a</strong></td>
<td>Inhibition Demonstrated (PMID: 16160607; PMID: 11567647); No Effect (PMID: 12820963; PMID: 9933636)</td>
</tr>
<tr>
<td><strong>pERK5</strong></td>
<td>Not Tested</td>
</tr>
<tr>
<td><strong>PI3K-Act</strong></td>
<td>No Effect on the Expression (PMID: 12766174)</td>
</tr>
<tr>
<td><strong>PIP3</strong></td>
<td>Not Tested</td>
</tr>
<tr>
<td><strong>Rapamycin</strong></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td><strong>pPDK1</strong></td>
<td>No Effect Indicated (PMID: 16377759) - Any Affect May be Substrate-Dependent (PMID: 21075311)</td>
</tr>
<tr>
<td><strong>pPKCz</strong></td>
<td>Not Tested Directly (PMID: 10339425)</td>
</tr>
<tr>
<td><strong>ppAKT</strong></td>
<td>Inhibition at Higher Doses (PMID: 20512842); Activation Possible at Lower Doses (PMID: 16103051; PMID: 23437362; PMID: 17698586; PMID: 21075311)</td>
</tr>
<tr>
<td><strong>ppS6K1</strong></td>
<td>Inhibition Demonstrated (PMID: 23437362; <strong>PMID: 21114628</strong>; PMID: 11567647; PMID: 15030312; PMID: 16763566; PMID: 15833867)</td>
</tr>
<tr>
<td><strong>eIF4B-Pi</strong></td>
<td>Increase Demonstrated at Lower Dose (PMID: 21075852); Prolonged Exposure - Inhibition Demonstrated (PMID: 16763566)</td>
</tr>
<tr>
<td><strong>p4E-BP1</strong></td>
<td>Partial Inhibition Demonstrated (PMID: 19402821; PMID: 17425689; PMID: 11799119)</td>
</tr>
<tr>
<td><strong>mTORC1-Act</strong></td>
<td>Inhibition Demonstrated (PMID: 15899889); <strong>No Effect Possible (PMID: 16103051)</strong></td>
</tr>
<tr>
<td><strong>pTSC2</strong></td>
<td>Not Tested</td>
</tr>
<tr>
<td>SB203580</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td><strong>ppMEK1</strong></td>
<td>No Effect (PMID: 10998351)</td>
</tr>
<tr>
<td><strong>ppMEK2</strong></td>
<td>No Effect (PMID: 10998351)</td>
</tr>
<tr>
<td><strong>ppERK1</strong></td>
<td>Activation Demonstrated (PMID: 10960075; PMID: 11297530; PMID: 16111636) Disputed (PMID: 17850214; PMID: 10998351)</td>
</tr>
<tr>
<td><strong>ppERK2</strong></td>
<td>Activation Demonstrated (PMID: 10960075; PMID: 11297530; PMID: 16111636) Disputed (PMID: 17850214; PMID: 10998351)</td>
</tr>
<tr>
<td><strong>pRSK1</strong></td>
<td>Partial Inhibition Demonstrated (PMID: 17906627)</td>
</tr>
<tr>
<td><strong>rpS6-Pi</strong></td>
<td>Partial Inhibition Demonstrated (PMID: 23315073)</td>
</tr>
<tr>
<td><strong>pMNK1</strong></td>
<td>Not Tested</td>
</tr>
<tr>
<td><strong>pMNK2</strong></td>
<td>No Effect (PMID: 11154262)</td>
</tr>
<tr>
<td><strong>eIF4E-Pi</strong></td>
<td>Partial Inhibition Demonstrated (PMID: 9545260; PMID: 12781867; PMID: 15075293)</td>
</tr>
<tr>
<td><strong>pJNK1</strong></td>
<td>Activation Demonstrated (PMID: 10960075) But Disputed (PMID: 9598985)</td>
</tr>
<tr>
<td><strong>cJun-Pi</strong></td>
<td>Activation Demonstrated (PMID: 10960075) But Disputed (PMID: 9598985)</td>
</tr>
<tr>
<td><strong>ppp38a</strong></td>
<td>Inhibition Demonstrated (PMID: 9753474; PMID: 18502741; PMID: 21903942)</td>
</tr>
<tr>
<td><strong>pERK5</strong></td>
<td>Not Tested</td>
</tr>
<tr>
<td><strong>PI3K-Act</strong></td>
<td>No Effect (PMID: 11367542)</td>
</tr>
<tr>
<td><strong>PI3P</strong></td>
<td>No Effect (PMID: 11367542)</td>
</tr>
<tr>
<td><strong>PDK1-Act</strong></td>
<td>Inhibition Demonstrated (PMID: 10702313)</td>
</tr>
<tr>
<td><strong>pPKCz</strong></td>
<td>Activation Demonstrated (PMID: 15665819)</td>
</tr>
<tr>
<td><strong>ppAKT</strong></td>
<td>Inhibition Demonstrated (PMID: 10702313; PMID: 11042204)</td>
</tr>
<tr>
<td><strong>ppS6K</strong></td>
<td>Partial Inhibition Demonstrated (PMID: 10455142; PMID: 9299480)</td>
</tr>
<tr>
<td><strong>eIF4B-Pi</strong></td>
<td>No Effect (PMID: 16763566)</td>
</tr>
<tr>
<td><strong>p4E-BP1</strong></td>
<td>Partial Inhibition Demonstrated (PMID: 23315073)</td>
</tr>
<tr>
<td><strong>mTORC1-Act</strong></td>
<td>Inhibition Demonstrated (PMID: 23315073)</td>
</tr>
<tr>
<td><strong>pTSC2</strong></td>
<td>Increase Demonstrated (PMID: 23315073)</td>
</tr>
<tr>
<td>Protein</td>
<td>Effect Description</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>ppMEK1</td>
<td>Partial Inhibition (PMID: 17850214)</td>
</tr>
<tr>
<td>ppMEK2</td>
<td>Partial Inhibition (PMID: 17850214)</td>
</tr>
<tr>
<td>ppERK1</td>
<td>Partial Inhibition - Disputed (PMID: 15454338; PMID: 20868664; PMID: 12603281; PMID: 12697808)</td>
</tr>
<tr>
<td>ppERK2</td>
<td>Partial Inhibition - Disputed (PMID: 15454338; PMID: 20868664; PMID: 12603281; PMID: 12697808)</td>
</tr>
<tr>
<td>pRSK1</td>
<td>Partial Increase (PMID: 16449644) Disputed (PMID: 12603281)</td>
</tr>
<tr>
<td>rpS6-Pi</td>
<td>Partial Inhibition (PMID: 21320501)</td>
</tr>
<tr>
<td>pMNK1</td>
<td>Partial Inhibition (PMID: 17850214)</td>
</tr>
<tr>
<td>pMNK2</td>
<td>Partial Inhibition (PMID: 17850214)</td>
</tr>
<tr>
<td>eIF4E-Pi</td>
<td>Partial Inhibition (PMID: 20359850; PMID: 20359507)</td>
</tr>
<tr>
<td>pJNK1</td>
<td>Inhibition Demonstrated (PMID: 17850214)</td>
</tr>
<tr>
<td>cJun-Pi</td>
<td>Inhibition Demonstrated (PMID: 20868664; PMID: 16449644; PMID: 16176902; PMID: 19305497)</td>
</tr>
<tr>
<td>ppp38a</td>
<td>No Effect (PMID: 16449644; PMID: 12603281; PMID: 17850214; PMID: 18359001); Very Early Increase Demonstrated (PMID: 12878189; PMID: 18359001)</td>
</tr>
<tr>
<td>pERK5</td>
<td>Partial Inhibition (PMID: 20643107)</td>
</tr>
<tr>
<td>PI3K-Act</td>
<td>Not Tested BUT Inhibition of δ Isoform of p110 Subunit - Subunit Specific (PMID: 19106158; PMID: 21558088)</td>
</tr>
<tr>
<td>PIP3</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PDK1-Act</td>
<td>Partial Inhibition (PMID: 17850214)</td>
</tr>
<tr>
<td>pPKCz</td>
<td>No Effect (PMID: 17850214)</td>
</tr>
<tr>
<td>ppAKT</td>
<td>No Effect (PMID: 17850214)</td>
</tr>
<tr>
<td>ppS6K</td>
<td>Partial Inhibition (PMID: 17850214; PMID: 16099428)</td>
</tr>
<tr>
<td>eIF4B-Pi</td>
<td>Partial Inhibition (PMID: 17850214; PMID: 21263197)</td>
</tr>
<tr>
<td>p4E-BP1</td>
<td>Partial Inhibition (PMID: 17850214 &amp; PMID: 21263197)</td>
</tr>
<tr>
<td>mTORC1-Act</td>
<td>Inhibition Demonstrated (PMID: 21320501; PMID: 22493283)</td>
</tr>
<tr>
<td>pTSC2</td>
<td>Not Tested</td>
</tr>
<tr>
<td>Protein</td>
<td>Description</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>ppMEK1</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppMEK2</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppERK1</td>
<td>No Effect (PMID: 20832753)</td>
</tr>
<tr>
<td>ppERK2</td>
<td>No Effect (PMID: 20832753)</td>
</tr>
<tr>
<td>pRSK1</td>
<td>Partial Inhibition (PMID: 20832753)</td>
</tr>
<tr>
<td>rpS6-Pi</td>
<td>Not Tested</td>
</tr>
<tr>
<td>pMNK1</td>
<td>Not Tested</td>
</tr>
<tr>
<td>pMNK2</td>
<td>Not Tested</td>
</tr>
<tr>
<td>eIF4E-Pi</td>
<td>Not Tested</td>
</tr>
<tr>
<td>pJNK1</td>
<td>Not Tested</td>
</tr>
<tr>
<td>cJun-Pi</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppp38a</td>
<td>Not Tested</td>
</tr>
<tr>
<td>pERK5</td>
<td>Inhibition Demonstrated (PMID: 20832753; PMID: 23428871)</td>
</tr>
<tr>
<td>PI3K-Act</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PIP3</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PDK1-Act</td>
<td>Not Tested</td>
</tr>
<tr>
<td>pPKCz</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppAKT</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppS6K</td>
<td>Not Tested</td>
</tr>
<tr>
<td>eIF4B-Pi</td>
<td>Not Tested</td>
</tr>
<tr>
<td>p4E-BP1</td>
<td>Not Tested</td>
</tr>
<tr>
<td>mTORC1-Act</td>
<td>Not Tested</td>
</tr>
<tr>
<td>pTSC2</td>
<td>Not Tested</td>
</tr>
</tbody>
</table>
Note the details of the effects of Ku0063794 are included as Ku0063794 was used in the initial model

<table>
<thead>
<tr>
<th>Protein</th>
<th>Effect</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ppMEK1</td>
<td>No Effect (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>ppMEK2</td>
<td>No Effect (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>ppERK1</td>
<td>No Effect (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>ppERK2</td>
<td>No Effect (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>pRSK1</td>
<td>No Effect (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>rpS6-Pi</td>
<td>Partial Inhibition (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>pMNK1</td>
<td>Partial Activation (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>MNK1/2-Pi</td>
<td>Partial Inhibition (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>eIF4E-Pi</td>
<td>Not Tested</td>
<td></td>
</tr>
<tr>
<td>pJNK</td>
<td>No Effect (JNK2α2) Partial Inhibition (JNK1α1) (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>cJun-Pi</td>
<td>Not Tested</td>
<td></td>
</tr>
<tr>
<td>ppp38a</td>
<td>No Effect (α-isoform) (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>pERK5</td>
<td>Not Tested</td>
<td></td>
</tr>
<tr>
<td>PI3K-Act</td>
<td>No Effect (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>PIP3</td>
<td>No Effect - Implied (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>PDK1-Act</td>
<td>Partial Inhibition (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>pPKCz</td>
<td>Partial Inhibition (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>ppAKT</td>
<td>No Effect (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>ppS6K</td>
<td>No Effect (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>eIF4B-Pi</td>
<td>Not Tested</td>
<td></td>
</tr>
<tr>
<td>P4E-BP1</td>
<td>Inhibition Demonstrated (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>mTORC1-Act</td>
<td>Inhibition Demonstrated (PMID: 19402821)</td>
<td></td>
</tr>
<tr>
<td>pTSC2</td>
<td>Not Tested</td>
<td></td>
</tr>
</tbody>
</table>
Torin-1, whilst not used in this project, is included as the list of effects is a tool that can be used for other works.

<table>
<thead>
<tr>
<th></th>
<th>Torin-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>ppERK1</td>
<td>Activation Demonstrated (PMID: 21508335; PMID: 21278786) - No Effect (PMID: 22343943; PMID: 21592956)</td>
</tr>
<tr>
<td>ppERK2</td>
<td>Activation Demonstrated (PMID: 21508335; PMID: 21278786) - No Effect (PMID: 22343943; PMID: 21592956)</td>
</tr>
<tr>
<td>ppMEK1</td>
<td>Not Tested</td>
</tr>
<tr>
<td>ppMEK2</td>
<td>Not Tested</td>
</tr>
<tr>
<td>pRSK1</td>
<td>Inhibition Demonstrated (At Higher Concentrations) (PMID: 22223645)</td>
</tr>
<tr>
<td>rpS6-Pi</td>
<td>Inhibition Demonstrated (PMID: 24424027; PMID: 22552098; PMID: 24008870; PMID: 21908613; PMID: 22065737; PMID: 23415771; PMID: 21576368; PMID: 23319332)</td>
</tr>
<tr>
<td>pMNK1</td>
<td>No Effect (PMID: 22223645)</td>
</tr>
<tr>
<td>pMNK2</td>
<td>No Effect (PMID: 22223645)</td>
</tr>
<tr>
<td>eIF4E-Pi</td>
<td>Inhibition Demonstrated (PMID: 20679199)</td>
</tr>
<tr>
<td>pJNK1</td>
<td>No Effect (PMID: 22223645)</td>
</tr>
<tr>
<td>cJun-Pi</td>
<td>Activation Demonstrated (PMID: 21508335)</td>
</tr>
<tr>
<td>ppp38a</td>
<td>No Effect (PMID: 21592956)</td>
</tr>
<tr>
<td>pERK5</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PI3K-Act</td>
<td>No Effect (PMID: 22015718; PMID: 23436801) - But Activation Maybe Possible (PMID: 22015718)</td>
</tr>
<tr>
<td>PIP3</td>
<td>Not Tested</td>
</tr>
<tr>
<td>Torin-1</td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>---</td>
</tr>
<tr>
<td>pPDK1</td>
<td>Not Tested</td>
</tr>
<tr>
<td>pPKCz</td>
<td>Inhibition Demonstrated (PMID: 24375676) - No Effect (PMID: 22223645)</td>
</tr>
<tr>
<td>ppAKT</td>
<td>Inhibition Demonstrated (PMID: 21508335; PMID: 21659604; PMID: 22223645; PMID: 22733130; PMID: 20089925; PMID: 23142081; PMID: 24311379; PMID: 20735411; PMID: 22006022; PMID: 24043828)</td>
</tr>
<tr>
<td>ppS6K1</td>
<td>Inhibition Demonstrated (PMID: 22552098; PMID: 21508335; PMID: 21659604; PMID: 22343943; PMID: 22223645; PMID: 21858812; PMID: 23707523; PMID: 23524850; PMID: 23888043)</td>
</tr>
<tr>
<td>eIF4B-Pi</td>
<td>Not Tested</td>
</tr>
<tr>
<td>p4E-BP1</td>
<td>Inhibition Demonstrated (PMID: 21508335; PMID: 21659604; PMID: 23547259; PMID: 23077579; PMID: 21858812; PMID: 23707523; PMID: 21841310; PMID: 21307192)</td>
</tr>
<tr>
<td>mTORC1 - Act</td>
<td>Inhibition Demonstrated (PMID: 19150980; PMID: 23934889; PMID: 23401004; PMID: 22053105; PMID: 21576368; PMID: 23963679)</td>
</tr>
<tr>
<td>pTSC2</td>
<td>Not Tested</td>
</tr>
</tbody>
</table>
Supplementary Information 3 – Dataset Against Which the Model Integrated with Metabolism was Tested (Metabolic Dataset)
<table>
<thead>
<tr>
<th></th>
<th>BIX02189</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-Myc</td>
<td>Not Tested</td>
</tr>
<tr>
<td>HIF-1α</td>
<td>Not Tested with the Inhibitor - ERK5 Maybe Important though (PMID: 16735500)</td>
</tr>
<tr>
<td>Scd1</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PFK</td>
<td>Not Tested</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>CGP57380</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-Myc</td>
<td>No Effect (PMID: 22370634)</td>
</tr>
<tr>
<td>HIF-1α</td>
<td>No Effect (PMID: 25115400)</td>
</tr>
<tr>
<td>Scd1</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PFK</td>
<td>Not Tested</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>GSK2334470</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-Myc</td>
<td>Not Tested</td>
</tr>
<tr>
<td>HIF-1α</td>
<td>Not Tested</td>
</tr>
<tr>
<td>Scd1</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PFK</td>
<td>Not Tested</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>LY294002</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-Myc</td>
<td>Inhibition Demonstrated (PMID: 22546857; PMID: 15467756; PMID: 25636967)</td>
</tr>
<tr>
<td>HIF-1α</td>
<td>Inhibition Demonstrated (PMID: 16849522; PMID: 12032158; PMID: 10749120; PMID: 14982927; PMID: 12070140; PMID: 11457733; PMID: 20194722; PMID: 15714461; PMID: 22787058)</td>
</tr>
<tr>
<td>Scd1</td>
<td>Inhibition Demonstrated (PMID: 20876744)</td>
</tr>
<tr>
<td>PFK</td>
<td>Inhibition Demonstrated (PMID: 16051672) - No Effect on Level (PMID: 20958264) - Activity May be Inhibited (PMID: 17302559; PMID: 20958264)</td>
</tr>
<tr>
<td></td>
<td>PD184352</td>
</tr>
<tr>
<td>---</td>
<td>----------</td>
</tr>
<tr>
<td>c-Myc</td>
<td>No Effect on mRNA Level (PMID: 15286700) - Inhibition Demonstrated (Inhibitor Not Used Alone) (PMID: 18497825)</td>
</tr>
<tr>
<td>HIF-1a</td>
<td>Inhibition Demonstrated (PMID: 17213817) - No Effect (PMID: 18199551)</td>
</tr>
<tr>
<td>SCD1</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PFK</td>
<td>Not Tested</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Rapamycin</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-Myc</td>
<td>Inhibition Demonstrated (PMID: 14576155; PMID: 18621930; PMID: 23612979; PMID: 15634685; PMID: 10993886; PMID: 25636967)</td>
</tr>
<tr>
<td>HIF-1α</td>
<td>Inhibition Demonstrated (PMID: 20670887; PMID: 21567203; PMID: 18781596; PMID: 17968710; PMID: 17502379; PMID: 12032158; PMID: 12242281; PMID: 10749120; PMID: 22900063) - No Effect (PMID: 16849522; PMID: 12149254; PMID: 20194722; PMID: 22787058)</td>
</tr>
<tr>
<td>Scd1</td>
<td>Inhibition Demonstrated (PMID: 20670887; PMID: 20876744)</td>
</tr>
<tr>
<td>PFK</td>
<td>Inhibition Demonstrated (PMID: 20670887)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SB203580</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-Myc</td>
<td>Inhibition Demonstrated (PMID: 20212154; PMID: 16365184; PMID: 10648601; PMID: 12080469; PMID: 11563982)</td>
</tr>
<tr>
<td>HIF-1α</td>
<td>Inhibition Demonstrated (PMID: 20194722; PMID: 11978547; PMID: 12482858) - No Effect (PMID: 15714461)</td>
</tr>
<tr>
<td>SCD1</td>
<td>No Effect (with 5-Fluorouracil Pretreatment) (PMID: 24135379)</td>
</tr>
<tr>
<td>PFK</td>
<td>Not Tested</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SP600125</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-Myc</td>
<td>Inhibition Demonstrated (PMID: 15890690; PMID: 20335523; PMID: 21525782; PMID: 25486532; PMID: 19418558; PMID: 24104553; PMID: 19996270)</td>
</tr>
<tr>
<td>HIF-1α</td>
<td>No Effect (PMID: 20194722)</td>
</tr>
<tr>
<td>SCD1</td>
<td>Increase Demonstrated (in Trans-10, cis-12 conjugated linoleic acid Cotreatment) (PMID: 21744278)</td>
</tr>
<tr>
<td>PFK</td>
<td>Not Tested</td>
</tr>
<tr>
<td>Protein</td>
<td>Activity/Status</td>
</tr>
<tr>
<td>---------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>c-Myc</td>
<td>Inhibition Demonstrated (PMID: 24880079)</td>
</tr>
<tr>
<td>HIF-1α</td>
<td>Not Tested</td>
</tr>
<tr>
<td>Scd1</td>
<td>Not Tested</td>
</tr>
<tr>
<td>PFK</td>
<td>Not Tested</td>
</tr>
</tbody>
</table>
Supplementary Information 4 – Ingredients List for Solutions Produced In-House
Running Buffer

14.4gm Glycine (Thermo Fisher Scientific Inc., Rockford, IL (USA))
3.0gm Tris-Base (Thermo Fisher Scientific Inc., Rockford, IL (USA))
5ml 20% SDS (Thermo Fisher Scientific Inc., Rockford, IL (USA))
Made up to 1l with Milli-Q® Ultrapure (EMD Millipore Corp., Merck KGaA, Darmstadt (DE))

Transfer Buffer

14.4gm Glycine (Thermo Fisher Scientific Inc., Rockford, IL (USA))
3.0gm Tris-Base (Thermo Fisher Scientific Inc., Rockford, IL (USA))
200ml Methanol (Sigma-Aldrich Co. LLC, St. Louis, MO (USA))
Made up to 1l with Milli-Q® Ultrapure (EMD Millipore Corp., Merck KGaA, Darmstadt (DE))

1x PBS-Tween Solution

10 Dulbecco A/PBS Tablets (Oxoid Ltd., Thermo Fisher Scientific Inc., Rockford, IL (USA))
1ml Tween-20 (Sigma-Aldrich Co. LLC, St. Louis, MO (USA))
Made up to 1l with Milli-Q® Ultrapure (EMD Millipore Corp., Merck KGaA, Darmstadt (DE))