ON THE USE OF PROCESS ALGEBRA TECHNIQUES IN COMPUTATIONAL MODELLING OF CANCER INITIATION AND DEVELOPMENT

by Oksana Schaeffer

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Computer Science
The University of Birmingham
February 2008

Abstract

Cancer research has been revolutionised by recent technological advances that allow scientists to produce extensive collections of experimental data, especially on the molecular and cellular level. Formal modelling is a necessary tool for integrating massive amounts of diverse measurement data into a coherent picture of disease development. Models can be used to test hypotheses about the role of cellular components in system function and in creating disease, and to make predictions which can then be tested experimentally.

This thesis evaluates process algebra techniques as description formalisms for a collection of cancer-related models. Process algebras view biology as a dynamic interactive communication network, in which an individual agent is performing a computation corresponding to the reaction. Agents typically represent entities such as molecules or cells. The stochastic extensions of process algebras allow the modeller to assign probability (or rate) to every reaction. The analysis of the resulting models is usually based on stochastic simulation. Alternatively, formal verification tools can be used to calculate exact quantitative properties of the underlying stochastic process.

We have explored the applicability of the process algebra formalism by analysing the dynamics of two cancer-related signalling pathways: Wnt/Wingless and FGF (Fibroblast Growth Factor). In addition to process algebra models, we have also derived continuous differential equation models for comparison. Systematic analysis of parameter spaces has revealed which variables have the most influence on temporal and steady state properties

of the system. By integrating feedback mechanisms, amplification factors, and different time scales we have demonstrated a resulting emergence of several unexpected properties of system dynamics. We were later able to confirm these by *in vitro* experiments for both pathways.

To examine the function of the specific signalling architecture in the cellular decision making process, we have constructed a model that couples Wnt signalling to the decision process within the cell and cell microenvironment. The model reveals signalling characteristics that ensure accuracy and robustness of Wnt-mediated determination of proliferative cell fate and lead to tissue architecture which is resistant to mutations. The main contribution of this thesis is, therefore, to systems biology; we have produced reusable and validated quantified models and demonstrated their value in designing, testing, and refining hypotheses about cancer.

Acknowledgements

I would like to express my deep gratitude to my supervisor Prof. Marta Kwiatkowska for her enthusiastic guidance in this work. I thank Prof. John Heath for greatly enriching the biological side of this study, for his data and general advice. Many thanks go to Dr. Glenn Matthews and Dr. Germaine Caldwell for their help with Wnt experiments and relevant discussions. I also acknowledge the support of the Integrative Biology project (EPSRC GR/S720231/01) and Microsoft Research Cambridge contract MRL 2005-44.

I am grateful to all my fellow graduate students for making my time in Birmingham more enjoyable. Finally, I give my special thanks to my family, particularly to my parents and most importantly to Alex, for love and support they have given me over these years.

Contents

1 Introduction **1**
- 1.1 Contributions . . . 3
- 1.2 Publications . . . 4
- 1.3 Outline . . . 7

2 Signalling networks **9**
- 2.1 Introduction . . . 9
  - 2.1.1 Computational models of signalling systems . . . 9
- 2.2 Modelling frameworks . . . 11
  - 2.2.1 Continuous deterministic approach . . . 13
  - 2.2.2 Discrete stochastic approach . . . 15
  - 2.2.3 Modelling signalling dynamics in π-calculus . . . 17
  - 2.2.4 Signalling networks in PRISM . . . 23
- 2.3 Analysis techniques . . . 28
  - 2.3.1 Simulation . . . 28
  - 2.3.2 Sensitivity analysis . . . 28
  - 2.3.3 Changes in the trajectory . . . 30
  - 2.3.4 Experimental validation . . . 30
- 2.4 Examples of signalling mechanisms . . . 32
  - 2.4.1 Receptor module . . . 32
  - 2.4.2 Positive feedback module . . . 39
  - 2.4.3 Negative feedback module . . . 46
- 2.5 Conclusions . . . 52

3 FGF signalling pathway **54**
- 3.1 Introduction . . . 54
  - 3.1.1 FGF biology . . . 54
- 3.2 Related work . . . 56
- 3.3 FGF model design . . . 57
  - 3.3.1 Components and interactions of the FGF pathway . . . 57
  - 3.3.2 Choice of parameter values . . . 61
  - 3.3.3 π-calculus model of FGF signalling . . . 61

3.3.4 ODE model of the FGF pathway . . . 65
3.4 Biological implications of the model . . . 65
3.4.1 In silico mutagenesis predicts the roles of pathway components . . . 67
3.4.2 Dual role of Src in signal regulation: from model to experiments and back . . . 68
3.4.3 Cross-regulation between receptor activation and attenuation . . . 71
3.5 Conclusions . . . 73

4 Wnt signalling **76**
4.1 Introduction . . . 76
4.1.1 Wnt biology . . . 76
4.2 Related models . . . 78
4.3 Model design . . . 79
4.3.1 Reactions included in the model . . . 79
4.3.2 Wnt model in π-calculus . . . 80
4.3.3 Wnt model in ODE . . . 85
4.4 Model predictions . . . 90
4.4.1 Two dynamic oscillatory regimes . . . 90
4.4.2 Experimental validation . . . 94
4.4.3 Inherent robustness and adaptability . . . 97
4.5 Biological significance of Wnt oscillations . . . 99
4.6 Conclusions . . . 101

5 Cellular decisions **102**
5.1 Introduction . . . 102
5.1.1 Crypt biology . . . 103
5.2 Related work . . . 104
5.3 Extending π-calculus to model cells . . . 105
5.3.1 Cells as mobile ambients . . . 105
5.3.2 Cell division . . . 107
5.3.3 Modelling space and cell movement . . . 107
5.4 A model of intra- and inter-cellular dynamics of the crypt . . . 110
5.4.1 Intracellular signal transduction network . . . 111
5.4.2 Proliferative and differentiated cell fate . . . 111
5.4.3 Wnt gradient in the tissue . . . 113
5.5 Robust cell fate determination by Wnt signalling . . . 115
5.6 Conclusions . . . 117

6 Conclusions and future work **119**

Bibliography **122**

Appendices **132**

A Implementation of elementary signalling modules **132**
A.1 Receptor module 132
A.2 Positive feedback module 133
A.3 Negative feedback module 134

B Modelling FGF pathway **135**

C Implementation of Wnt pathway model **141**

D Multi-scale model of cellular decisions **146**

E Glossary **151**

List of Figures

2.1 Basic stochastic π-calculus operations: (a) communication, (b) choice, (c) parallel composition, and (d) scope restriction primitives. 20
2.2 Biochemical reactions in PRISM 25
2.3 Implementation of the receptor module in π-calculus. 33
2.4 Temporal evolution of the receptor subsystem: a single stochastic run (a and b), an average of 100 stochastic runs (c and d) and deterministic solution (e and f) for $\#S(0) = 0.1 * \Omega$ and $\#S(0) = \Omega$, respectively. 35
2.5 Receptor module in PRISM 36
2.6 Properties of the receptor module: (a) The probability distribution of the number of active receptor molecules at time $t = 10$ min; (b) The expected number of bindings per stimulus molecule depending on the stimulus strength. 38
2.7 Linear positive feedback in π-calculus 40
2.8 General form of feedback in PRISM 41
2.9 Temporal evolution of the linear feedback system: single stochastic run (a and b), average of 100 stochastic runs (c and d) and deterministic solution (e and f) for $\#S(0) = 0$ and $\#S(0) = 0.1 * \Omega$. 42
2.10 Positive feedback and bistability: results of individual stochastic runs (a and b), average of 100 runs (c and d), and deterministic solution (e and f) for $\#S(0) = 0$ and $\#S(0) = 0.1 * \Omega$, respectively. 44
2.11 Probability distribution and the number of synthesised receptors in the positive feedback model. 45
2.12 Negative feedback module implementation in π-calculus. 47
2.13 Negative feedback and oscillations: results of the π-calculus (a, c, and e) and the ODE (b, d, and f) models at $\#S(0) = 0.01 * \Omega$, $\#S(0) = 0.1 * \Omega$ and $\#S(0) = \Omega$, respectively. 49
2.14 Signalling amplitude and duration depending on the value of incoming stimulus S for deterministic (a and c) and stochastic (b and d) systems, respectively. 51
2.15 Variability of oscillations: (a and b) the number of molecules B, (c and d) the corresponding limit cycle and (e and f) the histogram of the cycle periods for $\Omega = 100$ and $\Omega = 1000$, respectively. 53

3.1 Diagram of the possible bindings of the FGF receptor. 55

3.2 Schematic representation of molecular interactions in the FGF pathway . . 58
3.3 Representation of FRS2 molecule in π-calculus. 63
3.4 Src-induced receptor recruitment in π-calculus. 64
3.5 Varying FGFR kinase activity and FGF levels: (a) 10- and 100-fold decrease of FGFR kinase rate compared with the default rate; (b) 100%, 10% and 5% of the FGF level. 67
3.6 In silico mutagenesis: (a) simulations with inhibited Spry, Shp2, and Src compared to the model containing all these components; (b) inhibition of Src in the revised model; (c and d) experimental confirmation of predictions with inhibited Src (experiment was not repeated, see text for explanation). 69
3.7 Effects of FGFR mutations: simulations with 10-fold inhibition of FGF:FGFR disassociation rate compared to the normal rate in the initial (a) and revised (b) model; (c and d) experimental validation of the phenotypic changes in mutant. Experiments were repeated twice (thus no error bars) with consistent results. 72

4.1 Schematic description of molecular reactions of the Wnt pathway 81
4.2 Representation of Axin and APC molecules in π-calculus. 84
4.3 Oscillations in β-catenin concentration predicted by deterministic (a) and stochastic (b) versions of the model. (c) shows a schematic phase portrait of the deterministic oscillator. 91
4.4 Time evolution of β-catenin concentration in deterministic (a) and stochastic (b) systems. (c) shows a phase portrait of the deterministic systems that falls into a stable steady state ($\bullet$) and a noise-induced excursion through a phase space of the stochastic system. 93
4.5 Experimental validation of the model predictions: real-time analysis of Axin2 (a) and NKD1 (c) expression in HeLa cells following treatment with Wnt3A conditioned medium; periodogram computed for Axin2 (b) and NKD1 (d) time series. 95

5.1 Ambient capabilities 106
5.2 Communication directions between ambients 107
5.3 Cell division in π-calculus. The probability of reassigning a molecule to a daughter cell converges to 1/2, if the rates of channels `move` and `die` are equal. 108
5.4 Cell organisation: (a) linear array of cells referencing upper neighbour; (b) lattice representation in π-calculus. 110
5.5 Adaptation of the Wnt subsystem in a multi-scale model. 112
5.6 Stem cell evolution: a cell undergoes proliferation, differentiation, or death. Additionally, the cell is constantly receiving information about the environment, and adjusts its position within the spatial lattice to accommodate newly born cells. 114

5.7 Cell fate control by the Wnt pathway: (a) model predictions of the proliferative cell distribution in both healthy and mutant tissues agree well with (b) the experimental data. 117

List of Tables

2.1 Operational semantics of π-calculus. 19
2.2 Sensitivity coefficients of the receptor module calculated for the amplitude *Amp* and duration *Dur* of the response component A^*. 39
2.3 Sensitivity coefficients of the positive feedback model determined for the amplitude *Amp* and duration *Dur* of A^*. 46
2.4 Sensitivity coefficients of the negative feedback system, determined for for the amplitude *Amp* and duration *Dur* of B. 52

3.1 Kinetic parameter values of the FGF pathway 74
3.2 Sensitivity coefficients for the amplitude *Amp* and duration *Dur* of the signalling response component $FRS2 : Grb2$. **Bold** indicates sensitive parameters (in this particular example, sensitivity threshold of 0.05 is chosen). Values smaller than 10^{-5} are annotated as $\sim$0. 75

4.1 Kinetic parameter values of the Wnt pathway. Superscript$^{\text{e}}$ indicates that values were determined based on time courses presented in respective papers. Superscript$^{\text{u}}$ indicates that we remain uncertain about respective values because the measurements were not reproduced by independent studies. 82
4.2 Sensitivity coefficients that characterise the average value, frequency, and amplitude of β-catenin. **Bold** indicates highly sensitive parameters (in this particular example, sensitivity threshold of 0.12 is chosen). 98

Chapter 1

Introduction

For decades, cancer research has been focused on the identification of cellular parts, genes and proteins that become disregulated during cancer development. To a large extent, this direction was chosen based on the assumption that there is a gene for everything: a gene that causes cancer and a gene that causes heart disease. This is no longer tenable for complex organisms; what has been discovered instead is that each gene plays a role in multiple functions. Likewise, each function arises from the cooperation of many genes. Interactions of cellular parts ensure that the system exhibits robustness and fault tolerance in the presence of randomly occurring perturbations. Understanding of tumourigenesis and other biological phenomena would benefit greatly from a view of the cell as an interactive system where components work together in networks, continuously adjusting to the information they receive about the internal cell state and cellular environment. While an understanding of genes and proteins continues to be important, research focus has to be shifted to include higher levels of abstraction, involving pathways and networks.

This thesis describes the application of computational tools to understanding properties of cancer using systems approaches. By a systems approach to biology we mean gathering and integrating diverse measurement data about molecular components and interactions

taking place in a living cell. These data are incorporated into models which are then used to test specific hypotheses about component roles in responding to stimuli and in creating disease. Moreover, models are used to make biologically interesting predictions which are then tested and confirmed experimentally. A systems based approach is thus shown to give a novel understanding of what goes wrong in disease at a molecular level and how to fix the aberrant molecular network.

This thesis evaluates process algebra techniques as description formalisms for a collection of cancer-related models. Process algebras view biology as a dynamic interactive communication network, in which an individual agent is performing a computation corresponding to the reaction. Such descriptions of biological networks induce discrete models whose states are vectors of agents and transitions correspond to reactions. Agents typically represent entities such as molecules or cells. Agents undergo state changes as a result of consecutive, parallel, competitive, compartment- or context-dependent reactions. The stochastic extensions of process algebras allow the modeller to assign probability (or rate) to every reaction. The analysis of the resulting models is usually based on Monte Carlo numerical simulation, such as stochastic simulation using the Gillespie algorithm. Alternatively, formal verification tools can be used to calculate exact quantitative properties of the underlying stochastic process, for example, the probability that a particular complex is formed.

Using the process algebra approach we study the dynamics of two cancer-related signalling pathways: Wnt/Wingless and FGF (Fibroblast Growth Factor). We have chosen these pathways because, even though their components and interactions are relatively well understood, the knowledge of system dynamics and how it changes as cancer progresses is still lacking. Pathway models are explored to discover types of properties that emerge from integration of such concepts as feedback mechanisms, amplification factors

and different time scales. Systematic analysis of parameter spaces is carried out in order to reveal which variables have the most influence on temporal and steady state properties of signalling systems. As a result of extensive manipulations of these models, we are able to predict several unexpected properties of signalling system dynamics. We later confirm these by *in vitro* experiments for both pathways.

To examine the function of the specific intracellular signalling network architecture in the cellular decision making process, we have constructed a family of models that couple Wnt signalling to the cellular behaviour and cell microenvironment. Based on the state of the embedded signalling network, each cell decides whether to proliferate, differentiate, or stay quiescent. Environmental factors induce changes in the intracellular state which might change cell fate and thus feedback on the extracellular and intracellular state. Different scenarios of cell fate control are compared. The model reveals cellular characteristics that ensure a stable number and positioning of cells that assume proliferative fate. Moreover, we demonstrate the robustness of Wnt-mediated determination of cell fate that protects tissue against exponential growth when mutations occur.

1.1 Contributions

The main contribution of this thesis is to systems biology; we have produced controllable and manipulable models and used them to design, test, and refine hypotheses about cancer. These results produced important predictions, which were then validated by *in vitro* experiments. We were focusing on the following research areas:

- The FGF pathway collaborative project allowed us to discover counter-intuitive features of the dynamics of this pathway. Our analysis revealed the interplay between receptor activation and attenuation that controls proper timing of signal transduction. Understanding of this complex dynamics helped explain how pathogenic mu-

tations in FGF receptor lead to sustained signalling, thereby yielding novel insight into the function of oncogenic mutations.

- The Wnt pathway project allowed us to understand in detail a pathway that shows frequent abnormalities in a cancer cell, in particular in colon cancer. We analysed dynamic features of the system, predicted, and later confirmed previously unknown oscillatory behaviour of the pathway. We further demonstrated how this complex pathway dynamics can be utilised to control cell fate induction in developing and self-renewing tissues.

- We further built a multi-scale extension of the Wnt model by linking it to the regulation of cell proliferation and differentiation in the intestinal tissue. Results of this project demonstrate how decisions made by individual cells lead to emerging robust regulation of tissue architecture. Model predictions of the effects of mutations in the Wnt pathway are consistent with concrete colon cancer models (Familial Adenomatous Polyposis, hyperplastic and adenomatous polyps).

1.2 Publications

Article 1. J. Heath and M. Kwiatkowska and G. Norman and D. Parker and O. Tymchyshyn (Schaeffer). (2008) Probabilistic Model Checking of Complex Biological Pathways. *Theoretical Computer Science (Special Issue on Converging Sciences: Informatics and Biology)*, **391**, pages 239-257.

appeared earlier as

J. Heath, M. Kwiatkowska, G. Norman, D. Parker and O. Tymchyshyn (Schaeffer). (2006) Probabilistic model checking of complex biological pathways. In C. Priami (editor) *Proc. Computational Methods in Systems Biology (CMSB'06)*, volume 4210

of Lecture Notes in Bioinformatics, pages 32-47, Springer Verlag.

This paper presents the comprehensive computational framework for reasoning about the behaviour of signalling pathways. My contributions to this work include refining, analysis, and interpretation of results from the FGF pathway modelling case study.

Article 2. M. Kwiatkowska, G. Norman, D. Parker, O. Tymchyshyn (Schaeffer), J. Heath and E. Gaffney. (2006) Simulation and verification for computational modelling of signalling pathways. In L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto (editors) *Proc. 2006 Winter Simulation Conference*, pages 1666-1674, Winter Simulation Conference, Monterey, CA, USA.

This paper compares various computational frameworks in their ability to capture the dynamics of biological signalling pathways. My contributions to this paper include design of the FGF-based pathway reaction system, and derivation and analysis of the stochastic π-calculus model for this system.

Article 3. O. Tymchyshyn (Schaeffer) and M. Kwiatkowska. (2008) Combining intra- and inter-cellular dynamics to investigate intestinal homeostasis. In J. Fisher (editor) *Proc. Formal Methods in Systems Biology (FMSB'08)*, volume 5054 of Lecture Notes in Bioinformatics, pages 63-76, Springer Verlag.

This paper presents a multi-scale model of the regulation of cellular behaviour in the epithelial self-renewal in the intestine. Here my contributions include model design and analysis. The model is then validated using information from the literature.

Poster 1. O. Tymchyshyn (Schaeffer), G. Norman, J. Heath and M. Kwiatkowska. (2006) Computer assisted biological reasoning: the simulation and analysis of FGF signalling pathway dynamics. Poster at the NCRI Cancer Conference, Birmingham, UK.

This poster contains preliminary analysis of dynamic features of the FGF pathway including characterisation of the role of different pathway components in activating pathway amplitude and/or duration. The paper with complete analysis of the pathway is in preparation (listed as Article 4).

Poster 2. O.Tymchyshyn (Schaeffer) and M. Kwiatkowska. (2006) Computational approach for understanding molecular and cellular implications of defects in the Wnt pathway. Poster at the Winter Simulation Conference, Monterey, CA.

This poster outlines an adaptation of the stochastic π-calculus framework to modelling cellular dynamics at different levels of abstraction. Further details of model analysis are given in Article 3.

Poster 3. J. Heath, M. Kwiatkowska, G. Norman and O. Tymchyshyn (Schaeffer). (2005) Stochastic modelling of the FGF signalling pathway. Poster at the International Conference in Systems Biology, Harvard Medical School, Boston, MA.

This poster contains earlier results of the FGF modelling study.

The following papers are currently in preparation or under submission:

Article 4. O. Tymchyshyn (Schaeffer), S. Akbazardeh, D. McEwan, G. Norman, J. Heath and M. Kwiatkowska. Computational reasoning applied to FGF pathway: simulation, analysis and experimental validation.

This paper consolidates earlier work on the computational modelling of FGF signalling and discusses biological insight gained from the model. My contributions are derivation of the stochastic π-calculus model of the complete pathway, comprehensive simulation-based analysis of the model, and model refinement with additional information from experimental validation.

Article 5. O. Tymchyshyn (Schaeffer), G. Caldwell, G. Matthews, J. Heath and M. Kwiatkowska. Dynamic and noise-induced oscillations of the Wnt pathway: analysis and experimental validation.

This paper presents design, analysis and validation of the Wnt pathway model. My contributions include the design of models using deterministic ordinary differential equation and stochastic π-calculus frameworks, and model analysis including interpretation of their results in biological context.

1.3 Outline

This thesis is organised as follows:

Chapter 2 presents various modelling techniques for studying complex biological systems at the molecular level, and illustrates their application to analysis of several general types of reaction modules. We focus on qualitative changes of molecular dynamics emerging from feedback inclusion into the model.

Chapter 3 studies the FGF signalling pathway which plays a significant role in common forms of human cancer. We describe the computational model that captures biological details of the pathway and discuss results obtained from the model. Finally, we validate model predictions using an *in vitro* experimental system.

Chapter 4 contains a model of the Wnt signalling pathway primarily known for its role in colon cancer. We derive models of pathway dynamics using two different computational frameworks, and analyse pathway properties that help ensure robust function of the signalling system in the presence of random noise and mutations.

Chapter 5 describes the model-based study of cellular turnover in the intestine and the

changes that occur in cancer. We derive the extension of the stochastic π-calculus for modelling cells, and test the feasibility of hypotheses about cellular interactions proposed by different groups of biologists.

Chapter 6 concludes by summarising the main message of the thesis.

Chapter 2

Signalling networks

2.1 Introduction

In this chapter, we discuss existing approaches to modelling complex biochemical systems in particular signalling pathways. We show how systematic analysis of signalling systems can be performed using the deterministic approach based on ordinary differential equations and the stochastic approach based on process algebras. We explore the applicability of such analysis to characterising general types of recurring reaction patterns in biological networks. Having knowledge of these, in the next chapters we construct more complex models of realistic signalling pathways: Wnt and FGF, which are primarily known for their role in the initiation and progression of cancer.

2.1.1 Computational models of signalling systems

Cell signalling is a set of communication mechanisms that allows a cell to sense and respond to its environment. A cell receives information from the environment through a class of proteins known as receptors. The signal is then propagated through a network of genetic

and molecular components together forming a signalling pathway. Regulation of signalling pathways is crucial for coordination of basic cellular activities and cellular decisions about division, differentiation, and death.

Traditional analysis of signalling mechanisms in molecular biology has focused on a reductionist approach. This approach tries to explain the complex function of a signalling network by studying properties of its individual parts. This approach ignored quantitative information essential for understanding the nonlinear character of system dynamics which is very common in biological systems.

A more modern systems biology approach relies on computational modelling to understand the behaviour of molecular networks. Modelling emphasises the emergent aspects of signalling networks which disappear if the components are studied in separation. If feedback loops come into play or if the relative timing of processes makes a difference, then computational analysis can be used to infer counter-intuitive behaviour. Moreover, modelling helps uncover principles of biological design and control by comparing different network architectures under wide ranges of the numerical values that quantify those architectures. The extensive characterisation of the dynamics of cellular response to external or internal changes can then be used to predict the conditions for the breakdown of the signalling mechanism present in mutant phenotypes and the effect of drugs on it.

As computational studies are becoming more common in the field of molecular biology, the question of what constitutes novel insight arises. In many cases, models are used to generate predictions which are then validated by new experimental data. In this thesis, we build computational models that are consequently used to design *in silico* experiments, develop new hypotheses, and test them by *in vitro* experiments that provide further modelling input. We show how models can be used to identify all possible classes of behaviour that could arise from a given network topology, and help uncover useful properties and

the potential purpose of a specific biochemical system.

Due to the lack of quantitative data in experiments and high level of noise and variability in biological data, model validation by traditional fitting of model results to any particular set of experimental molecular traces is often of limited use. An assumed model and experimental data have to be compared by analysing different model structures and identifying the one consistent with the known qualitative effect of changing the operating conditions, such as mutagenesis and drugs. Using this approach, we were able to validate and update the knowledge of molecular components and their interactions in the context of realistic signalling pathways.

2.2 Modelling frameworks

Intracellular signalling networks involve interactions of numerous molecular components. The role of each molecule in a signalling network is to communicate the signal from one partner to the next. To accomplish this the molecule has to be in a defined signalling *state*. The state of a signalling molecule is characterised by covalent modification (e.g., phosphorylation), its state of association with other molecules, and possibly its location in the cell. The state of the entire molecular network can be thought of as a vector of states of participating molecules. The process of signal transduction can then be viewed as a series of transitions between such network states. These state transitions are not static but are determined by dynamically varying context in which signalling occurs. Network states and rules for transitions between them provide the building blocks for the modelling of a signalling network.

Therefore, the starting point for the model construction is the definition of a set of network states and a set of reaction rules for signal transduction, together with a list of parameters and initial conditions. We consider a well-mixed biochemical system of

constant temperature and fixed volume that consists of molecular species $\mathbf{X}_n$, $n = 1, ..., N$. Molecular species can assume different states which correspond to covalent modification or occupation of binding site by a partner protein. Let $\{X_n^s\}$, $s = 1, ..., S_n$ be the set of possible states of molecular species $\mathbf{X}_n$, where S_n is the total number of different states.

Molecular species interact by participating in reactions that lead to molecular association, state changes (covalent modification), synthesis, or degradation. Each elementary reaction R_m, $m = 1, ..., M$, is given in the following form:

$$R_m : \mathbf{l}_{1m}\mathbf{X}_1 + \mathbf{l}_{2m}\mathbf{X}_2 + ... + \mathbf{l}_{Nm}\mathbf{X}_N \stackrel{k_m}{\rightarrow} \mathbf{r}_{1m}\mathbf{X}_1 + \mathbf{r}_{2m}\mathbf{X}_2 + ... + \mathbf{r}_{Nm}\mathbf{X}_N \tag{2.1}$$

where $\mathbf{l}_{nm}$ and $\mathbf{r}_{nm}$ are non-negative integer coefficients determining reactants and products of every reaction, respectively, defined as follows:

$$\mathbf{l}_{nm}(\mathbf{r}_{nm}) = \begin{cases} i, & \text{if } i \text{ molecules of } \mathbf{X}_n \text{ are reactants (products) in a reaction } R_m \\ 0, & \mathbf{X}_n \text{ are not among reactants (products) of } R_m, \end{cases}$$

and k_m is a rate constant associated with the reaction R_m. The number of reactant molecules, i.e. $\sum_{n=1}^{N} \mathbf{l}_{nm}$ denotes the order of the reaction R_m.

At this time, no standard representation of biochemical pathways exists. Several efforts are underway to build databases of biochemical signalling networks ([Science STKE, 1996, KEGG, 1995, PID, 2006]) together with tools for visualising these networks ([Kitano et al., 2005, Kohn et al., 2006]). While such databases are a step toward organising and visualising pathway data, at this time they do not provide tools for simulation and unifying dynamic information.

Multiple computational frameworks are currently used to provide a quantitative understanding of complex biochemical networks. They differ in the way the translation

between (2.1) and formal model entities is made, as well as in the applicable analysis techniques. The more traditional approach of ordinary differential equations describes the system in terms of continuous changes of concentrations of molecular species $\mathbf{X}_n$. This approach is implemented by several biochemical modelling packages [Shapiro et al., 2002, Mendes, 1997]. Its alternative is the stochastic approach that describes the evolution of molecular numbers in terms of probability of reactions. This approach is implemented by the stochastic π-calculus [Regev and Shapiro, 2004] and PRISM [Kwiatkowska et al., 2007] computational platforms and others [Le Novere and Shimizu, 2001, Fages et al., 2004] (reviewed in [Hlavacek et al., 2006, Kwiatkowska et al., 2006]).

A promising effort attempts to build an XML-based exchange language that would integrate models in different formalisms from various sources [SBML, 2000]. However, in its current state, SBML cannot be used for representing models which contain complex molecular entities composed from multiple components that undergo independent state transitions.

Below we present a summary of different computational frameworks, together with simple examples of biochemical reaction systems, which are subsequently analysed using these formalisms.

2.2.1 Continuous deterministic approach

The continuous deterministic approach and, in particular, ordinary differential equation models (ODEs) define the state of the system at time t by a vector of concentrations $[\mathbf{X}_n]$ of biochemical species $\mathbf{X}_n$. Model equations describe the dependency of fluxes $\Delta[\mathbf{X}_n]$ on the populations of participating biochemical species $\mathbf{X}_n$ and kinetic reaction constants k_m. The main assumption required to justify this approach states that reactants are present in large numbers in well-mixed spatially uniform reaction volume, allowing us to treat

populations of species as variables varying on a continuous scale.

In such a population-based approach, each state of $\mathbf{X}_n$, X_n^s, $s = 1, ..., S_n$, is treated as a distinct molecular species and demands its own population pool. This effectively requires the modeller to specify the reaction rules for all modification states X_n^s for every reactant or product molecule $\mathbf{X}_n$. Generally, if a molecule contains j independent binding and phosphorylation sites, and therefore can be in one of 2^j states, the number of variables and equations needed to describe the dynamics of such a system grows by 2^j. Exponential explosion in the number of equations is one of the major drawbacks of applying a continuous deterministic approach to modelling complex signalling pathways.

Let us consider a set of single-state variables Y_n, $n = 1, ..., \sum_{i=1}^{N} S_i$, defined through renaming relationship $X_n^s = Y_{s+\sum_{j=1}^{n-1} S_j}$. Let $[Y_n]$ denote the concentration of the molecular species Y_n. Using an ODE approach, the rates of production and consumption of individual biomolecular species, $d[Y_n]/dt$, are represented in terms of *mass action kinetics*. A law of mass action states that rates of a reaction are proportional to the concentrations of the reacting species. ODE models for (2.1) can be written as

$$\frac{d[Y_n]}{dt} = \sum_{m=1}^{M} \nu_{nm} k_m \prod_{i=1}^{N} [Y_i]^{l_{im}} \tag{2.2}$$

where l_{im} and r_{im} are stoichiometric coefficients determining whether Y_i is a reactant or product of the reaction R_m. $\nu_{im} = r_{im} - l_{im}$ denotes the change of molecules of Y_i resulting from a reaction R_m.

After providing values of rate constants k_m and initial concentrations of all biochemical species $[Y_n](0)$, the system of equations (2.2) is solved yielding the dynamic time profile of concentrations of all species. Although analytical solutions are not guaranteed, numerical methods for solving a system with a large number of coupled nonlinear ODEs are well developed. Several integrated development environments exist that provide the convenience

of front-end data entry as well as support for efficient simulation and data manipulation algorithms [Shapiro et al., 2002, Mendes, 1997, Bower and Beeman, 1998].

The continuous deterministic approach suffers from an exponential explosion in the number of equations, hence the time needed to solve the model. Moreover, the continuous approach is unable to describe the fluctuations in the molecular population levels. As we will see further (section 2.4.2), it is not even guaranteed that the model equations will provide a sufficiently accurate account of the average molecular population levels. For a biochemical system to be compatible with the continuity assumption, the number of molecules of each species must be large compared to thermal fluctuations in concentrations, and the number of reactions of each type needs to be large in each observation interval. The deterministic analysis breaks when fluctuations are amplified by the signalling network.

2.2.2 Discrete stochastic approach

A conventional ODE method represents a continuous approximation of reactions that actually involve interactions between individual molecules, which is a probabilistic process. An alternative is an event-based approach which deals with the probabilistic description of the system. In the stochastic approach, the state of the system at time t is defined by the integer vector of numbers of molecules of each biochemical species $\mathbf{X}_n$:

$$\mathbf{X}(t) = (\#\mathbf{X}_1(t), \#\mathbf{X}_2(t), ..., \#\mathbf{X}_N(t)).$$

The number of molecules divided by the reaction volume is the concentration $[\mathbf{X}_i] = \#\mathbf{X}_i/\Omega$. Ω is defined as the volume of the reaction system multiplied by the Avogadro number. The common values of Ω vary between 1 and 1000.

The behaviour of the counts of different species is described by their joint probability density function $P(\mathbf{x}, t) = Prob\{\#\mathbf{X}_i(t) = \mathbf{x}_i, i = 1, ..., N\}$, defined as the probability

of having $\mathbf{x}_i$ molecules of species $\mathbf{X}_i$ at time t, as opposed to deterministic rate equation descriptions of the absolute concentration of these molecules.

In order to define the probability $P(\mathbf{x}, t)$ in terms of the system state $\mathbf{X}$, one needs to know the probability of the occurrence of every elementary reaction step in the biochemical system. Each transition between system states is represented by a reaction step and the probability of that transition is determined by the probability of the corresponding reactions. This relation is formulated as a set of partial differential equations known as the chemical master equation (CME). For the reaction model in (2.1), the chemical master equation is

$$\frac{\partial P(\mathbf{x}, t)}{\partial t} = \sum_m (a_m(\mathbf{x} - \boldsymbol{\nu}_m) P(\mathbf{x} - \boldsymbol{\nu}_m, t) - a_m(\mathbf{x}) P(\mathbf{x}, t)). \tag{2.3}$$

where $a_m(\mathbf{x})$, called the *propensity function*, is the probability per unit of time that one reaction R_m occurs within this time interval if the system is in the state $\mathbf{x}$. $\boldsymbol{\nu}_m = (\boldsymbol{\nu}_{1m}, ..., \boldsymbol{\nu}_{Nm}) = (\mathbf{r}_{1m} - \mathbf{l}_{1m}, ..., \mathbf{r}_{Nm} - \mathbf{l}_{Nm})$ is the stoichiometric vector defining the result of the mth reaction. The propensity function is further defined as

$$a_m(\mathbf{x}) = c_m h_m(\mathbf{x}),$$

where $c_m > 0$ is the *stochastic rate constant* of the mth reaction, and $h_m(\mathbf{x})$ is the number of distinct combinations of molecular reactants associated with the reaction R_m in state $\mathbf{x}$.

The stochastic reaction rate constant c_m is closely related to the conventional macroscopic reaction rate constant k_m, used in the deterministic approach. The stochastic rate constant no longer describes the concentration of chemical species being produced or consumed per unit time, but rather the probability per unit time that a randomly chosen combination of reactant molecules will react. Therefore c_m is inversely proportional to the

reaction volume and satisfies the relationship $c_m = k_m \Omega^{1-n}$, where n is the reaction order.

Analytical solutions to the CME are only practical for simple reaction systems. However, one can resort to a Monte Carlo type of numerical simulation, such as the Gillespie algorithm [Gillespie, 1976], also called Stochastic Simulation Algorithm (SSA). Based on the theory of thermodynamics, the Gillespie algorithm leads to an exact procedure for numerical simulation of the dynamic evolution of a reacting system. Based on the propensity functions a_m at any given time, the algorithm computes the reaction that will occur next in the system, as well as the time interval within which this reaction occurs. The number of molecules of reacting species as well as reaction probabilities are updated at each time step. Iteration of this procedure until a final time is reached provides numerical realisations possessing the same statistical properties as those described by the CME. The method is accurate even at low numbers of reactants where the assumption of continuity used in ODEs breaks down.

The Gillespie algorithm does not easily handle the reactions of multistate molecules. Similar to the ODE approach, this method faces an exponential performance decrease under linear increase of the number of independent protein sites. Moreover, because the Gillespie algorithm does not represent each molecule in the system separately, it can neither associate physical quantities with each molecule, nor trace the fate of particular molecules over period of time.

2.2.3 Modelling signalling dynamics in π-calculus

The next two sections describe process algebra approaches to the stochastic modelling of signalling networks. The process-algebra approaches were originally developed in computer science for describing and reasoning about networks of concurrent components. They have been successfully applied for analysing several types of molecular and genetic

systems [Regev and Shapiro, 2004, Heath et al., 2007, Calder et al., 2006]. Once a biological system has been described using process algebra language constructs, the model can be stochastically simulated to derive the properties under study over time. Stochastic π-calculus [Priami, 1995] is a stochastic process algebra where interactions are assigned rates controlled by exponential distributions. In this thesis, we use the BioSPI [Regev and Shapiro, 2004] as the platform which performs simulations of the π-calculus code using an adaptation of the Gillespie algorithm [Gillespie, 1976].

A model in the stochastic π-calculus is a composition of communicating components (called *processes*). Communication between a pair of processes occurs on complementary *channels*. Output $\mathtt{r}!$ and input $\mathtt{r}?$ prefixes, where $\mathtt{r}$ is a channel name, are elementary constituents of communication capabilities. Following communication between process $\mathtt{X} ::= \mathtt{r}!\{\mathtt{y}\}, \mathtt{X}'$, containing an output capability $\mathtt{r}!\{\mathtt{y}\}$, and process $\mathtt{Y} ::= \mathtt{r}?\{\mathtt{z}\}, \mathtt{Y}'$, containing an input capability $\mathtt{r}?\{\mathtt{z}\}$, the action involving channel $\mathtt{r}$ is removed, process $\mathtt{X}$ changes its state to $\mathtt{X}'$, while process $\mathtt{Y}$ substitutes channel name $\mathtt{y}$ for $\mathtt{z}$ and changes its state to $\mathtt{Y}'$ (see also Table 2.1 for the formal reduction rules of π-calculus). Figure 2.1(a) displays graphically communication between processes $\mathtt{X}$ and $\mathtt{Y}$.

If the set of channel names transmitted over a channel is empty, communication capabilities are displayed as $\mathtt{r}![\,]$ and $\mathtt{r}?[\,]$. Alternatively, communication between processes may carry a set of channel names that further changes interaction capabilities of participating processes. Such a change in future communication capabilities as a result of passing channel names is termed *mobility*.

Furthermore, a process $\mathtt{X}$ may be defined as a *choice* between a set of processes $\mathtt{X_1}, \ldots \mathtt{X_n}$:

$$\mathtt{X} ::= \pi_1, \mathtt{X_1} + ... + \pi_n, \mathtt{X_n}$$

may evolve as either of $\mathtt{X_i}$, depending on which of the communication capabilities π_i is the

first one to complete in the current context, thus representing a race condition between a set of processes. In a graphical presentation, each edge from node $\mathtt{X}$ to node $\mathtt{X_i}$ is labelled with an action $\pi_\mathtt{i}$ and denotes an alternative execution path in the system (Fig. 2.1(b)).

A process $\mathtt{X}$ given by

$$\mathtt{X} ::= \mathtt{X_1} \mid \ldots \mid \mathtt{X_n}$$

denotes a *composition* of processes $\mathtt{X_1}$, ..., $\mathtt{X_n}$ running in parallel. Graphically, $\mathtt{X}$ can be modelled as in Fig. 2.1(c), where each edge from node $\mathtt{X}$ to node $\mathtt{X_i}$ denotes a concurrent execution path in the system. Node $\mathtt{X}$ is represented as a solid rectangle.

The creation of a *private channel* $\mathtt{r}$ within the scope of the given process $\mathtt{X}$ is achieved by the expression

$$(\mathtt{new}\ \mathtt{r}(\lambda))\,\mathtt{X}$$

which yields a new process in which channel $\mathtt{r}$ with rate λ is bound. Only processes that share a private channel may interact using that channel. Name binding is represented graphically as a bubble labelled with the restricted name $\mathtt{r}$ around the node $\mathtt{X}$ (Fig. 2.1(d)). Timing is incorporated into π-calculus models by associating each channel $\mathtt{r}$ with the rate governed by the exponential distribution with the mean $1/\lambda$.

$$\begin{array}{c}
\mathtt{r}!\{\mathtt{y}\},\ \mathtt{X}' \mid \mathtt{r}?\{\mathtt{z}\},\ \mathtt{Y}' \quad \rightarrow \quad \mathtt{X}' \mid \mathtt{Y}'[\mathtt{z}/\mathtt{y}] \\
\ldots + \mathtt{r}!\{\mathtt{y}\},\ \mathtt{X} + \ldots \mid \ldots + \mathtt{r}?\{\mathtt{z}\},\ \mathtt{Y} + \ldots \quad \rightarrow \quad \mathtt{X} \mid \mathtt{Y}[\mathtt{z}/\mathtt{y}] \\
\text{if } \mathtt{X} \quad \rightarrow \quad \mathtt{X}' \text{ then } \mathtt{X} \mid \mathtt{Y} \quad \rightarrow \quad \mathtt{X}' \mid \mathtt{Y} \\
\text{if } \mathtt{X} \quad \rightarrow \quad \mathtt{X}' \text{ then } (\mathtt{new}\ \mathtt{r})\,\mathtt{X} \quad \rightarrow \quad (\mathtt{new}\ \mathtt{r})\,\mathtt{X}'
\end{array}$$

Table 2.1: Operational semantics of π-calculus.

A translation scheme that maps molecular signalling pathways into π-calculus programs was first introduced in [Regev and Shapiro, 2002, Regev and Shapiro, 2004]. Molecular entities can be coded in π-calculus as processes that participate in reactions by communi-

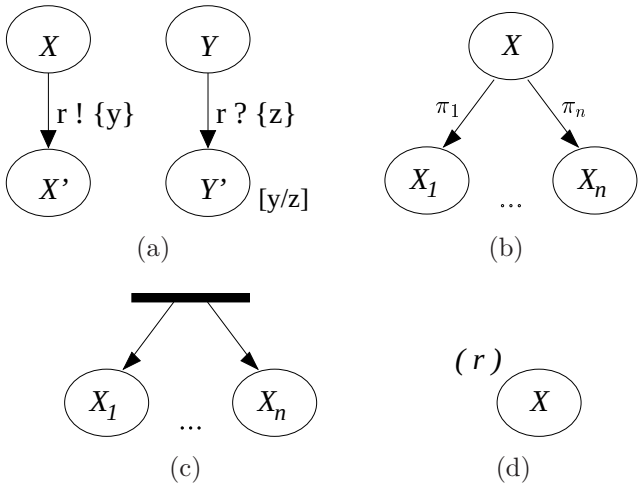


Figure 2.1: Basic stochastic π-calculus operations: (a) communication, (b) choice, (c) parallel composition, and (d) scope restriction primitives.

cating over channels. State transitions resulting from process communication correspond to covalent modification, association/dissociation, or degradation of signalling molecules. By design, all actions in π-calculus necessitate precisely two participants. Let us consider how reactions of different orders are modelled. *Zero-order* reactions require definition of a fictitious constant reactant process that transforms them into first-order reactions. *First-order* reaction R_m of the form

$$R_m : \mathbf{X} \rightarrow \sum_{n=1}^{N} \mathbf{r}_n \mathbf{X}_n$$

can be represented in π-calculus by defining the following processes:

$$\mathtt{X} ::= \mathtt{m}\,!\,[\,],\mathtt{Y}.$$

$$\mathtt{Y} ::= \underbrace{\mathtt{X_1} \mid \ldots \mid \mathtt{X_1}}_{\mathtt{r_1\ times}} \mid \ldots \mid \underbrace{\mathtt{X_N} \mid \ldots \mid \mathtt{X_N}}_{\mathtt{r_N\ times}}.$$

where process $\mathtt{X_n}$ is the realisation of biochemical species $\mathbf{X}_n$. The model has to be supplied

with a constant process Clock that enables a pairwise communication on m:

$$\texttt{Clock} ::= \ldots + \texttt{m}\,?\,[\,], \texttt{Clock} + \ldots$$

For example, first-order reaction $A \rightarrow A^* + B$ in which molecule A phosphorylates and recruits molecule B, can be represented in π-calculus as

$$\texttt{A} ::= \texttt{m}\,!\,[\,],\ \texttt{A}^* \,|\, \texttt{B}.$$

Second-order reaction of the form

$$R_m : \mathbf{X}_1 + \mathbf{X}_2 \ \rightarrow \ \sum_n \mathbf{r}_n \mathbf{X}_n$$

is modelled by complementary binding motifs between molecular species that allows them to participate in an association reaction. We denote these motifs by the pair of input and output communication capabilities on channel m:

$$\texttt{X}_1 ::= \texttt{m}\,!\,[\,], \texttt{Y}_1.$$

$$\texttt{X}_2 ::= \texttt{m}\,?\,[\,], \texttt{Y}_2.$$

The parallel composition of the resulting processes $\texttt{Y}_1$ and $\texttt{Y}_2$ has to contain r_n instances of process $\texttt{X}_\texttt{n}$ for each n, running in parallel. One possibility to represent these is shown below:

$$\texttt{Y}_1 ::= \underbrace{\texttt{X}_1 \,|\, \ldots \,|\, \texttt{X}_1}_{\texttt{r}_1\ \texttt{times}} \,|\, \ldots \,|\, \underbrace{\texttt{X}_\texttt{N} \,|\, \ldots \,|\, \texttt{X}_\texttt{N}}_{\texttt{r}_\texttt{N}\ \texttt{times}}.$$

$$\texttt{Y}_2 ::= \texttt{0}.$$

Cases of higher order reactions with three or more reactants cannot be directly implemented in the π-calculus. Such situations can be approximated by introducing the intermediate complexes corresponding to the cascade of pairwise component bindings. Since the reaction occurrence is based on the probability of collision between reactant molecules within a small interval of time, which is small for larger reactant numbers, higher order reactions are rare.

Molecules with several independent functional domains are represented as a parallel composition of π-calculus processes. If molecule $\mathbf{X}$ contains j independent binding and phosphorylation sites, the process $\mathtt{X}$ will be implemented as a parallel composition of j processes:

$$\mathtt{X} ::= \mathtt{X_1} \mid \ldots \mid \mathtt{X_j}.$$

The state of each $\mathtt{X_1}$, ..., $\mathtt{X_j}$ reflects the set of activities in which the site can participate. In the ODE approach the modeller would be required to write reaction rules and equations for all modification states of $\mathbf{X}$ which grow exponentially with the number of independent sites j. In π-calculus, each process $\mathtt{X_1}$, ..., $\mathtt{X_j}$ contains the definition for reaction rules in which the respective site participates. In this way, the number of reactions needed to describe $\mathbf{X}$ in π-calculus grows linearly with the number of sites.

There are two major simulation platforms available for the analysis of π-calculus models, BioSPI [Regev and Shapiro, 2004] and SPiM [Phillips and Cardelli, 2005]. In this thesis, we use BioSPI for performing simulations of the constructed pathway models because it enables compartment-based models (analysed in Chapter 5), in contrast to SPiM. The BioSPI platform receives as input π-calculus code in textual format. A complete π-calculus program consists of a set of public channel declarations, a set of process definitions, and an initial process $\mathtt{S}$ which is to be reduced with respect to these declarations and definitions. Simulations are performed using an adaptation of the Gillespie algorithm. By taking a

component-based, rather than reaction-based approach implemented by the Gillespie algorithm, π-calculus simulations avoid the problem of combinatorial explosion. There is no global space that lists all states of $\mathbf{X}$; instead local state spaces are used to derive values of each subcomponent of $\mathbf{X}$.

Consider a simple biochemical system consisting of two molecular species, **A** and **B**. Species **A** uses its binding site to associate with species **B**. Assume as well that in addition to binding site, molecule **A** possesses n independent phosphorylation sites. Using BioSPI notation, this system can be described as follows:

$$\begin{aligned} \mathtt{A} &::= \mathtt{Bind} \mid \mathtt{Ph_1} \mid \ldots \mid \mathtt{Ph_n}. \\ \mathtt{Bind} &::= \mathtt{bind!}\{\mathtt{rel}\}, \mathtt{rel?}[\,], \mathtt{Bind}. \\ \mathtt{B} &::= \mathtt{bind?}\{\mathtt{rel}\}, \mathtt{rel!}[\,], \mathtt{B}. \end{aligned} \tag{2.4}$$

Reactions involving subprocess `Bind` which describes the state of binding between **A** and **B**, are modelled independently of reactions involving phosphorylation sites $\mathtt{Ph_1}$, ..., $\mathtt{Ph_n}$. This is in contrast to approaches that represent the global state space of molecule **A**, such as deterministic ODE-based approach and stochastic approach based on the Gillespie algorithm. The global state space of molecule **A** contains all possible combinations of binding and phosphorylation sites, thus 2^{n+1} states in total. The number of reactions needed to describe binding between **A** and **B** grows proportionally.

2.2.4 Signalling networks in PRISM

Probabilistic model checking and the probabilistic model checker PRISM [Kwiatkowska et al., 2007] is an alternative framework for modelling and analysis of biochemical networks. One of the benefits of this approach is the ability to automatically calculate *exact* quantitative measures of the stochastic system under study based on its

systematic and exhaustive exploration. The applicability of these techniques has been demonstrated for complex systems from a broad range of domains, including security and communication protocols, distributed algorithms and power management. Recently, PRISM has also been applied to modelling and analysis of several biochemical reaction networks such as MAPK [Calder et al., 2006] and FGF pathways [Heath et al., 2007].

As with with the π-calculus approach, the system under study has first to be described in the PRISM modelling language. Each of the basic molecular and genetic components of the pathway is represented by a separate PRISM *module*. The different forms which each protein can take are represented by one or more finite-ranging *variables* within the module.

Synchronisation between modules is used to model reactions involving interactions of multiple elements. PRISM is based on multi-way synchronisation rather than on binary channel communication of π-calculus. PRISM model for (2.1) can be written as shown in Fig. 2.2. Here we model the state of the system by the vector of counts of different molecular species $\mathbf{X}_n$, $n = 1, ..., N$, represented by N modules, each with one variable x_n storing the count of $\mathbf{X}_n$ present in the system. The behaviour of each module X_n is specified by a number of guarded commands of the form $[a]\ g \rightarrow r : u$. If the predicate, or guard, g is true, the system is updated according to u at the rate r. For example, if the value of variable x changes as a result of this command, this is denoted by $x' = \ldots$. The action label a denotes that the state of several modules changes simultaneously. When the reaction R_m occurs, variable x_n changes as described by stoichiometric coefficients $\mathbf{l}_{nm}$ and $\boldsymbol{\nu}_{nm} = l_{nm} - r_{nm}$, for each $n = 1, ..., N$. The rate of reaction R_m is proportional to the product of counts of reactant species x_n factored by l_{nm}, denoted as $x_n^{l_{nm}}$. PRISM computes the rate of a combined transition as the product of the rates for all transitions synchronised on the same guard R_m. We include the stochastic rate constant c_m in a

```
module X_n
  x_n : [0..N_x] init N_x^0;

  ...
  //if X_n is a substrate in the reaction R_m, ie l_nm > 0
  [R_m] x_n ≥ l_nm & x_n ≤ N_x − ν_nm → x_n^{l_nm} : (x'_n = x_n + ν_nm);

  //otherwise
  [R_m] x_n ≤ N_x − ν_nm → 1 : (x'_n = x_n + ν_nm);
  ...
endmodule

module basal
  dummy : [0..1] init 0;

  ...
  [R_m] true → c_m : true;
  ...
endmodule
```

Figure 2.2: Biochemical reactions in PRISM

separate module. Note that the action label a may be omitted for transitions involving a single module.

Different states $X_n^s,\ s = 1, ..., S_n$, of molecule $\mathbf{X}_n$ can be represented by values of the variable x_n within module X_n. Independent domains of the same molecule can be modelled as independent variables within one module. Example 2.4 can be modelled in PRISM as follows: In this way, a model description in PRISM avoids exponential state explosion [Heath et al., 2007].

From a model description, PRISM builds a system representation, typically a labelled state-transition system in which each state represents a possible configuration and the transitions represent the evolution of the system from one configuration to another over time. Transitions are annotated with rates which are interpreted as parameters of a negative exponential distribution. Models are analysed by exhaustive exploration of the paths of the system in order to reason about the probability that a model behaves in a

```
module A
  bind : [0..N_a] init 0;
  ph_1 : [0..N_a] init 0;
  ...
  ph_n : [0..N_a] init 0;

  [bind] bind ≤ N_a − 1 → (N_a − bind) : (bind′ = bind + 1);

  [rel] bind ≥ 1 → bind : (bind′ = bind − 1);
  ...
endmodule

module B
  b : [0..N_b] init 0;

  [bind] b ≤ N_b − 1 → (N_b − b) : (b′ = b + 1);

  [rel] b ≥ 1 → b : (b′ = b − 1)
endmodule
```

certain fashion, as well as other quantitative measures relating to the temporal evolution of the model.

The main contribution of PRISM is the formalisation of qualitative and quantitative properties of the system in temporal logic. PRISM models can be augmented with rewards associated either with system states or transitions. A state reward is accumulated in proportion to the time spent in a given state. A transition reward is accumulated every time the transition is taken. PRISM can then analyse properties which relate to the expected values of these rewards. Properties are specified using $\mathcal{R}$ operator, e.g., expression $\mathcal{R}_{=?}[\,\texttt{rewardprop}\,]$ returns the expected value of the reward associated with `rewardprop` (see [Kwiatkowska et al., 2007] for formal definition of reward-based properties). For example, one might be interested in the following properties:

- What is the *expected number of molecules* of species $\mathbf{X}$ at time instant t. This is

expressed by the formula:

$$\mathcal{R}_{=?}[\,\mathcal{I} = t\,]$$

with reward x indicating the number of molecules $\mathbf{X}$ is associated with every state.

- What is the *expected number of reactions* of a certain type that occurred by time t. The respective formula is:

$$\mathcal{R}_{=?}[\,\mathcal{C} \leq t\,]$$

with reward of 1 associated with every transition corresponding to the reaction of interest.

- What is the *expected time* until the number of molecules $\mathbf{X}$ reaches a threshold T_x. This is expressed as:

$$\mathcal{R}_{=?}[\,\mathcal{F}(x \geq T_x)\,]$$

and reward of 1 associated with every state.

The formalisation allows one to automatically verify whether or not each property is satisfied.

Because probabilistic model checking carries out analysis of the global state of the system, at the stage of model analysis it faces a problem of an exponential state explosion in the number of components of the underlying model. PRISM supports state-of-the-art techniques which tackle this problem, thus increasing the size of systems that can be analysed [Kwiatkowska et al., 2007].

The stochastic π-calculus and PRISM approaches differ in their ability to represent non-linear kinetics of biochemical reactions. π-calculus is limited to two-way synchronisation between system components which allows to represent basic zero-, first- and second-order reactions. On the other hand, PRISM model may contain reactions with

more complex non-linear dynamics, such as Michaelis-Menten enzymatic reaction or Hill exponent reactions. However, π-calculus notion of mobility may be used to model infinite arrays of processes (molecules, cells, etc) and channels (reactions), which cannot be expressed in PRISM.

2.3 Analysis techniques

2.3.1 Simulation

The majority of analysis techniques used in both deterministic and stochastic frameworks rely on simulation which generates time trajectories of chosen properties of the system. Simulation is a simple but powerful tool for studying behaviour of signalling networks. Using simulation, time-dependent profile of molecular species can be compared over a wide range of rate parameters, initial concentrations, and reaction network topologies. To simulate the effect of a kinase inhibitor, for example, a model is run with the reduced rate of phosphorylation reaction. The rate of protein interactions with its partners can be varied to study the effect of mutations affecting the structure of this protein. In each case, simulated data can be compared to experimental data such as quantitative western blotting, ELISA test, or flow cytometry.

2.3.2 Sensitivity analysis

Sensitivity analysis is used to determine which concentrations and rate constants in a model have the most influence on the overall system behaviour. In addition, this type of analysis is valuable in ascertaining which parameters should be the focus of direct measurement or experimental perturbation. In order to determine sensitivity of the model, we need to formulate the objective function Z (e.g., maximum number of signalling molecules)

and determine the set of model variables p_j over which the sensitivity analysis should be performed. The sensitivity coefficients $S^Z_{p_j}$ describe the change in the objective function Z due to variations in the parameters p_j. Sensitivity measure independent from the units of Z and p_j is as follows:

$$S^Z_{p_j} = \frac{p_j}{Z}\frac{\partial Z}{\partial p_j} = \frac{\partial \ln Z}{\partial \ln p_j} \tag{2.5}$$

Sensitivity coefficients can be found analytically for relatively simple systems or calculated numerically using a variety of methods such as the direct method, the decoupled direct method, Green functions, and others [Varma et al., 1999].

Sensitivity analysis in the case of stochastic systems measures the overall changes in the probability density function. A system state is a random variable with the probability density function $P(x,t)$ which follows a CME (2.3). We define the output as a function of the state $Z(x,t)$. A direct analog of sensitivity coefficients (2.5) for a discrete stochastic system with respect to the output Z, can be defined as the expectation of sensitivity measure:

$$S^Z_{p_j} = E\left[\frac{\partial \ln Z(x,t)}{\partial \ln p_j}\right] = \sum_x \frac{\partial \ln Z(x,t)}{\partial \ln p_j} P(x,t). \tag{2.6}$$

The dependence of the state x on the parameter p_j is implicitly assumed.

Sensitivity as defined by (2.5) or (2.6) is context specific, that is, it is performed around a particular point in the parameter space. However, calculating sensitivity by simultaneously altering multiple parameters provides a more global view of network behaviour. To estimate the total sensitivity of Z to p_j, a reference state is perturbed by random modifications of system variables p_j. Each alteration of the reference system is characterised by the total parameter variation σ defined as $\log(\sigma) = \sum_j |\log(p^*_j/p_j)|$, where p^*_j are the biochemical parameters of the altered system. Z is plotted as a function of the total parameter variation σ, for an ensemble of model systems.

2.3.3 Changes in the trajectory

Another approach to analysing models is qualitative and based on determining the classes of physiologically distinct behaviours that the system can produce. Bistability, oscillations, wave propagation, and pattern formation are among nonlinear dynamic phenomena crucial for the regulation of biological processes. *Bifurcation analysis* focuses on determining whether parameter changes can lead to a dramatic transition between different trajectories through the phase space which is defined as a space in which each coordinate stands for a time-dependent variable in the model. For example, a steady-state solution corresponding to a single point in the phase space may lose stability and be replaced by a stable limit cycle solution (a closed curve in the phase space corresponding to periodic behaviour of the model). This is called a Hopf bifurcation. A bifurcation diagram is used to graphically represent how the trajectory changes as a function of a model parameter.

2.3.4 Experimental validation

This section summarises measurement technologies that are available for proving very specific predictive hypothesis resulting from computational modelling. Different experimental technologies can be used to verify gene or protein properties predicted by the model under study (reviewed in [Albeck et al., 2006, Kingsmore, 2006]). They vary in their cost, throughput, sample size, and their applicability to a single cell or population of cells.

Because of the relative ease of *immunoblotting*, this technique has become the most widely used method for characterising protein size and the relative abundance. Immunoblots are primarily used to compare relative abundance of a particular protein present in samples under different physiological and experimental conditions; thus different immunoblots cannot be directly compared to each other. Immunoblotting can be used to determine protein size, and therefore it has been a key method for detecting mutations

that result in abnormal protein size. The main limitations of this method stem from the inability to fully control the experimental conditions which frequently results in large fluctuations in the output.

Enzyme-linked immunosorbent assay (ELISA) is an experimental technique for quantitation of antigen or antibody in which enzyme-labeled antibody or antigen is bound to a solid support. ELISA shows sensitivity more than 100-fold higher than immunoblot and can predict smaller changes occurring during signalling. Immunoblots are often used to confirm the specificity of antibodies which are detected by ELISA procedures.

Polymerase chain reaction (PCR) is a highly efficient method to amplify low levels of specific DNA sequences in a sample in order to reach the threshold of detection. The extreme sensitivity of PCR has made it the favoured technique in quantitation of low abundance DNA. In designing a quantitative PCR assay, several variables that influence DNA amplification must be determined and controlled, thus complicating the use of PCR.

Fluorescence-based techniques and imaging can be used to quantify signals at cellular or subcellular resolutions. Sensitivity of detection is one limiting factor of such techniques, as is low throughput. Simultaneous measurement of a maximum of 10 different signals is practical. However, fluorescence-based techniques are necessary when it is essential to monitor changes in a single cell.

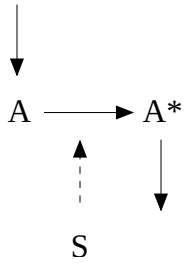
The optimal choice of the experimental technique depends on the system under study. Immunoblotting is generally a logical first step in the characterization of protein concentration as it allows visualisation of intact protein and provides information about protein size. For applications in which sensitivity or absolute concentrations are important, ELISA and PCT are more suitable choices.

2.4 Examples of signalling mechanisms

In this section, we use the described modelling frameworks to study the properties of simple biochemical systems. In subsequent discussion, we use the term *module* to describe the set of interacting components which carry out a specific cellular function. Such modules are often reused in the complex biochemical networks to execute the respective function.

2.4.1 Receptor module

We start with a simple modelling system, in which external stimulus S binds and promotes activation by phosphorylation of the receptor molecule A. Activated receptor A^* sends a signal to downstream components. To ensure termination of successful signal transduction, the phosphorylated receptor complex is degraded. We also include a low-level synthesis and phosphorylation of non-active receptor molecules. The reaction network is represented graphically as



Solid arrows indicate activation of system components and dashed arrows represent stimulus-induced changes of these activation events.

The π-calculus model for the receptor module is shown in Fig. 2.3 (please refer to the Appendix A for the textual representation). It includes processes `A` and `S` corresponding to unbound and unphosphorylated receptor and stimulus molecules. A constant process `Syn` transits periodically on channel `syn` to create a new instance of the receptor process. The non-active receptor synchronises with the stimulus on channel `bind` and changes its

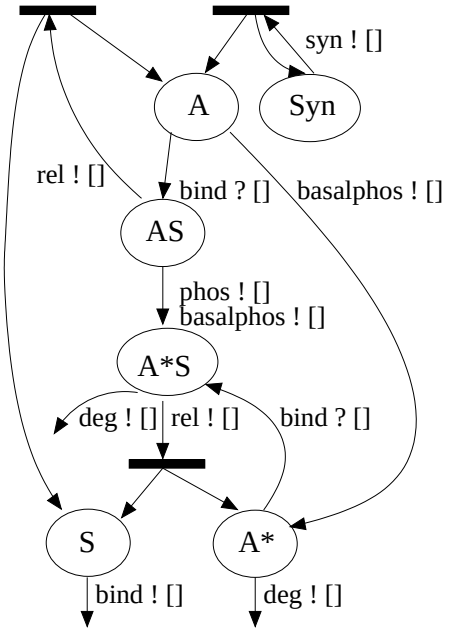


Figure 2.3: Implementation of the receptor module in π-calculus.

state to `AS`. The resulting process `AS` can either release the stimulus (transition marked by channel `rel`) or be phosphorylated to become $\texttt{A}^*\texttt{S}$ (channel `phos`). An active form of the receptor complex can either disassociate (channel `rel`) or degrade (channel `deg`). A low-level phosphorylation of both unbound and bound forms of the receptor, `A` and `AS`, proceeds on channel `basalphos`. We also define a constant auxiliary process `Clock` which contains input capabilities on channels `syn`, `rel`, `basalphos`, `phos` and `deg` (not shown).

Rate constants associated with first-order reactions (c_{rel}, $c_{basalphos}$, c_{phos}, and c_{deg}) do not depend on the units of reaction species and are therefore the same for stochastic and deterministic models. Association reaction is binary and hence c_{bind} has to be adjusted by reaction volume Ω to be equal to the deterministic rate of the reaction $c_{bind} = k_{bind}/\Omega$. Zero-order reaction rate c_{syn} is similarly adjusted $c_{bind} = k_{bind} * \Omega$.

The model is first analysed using BioSPI engine to reveal how the response component A^* is regulated by the incoming signal S. In Fig. 2.4(a)-2.4(d), we plot the transient

response of the receptor module at two different values of S (both single run and an average over 100 runs is shown). The rate constants used in simulations are as follows: $c_{bind} = 0.1/\Omega, c_{rel} = 0.1, c_{basalphos} = 0.0005, c_{phos} = 0.1, c_{deg} = 0.005, c_{syn} = 0.005 * \Omega$, with the initial number of receptor molecules $\#R(0) = \Omega$ and $\Omega = 100$.

In order to determine the characteristics of the underlying stochastic process, we implement the receptor reactions in the PRISM probabilistic checker. In the PRISM model, we adopt a population-based approach by describing the evolution of molecular counts, rather than interactions between individual molecules. PRISM modules representing populations of S and A are given in Fig. 2.5. The receptor module contains variables representing counts of free, bound, and phosphorylated receptor molecules. The only variable in the stimulus module corresponds to the number of unbound stimulus molecules. These modules synchronise on actions *bind* and *rel*, which increase and decrease the number of receptor complex molecules, respectively. The other commands of the module *receptor* are zero-order synthesis, first-order phosphorylation and first-order degradation which are independent of *stimulus* (thus no action label).

To check whether the PRISM model produces results similar to the π-calculus model, we first examined the expected number of active receptors A^* using the temporal logic property $\mathcal{R}_{=?}[\mathcal{I} = t]$ with the reward of the number of $a^* + a^*s$ molecules assigned to each state and t ranging between 0 and the maximum observation time of one hour. The results closely match those in Fig. 2.4 (data not shown). Using probabilistic model checking we further examine the probability distribution of molecular counts. The properties that we check are as follows (see results in Fig. 2.6):

- What is the probability distribution of the number of signalling molecules A^* at time

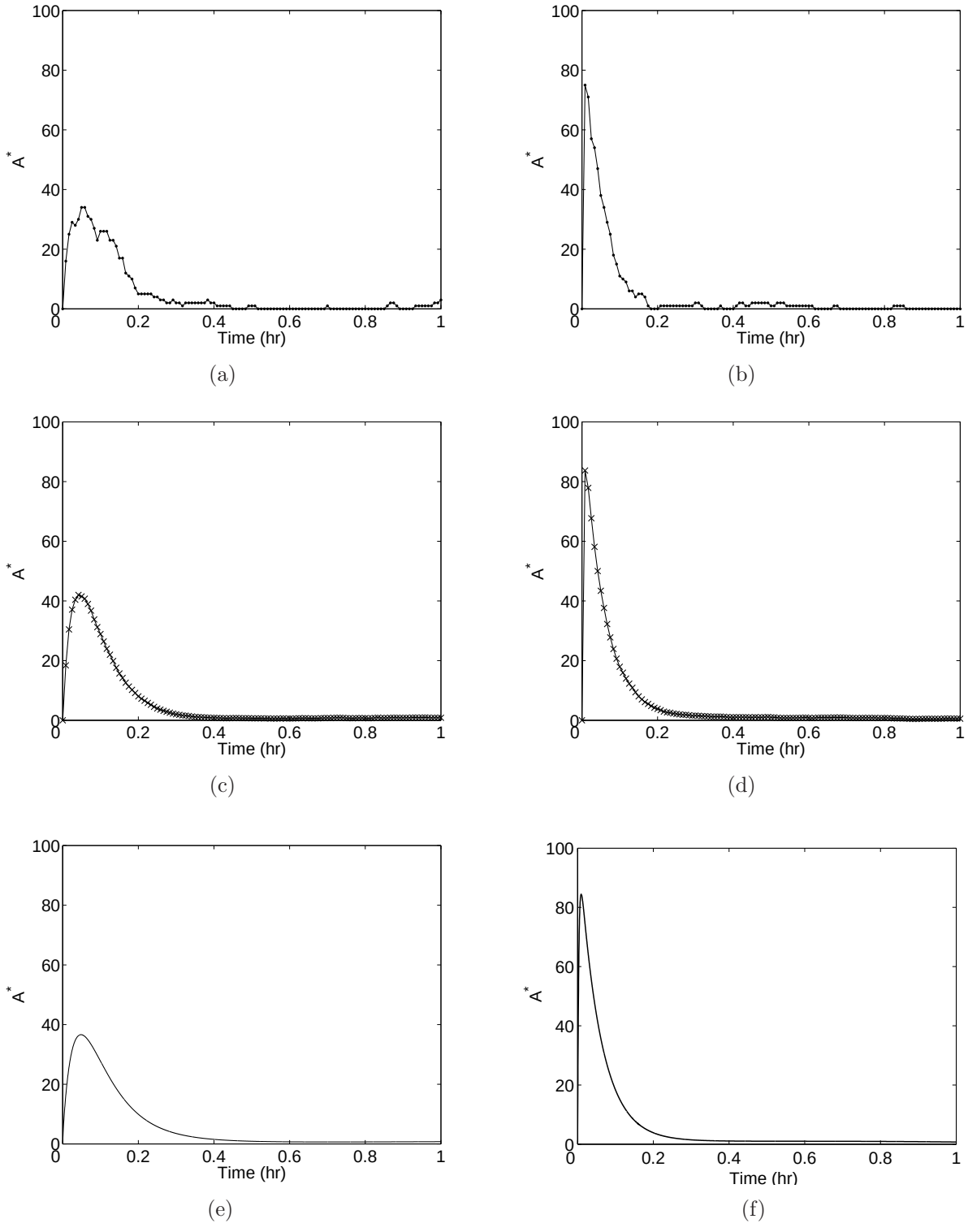


Figure 2.4: Temporal evolution of the receptor subsystem: a single stochastic run (a and b), an average of 100 stochastic runs (c and d) and deterministic solution (e and f) for $\#S(0) = 0.1 * \Omega$ and $\#S(0) = \Omega$, respectively.

```
module receptor
  a  : [0..N_a] init N_a^0;
  as : [0..N_a] init 0;
  a* : [0..N_a] init 0;
  a*s : [0..N_a] init 0;

  //receptor synthesis
  [] a < N_a -> c_syn : (a' = a + 1);

  //association and disassociation
  [bind] a > 0   & as < N_a  -> c_bind * a  : (a' = a - 1)  & (as' = as + 1);
  [bind] a* > 0  & a*s < N_a -> c_bind * a* : (a*' = a* - 1) & (a*s' = a*s + 1);
  [rel]  a < N_a & as > 0    -> c_rel * as  : (a' = a + 1)  & (as' = as - 1);
  [rel]  a* < N_a & a*s > 0  -> c_rel * a*s : (a*' = a* + 1) & (a*s' = a*s - 1);

  //phosphorylation
  [] as > 0 & a*s < N_a -> c_phos * as      : (as' = as - 1) & (a*s' = a*s + 1);
  [] as > 0 & a*s < N_a -> c_basalphos * as : (as' = as - 1) & (a*s' = a*s + 1);
  [] a > 0  & a* < N_a  -> c_basalphos * a  : (a' = a - 1)   & (a*' = a* + 1);

  //degradation
  [] a* > 0  -> c_deg * a*  : (a*' = a* - 1);
  [] a*s > 0 -> c_deg * a*s : (a*s' = a*s - 1);
endmodule

module stimulus
  s : [0..N_s] init N_s;

  [bind] s > 0 -> s : (s' = s - 1);
  [rel] s < N_s -> 1 : (s' = s + 1);
endmodule
```

Figure 2.5: Receptor module in PRISM

point t? This is given by the formula:

$$\mathcal{P}_{=?}[true\, \mathcal{U}^{[t,t]}\ a^* + a^*s = i]$$

where t is time and i is an integer parameter from the range $[0, N]$ which stands

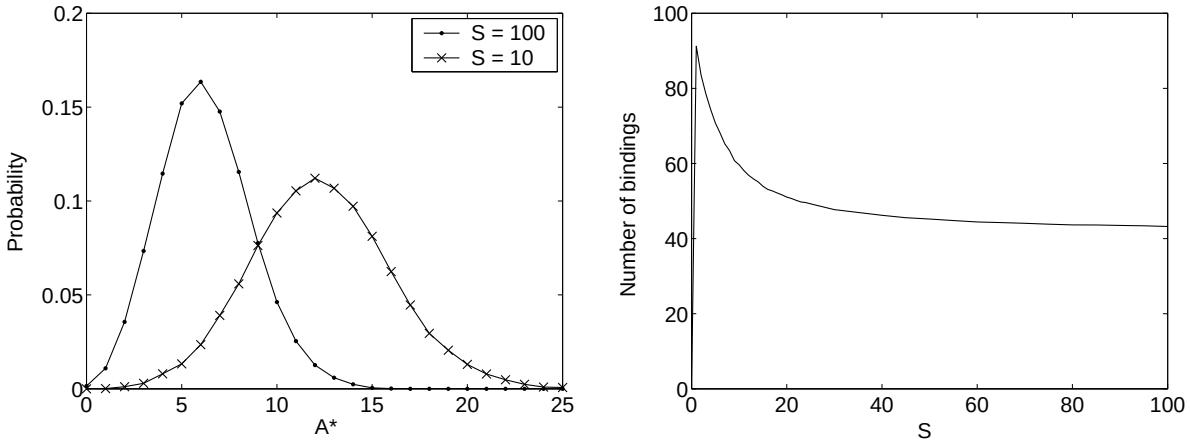
for the number of active receptor molecules A^*. As expected, higher stimulus levels lead to higher numbers of activated receptors. At the same time, this increases the A^* degradation rate, thereby leading to the faster elimination of an active signal. As shown in Fig. 2.6(a), at time point $t = 10\,\text{min}$, the average number of signalling molecules A^* is 12 for $S = 10$ and only 6 for $S = 100$. The probability distribution becomes narrower as its mean decreases.

- What is the expected number of bindings for each stimulus molecule S before degradation occurs? This is given by:

$$\mathcal{R}_{=?}[\mathcal{F}(s = 0 \wedge as = 0 \wedge a^*s = 0)]$$

Predicate $s = 0 \wedge as = 0 \wedge a^*s = 0$ is true if there are no stimulus molecules in the system (in either free or bound form). Reward of $1/N_s$ is associated with every transition labelled by *bind*. The curve in Fig. 2.6(b) peaks at $S = 1$ reaching about 90 bindings per stimulus molecule. As S increases, the number of bindings decreases, levelling out at about 40.

The deterministic analog of the receptor model can be derived using the law of mass action. The ODE system (2.7) states that the rate of change of the concentration $[A]$ is made up of a gain from basal synthesis and dissociation of the $[AS]$ complex, as well as a loss proportional to basal phosphorylation and binding to stimulus. Recall that deterministic reaction rate constants k_m relate to stochastic rate constants c_m by the factor Ω^{1-n}, where Ω is the reaction volume and n is the order of the reaction. Rates of



(a) (b)

Figure 2.6: Properties of the receptor module: (a) The probability distribution of the number of active receptor molecules at time $t = 10$ min; (b) The expected number of bindings per stimulus molecule depending on the stimulus strength.

change of other components are defined similarly:

$$\begin{aligned}
\frac{d[A]}{dt} &= k_{syn} - k_{basalphos}[A] + k_{rel}[AS] - k_{bind}[A][S] \\
\frac{d[AS]}{dt} &= -(k_{basalphos} + k_{phos} + k_{rel})[AS] + k_{bind}[A][S] \\
\frac{d[A^*S]}{dt} &= -(k_{deg} + k_{rel})[A^*S] + (k_{basalphos} + k_{phos})[AS] + k_{bind}[A^*][S] \\
\frac{d[A^*]}{t} &= k_{basalphos}[A] - k_{deg}[A^*] + k_{rel}[A^*S] - k_{bind}[A^*][S] \\
\frac{d[S]}{dt} &= k_{rel}[A^*S] + k_{rel}[AS] - k_{bind}[A][S] - k_{bind}[A^*][S]
\end{aligned} \tag{2.7}$$

The ODE system (2.7) is solved numerically to derive the temporal evolution of the concentration of the activated receptor. As can be seen in Fig. 2.4(e) and 2.4(f), the results match closely the results of the stochastic simulation, after averaging the latter.

Table 2.2 shows the relative importance of the rate constants in determining the maximum amplitude and duration of signalling. Signalling amplitude is regulated primarily by the strength of the stimulus-receptor complex disassociation (c_{rel}), the rate of receptor

Parameter	S^{Amp}	S^{Dur}
c_{bind}	0.005	-0.003
c_{rel}	0.274	-0.067
$c_{basalphos}$	0.057	-0.013
c_{phos}	0.22	-0.1
c_{deg}	-0.557	-0.848
c_{syn}	0.003	0.028

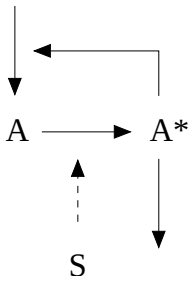
Table 2.2: Sensitivity coefficients of the receptor module calculated for the amplitude Amp and duration Dur of the response component A^*.

phosphorylation (c_{phos}), and the rate of active receptor degradation (c_{deg}). Duration of signalling is controlled almost exclusively by the rate of receptor degradation (c_{deg}). We also observe that the increase of c_{rel} and c_{phos} leads to amplitude upregulation, while the duration is downregulated.

2.4.2 Positive feedback module

To maintain, optimise or adapt to the conditions, a signalling process may use information about its current state, feeding the measurement of the output back into the decision process. Feedback loops are found frequently in biochemical networks. A component of the system is subject to positive feedback if it increases its own level of activity. An important system-level property of positive-feedback systems is the potential to produce bistability. A bistable system is one that switches between two discrete, alternative stable steady states. Bistable systems are thought to be involved in the generation of switch-like responses and the production of self-sustaining biochemical memories of transient stimuli.

The reaction diagram of the previously described receptor module is modified to include the induction of a non-active form of A molecules by the activated response component A^*. Depending on the characteristics of the feedback loop, such a system can exhibit a behaviour qualitatively different from that of a simple linear module. We further study



the properties of positive feedback systems by building and analysing stochastic and deterministic models.

The π-calculus model includes an additional transition (marked by channel `syn`) that leads to the creation of further instances of process `A`. The rate of this transition is proportional to the total number of activated receptors in the system. Because π-calculus models contain only binary communication channels that rely on exponential distribution, there is no direct method for representing more complicated cases of feedback dependency in this formalism.

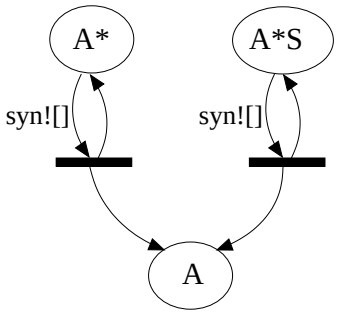


Figure 2.7: Linear positive feedback in π-calculus

A more general case of feedback is represented in PRISM and ODEs. Models from previous sections are modified as shown in Fig. 2.8 and equation (2.8).

$$\frac{d[A]}{dt} = \ldots + \mathbf{f}([A^*] + [A^*S]) \tag{2.8}$$

```
module receptor
...
   //synthesis of A
   [feed] a < N_a & a* > 0 & a*s > 0  →  c_syn * f(a* + a*s)  : (a' = a + 1);
...
endmodule
```

Figure 2.8: General form of feedback in PRISM

We consider two forms of feedback function $\mathbf{f}$: linear

$$\mathbf{f}(x) = c_{syn}x$$

and saturated

$$\mathbf{f}(x) = \frac{c_{syn}x^{c_h}}{c_{syn}x^{c_h} + c_m^{c_h}}.$$

Figure 2.9 demonstrates the transient response of the linear feedback function module for two different values of stimulus strength S. Parameters used in these simulations are as follows: $c_{syn} = 0.005$, $c_{deg} = 0.005$, $c_{bind} = 10/\Omega$, $c_{rel} = 0.1$, $c_{phos} = 0.1$, $c_{basalphos} = 0.0005$, $\#A(0) = 10/\Omega$, $\Omega = 100$. The primary effect of the feedback is prolonged signalling (compare signal decay $t_{1/2}$ =0.2 hr compared to only 0.07 hr in the receptor module). Once the signal is activated, it induces the expression of the feedback protein that helps maintain the activity of the pathway. Feedback proteins cannot, however, maintain signalling activity without continued stimulation.

We next test the case when the dependency of $\mathbf{f}(x)$ on x has a sigmoid shape: activity of feedback is near zero for low values of x. Over a range of x centered about c_m, the activity increases rather steeply to the maximal value of one. The steepness is proportional to c_h (called the Hill coefficient). In Fig. 2.10, we show simulation results of the pathway with saturated feedback (parameters used in simulations are as follows: $c_{syn} = 0.1$, $c_m = 10$,

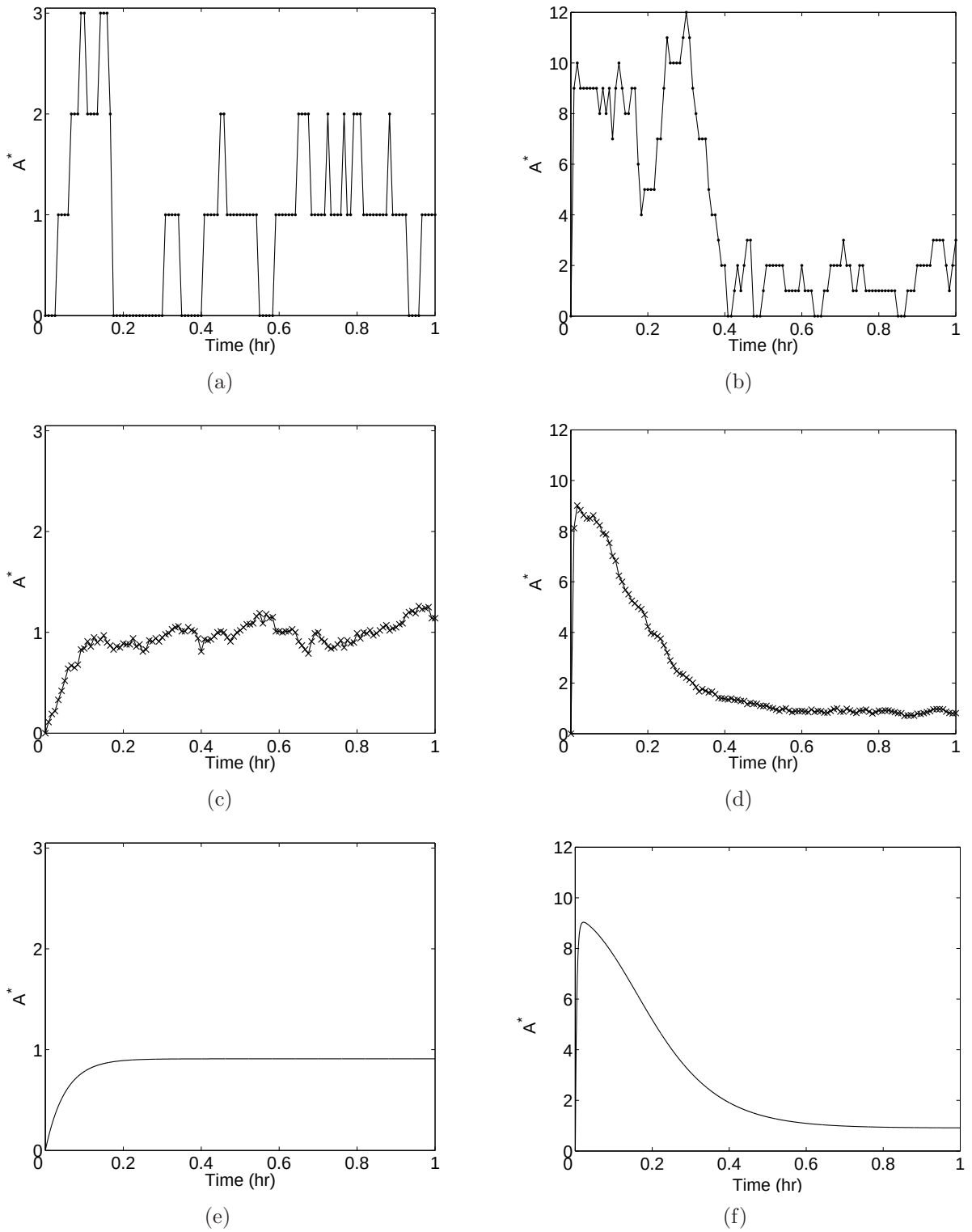


Figure 2.9: Temporal evolution of the linear feedback system: single stochastic run (a and b), average of 100 stochastic runs (c and d) and deterministic solution (e and f) for $\#S(0) = 0$ and $\#S(0) = 0.1 * \Omega$.

$c_h = 3$, and the rest as in Fig. 2.9).

If feedback is a saturating function, the system can produce a bistable response; depending on the stimulus strength, the system stabilises at low (Fig. 2.10(a), 2.10(c) and 2.10(e)) or high values (Fig. 2.10(b), 2.10(d) and 2.10(f)). High stimulus is not only sufficient for initial upregulation of the signalling amplitude, but also for the generation of a sustained response in which the system "memorises" the effect of transient stimuli. Increased sensitivity of the network to changes in the incoming signal is crucial for producing sustained signal.

We also observe that the results of the deterministic and stochastic systems differ. The average value of the amplitude in the long run is 200 nM in case of the deterministic model, and only about 120 nM for the stochastic model. Figure 2.10(b) shows two stochastic simulation runs of the model. In one case, the system evolution closely resembles that of the deterministic system. The signal approaches the value of 200 nM within 2 hr of stimulus induction, and stays in vicinity of this value for the rest of simulation. But inherent noise can also result in complete abolishment of signalling through degradation of initially small number of the activated receptor A^*. Coexistance of these separate phenotypes leads to significantly lower signalling amplitude when averaged over many stochastic runs (Fig. 2.10(d)). Thus, the stochastic approach is more accurate for representing positive feedback systems in which the incoming signal and random fluctuations are amplified.

Using PRISM, we check the properties of the probability distribution of the system state. The following properties were formulated:

- What is the steady state probability distribution of activated signalling molecules A^*? This is expressed as:

$$\mathcal{S}_{=?}[a^* + a^* s = i]$$

  where i is in the range $[0, N]$. Figure 2.11(a) confirms that the A^* probability

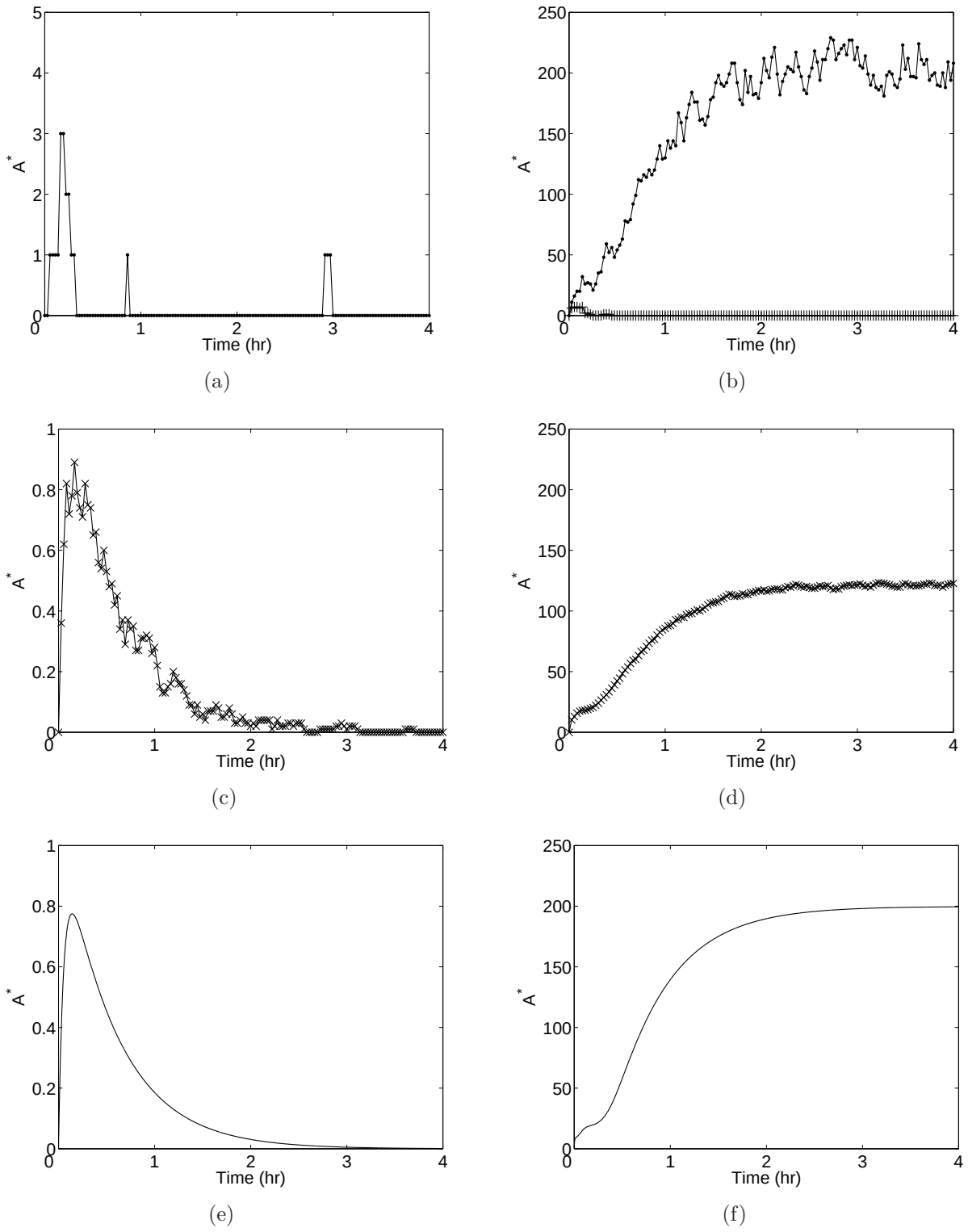
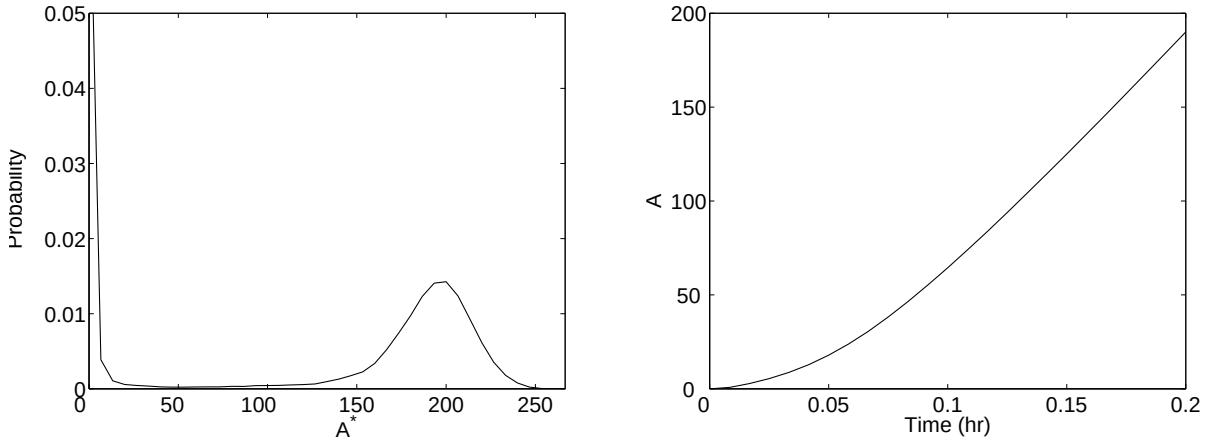


Figure 2.10: Positive feedback and bistability: results of individual stochastic runs (a and b), average of 100 runs (c and d), and deterministic solution (e and f) for $\#S(0) = 0$ and $\#S(0) = 0.1 * \Omega$, respectively.



(a) (b)

Figure 2.11: Probability distribution and the number of synthesised receptors in the positive feedback model.

distribution peaks at both 0 and 200 nM.

- What is the expected number of receptor molecules produced by time t due to the feedback activation? The formula is:

$$\mathcal{R}_{=?}[\mathcal{C} \leq t]$$

with the reward of 1 associated with each transition marked as $feed$.

Table 2.3 shows the relative importance of the rate constants in determining maximum amplitude and duration of signalling. Contrary to the simple activation module in the previous section, the system exhibits significant sensitivity for parameters c_{syn} and c_{deg}. The ability of the system to generate sustained response is controlled by these parameters.

Parameter	S^{Amp}	S^{Dur}
c_{bind}	2×10^{-6}	0.003
c_{rel}	4×10^{-5}	0.06
$c_{basalphos}$	0.001	0.085
c_{phos}	5×10^{-5}	0.066
c_{deg}	-5.25	-5.13
c_{syn}	4.78	5.14

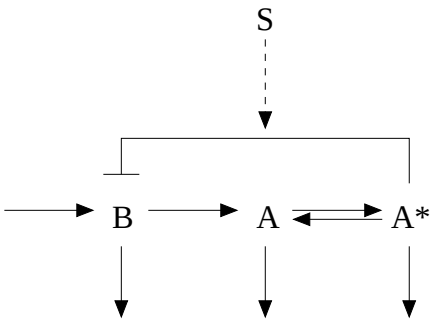
Table 2.3: Sensitivity coefficients of the positive feedback model determined for the amplitude *Amp* and duration *Dur* of A^*.

2.4.3 Negative feedback module

Finally, we consider a system with negative feedback control. Negative feedback is described by the inhibition of the component's activity by itself. While the response of the receptor system changes proportionally to the change in input, the response of the system containing negative feedback is generally not changed significantly with changes of the stimulus. This condition is also known as homeostasis, where the response to an external signal is approximately constant over a wide range of signal strength. Negative feedback generally stabilises the behaviour of the process and makes it resistant to perturbations.

Similar to the examples in previous sections, we consider a reaction system in which molecule A changes between phosphorylated and dephosphorylated states. When phosphorylated, A inhibits a signalling response molecule B. The incoming signal S regulates the ability of A^* to bind and inhibit B, leading to accumulation of B in the cell. To close the negative feedback loop, B activates the transcription of its inhibitor A.
Bar-headed arrows indicate inhibition of system components.

In the stochastic model, the mechanism for inhibition of B follows a three-step model that includes binding of A^* to B, followed by either disassociation of the $A^* : B$ complex or degradation of B while bound to A^*. The stochastic π-calculus model for negative feedback reaction system is shown in Fig. 2.12. Again, molecules A and B are implemented as π-



calculus processes `A` and `B`. They communicate on channel `inhibit` to create a complex which can subsequently dissociate on channel `rel` or eliminate `B` by degradation on channel `deg`. Zero-order synthesis of B, first-order degradation of both A and B, and first-order phosphorylation/dephosphorylation of A are represented similar to previous examples.

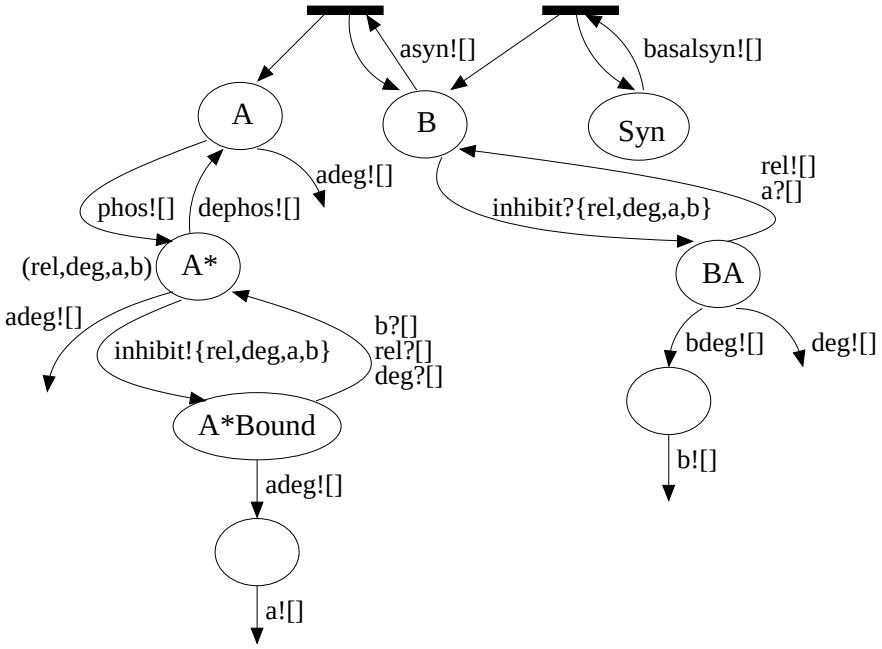


Figure 2.12: Negative feedback module implementation in π-calculus.

In the ODE model, we assume that inactivation of B by a phosphorylated form of A

follows a Michaelis-Menten kinetic model, leading to the following system of equations:

$$\frac{d[A]}{dt} = -(k_{phos} + k_{adeg})[A] + k_{dephos}[A^*] + k_{asyn}[B], \tag{2.9a}$$

$$\frac{d[A^*]}{dt} = k_{phos}[A] - (k_{dephos} + k_{adeg})[A^*], \tag{2.9b}$$

$$\frac{d[B]}{dt} = k_{bsyn} - k_{bdeg}[B] - \frac{V_{\max}^{inhibit}[S][A^*][B]}{K_m^{inhibit} + [B]} \tag{2.9c}$$

where $V_{\max}^{inhibit}$ and $K_m^{inhibit}$ are parameters of the Michaelis-Menten kinetic function.

We run simulations using the following parameter values: $c_{bsyn} = \Omega$, $c_{bdeg} = 0.01$, $c_{asyn} = 0.2$, $c_{adeg} = 0.1$, $c_{inhibit} = 10 * S/\Omega$, $c_{rel} = 1$, $c_{deg} = 5$, $c_{phos} = 0.05$, $c_{dephos} = 0.001$, $\#A(0) = \Omega$, $\#B(0) = 0.1 * \Omega$, and $\Omega = 100$. In Fig. 2.13, we plot the solution for response component B at different values of the incoming stimulus S for the deterministic and stochastic variants of the model. We observe that varying the level of the incoming stimulus induces qualitative changes of system behaviour. At low stimulation, the system reaches the steady state (Fig. 2.13(a) and 2.13(b)). As the stimulus increases, the system displays damped and stable oscillations around the steady state (Fig. 2.13(c) and 2.13(d)). The value of S influences the strength of damping and amplifying and the amplitude of the limit cycle. For a strong enough stimulus, the system reaches a stable limit cycle (Fig. 2.13(e) and 2.13(f)).

Again, we observe that the stochastic system behaves qualitatively differently compared to its deterministic counterpart. In the stochastic model, oscillations are produced at a much wider region of S values. Random fluctuations present in the stochastic system cause the solution to deviate from the steady state. The system then undergoes a long excursion in the phase space before returning to the steady state. The deterministic model does not include random perturbations that initiate a new cycle, and thus the trajectory stays in a steady state.

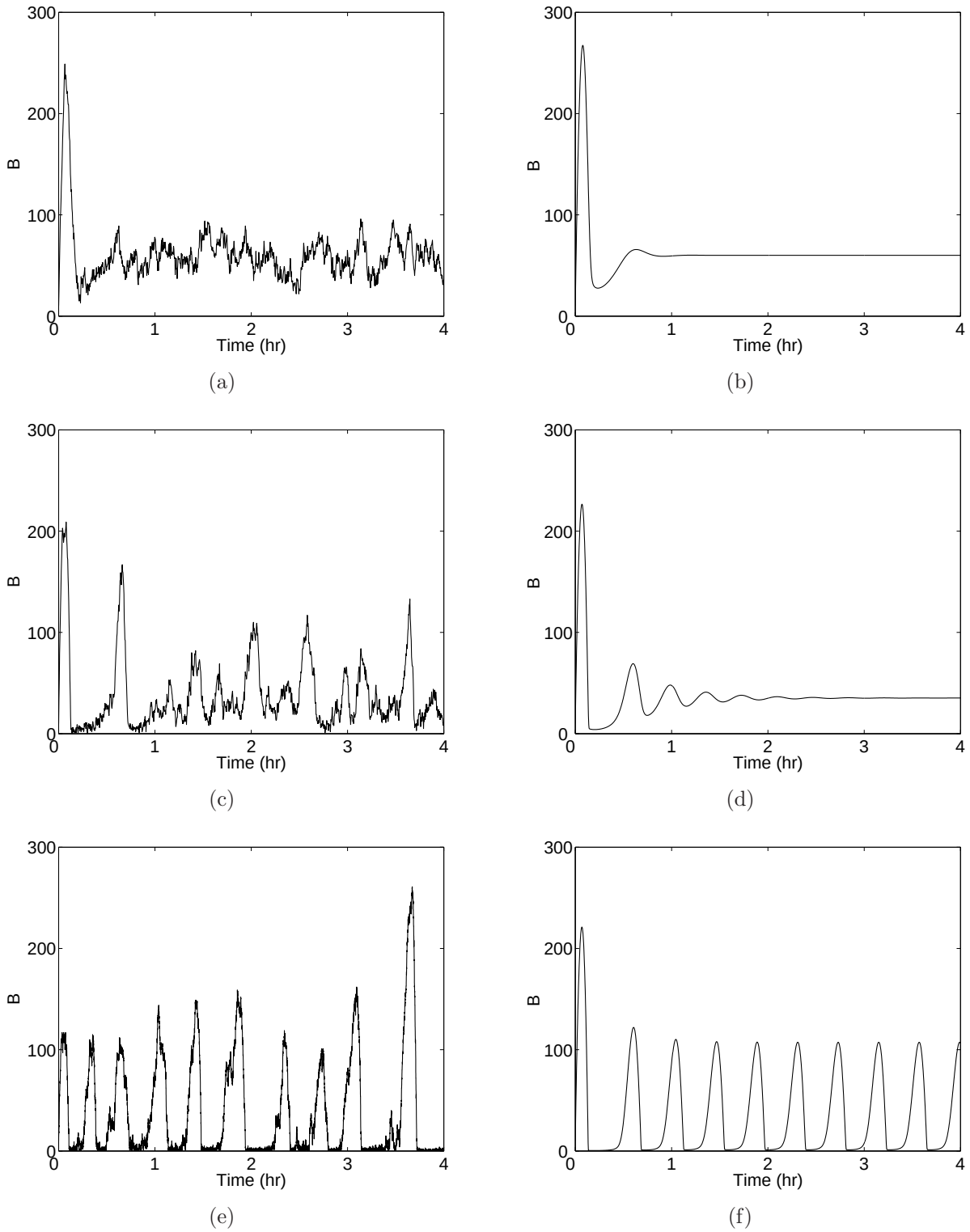


Figure 2.13: Negative feedback and oscillations: results of the π-calculus (a, c, and e) and the ODE (b, d, and f) models at $\#S(0) = 0.01 * \Omega$, $\#S(0) = 0.1 * \Omega$ and $\#S(0) = \Omega$, respectively.

In Fig. 2.14, we plot the temporary evolution of the amplitude and duration of the response component B. In the deterministic model, there is a clear separation between stable (solid line) and oscillating ranges (dashed line) of B amplitude (Fig. 2.14(a)). The amplitude of the oscillations depends on the stimulus S as shown by a dashed curve that corresponds to maximal and minimal values of the amplitude. In the case of the stochastic model, the amplitude shows oscillations at a much broader range of S (Fig. 2.14(b)). As S decreases, oscillations become obliterated with noise. As a result, the duration graph shows lesser variability as stimulus increases (Fig. 2.14(d), compared with the deterministic duration in Fig. 2.14(c)).

Another parameter influencing oscillatory behaviour of the stochastic system is system volume Ω. We expect that at high values of Ω the results of deterministic and stochastic models would be similar. Figure 2.15 illustrates the results of stochastic simulations during which the reaction volume Ω varies from 100 to 1000. At $\Omega = 100$ the number of B molecules varies from 0 to 250 (mean 40). For such small numbers of molecules, oscillations can still be clearly distinguished. Rhythmicity of the limit cycle (Fig. 2.15(e)) becomes obliterated by noise. For a larger reaction volume $\Omega = 1000$, robust oscillations are obtained, in which the number of B molecules is varied in the range of 0-1700. In this case, the effect of molecular noise is to merely induce variability in the maximum of the oscillations (Fig. 2.15(f)). This effect is reflected by the noisy appearance of the limit cycle and a thickening of its upper portion linking the maximum expression in both proteins. The histogram of periods in Fig. 2.15(d) indicates that the average period becomes closer to the deterministic period value, and the standard deviation is decreased greatly, compared to the previous case of smaller reaction volume $\Omega = 100$ (Fig. 2.15(c)).

The sensitivity coefficients of the negative feedback module are shown in Table 2.4.

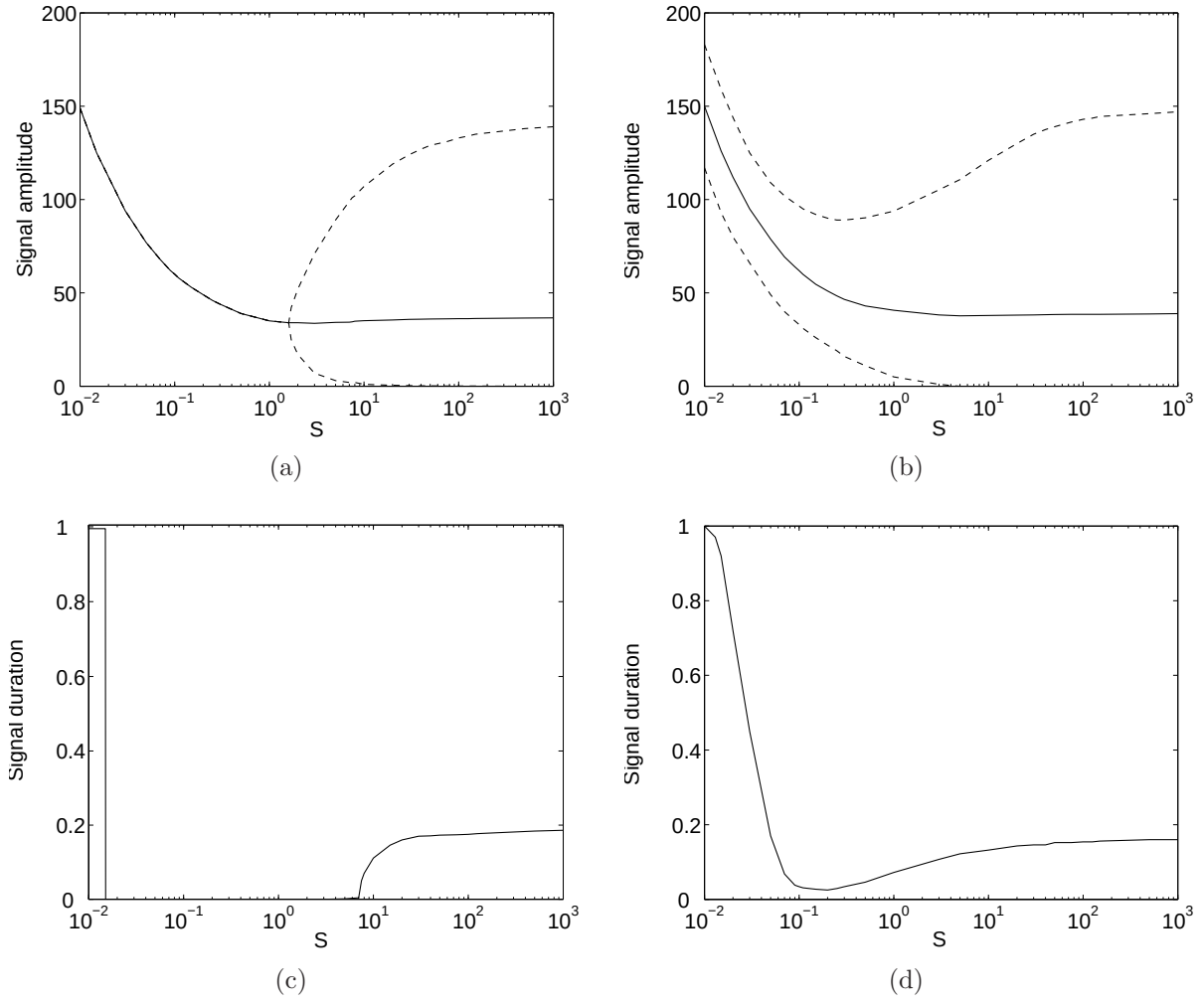


Figure 2.14: Signalling amplitude and duration depending on the value of incoming stimulus S for deterministic (a and c) and stochastic (b and d) systems, respectively.

Parameter	S^{Amp}	S^{Dur}
c_{bsyn}	1	0.12
c_{bdeg}	-0.02	-0.02
$c_{inhibit}$	0.02	0.08
c_{rel}	-0.02	-0.08
c_{deg}	-0.7	-0.1
c_{asyn}	-0.7	-0.13
c_{adeg}	1.5	0.72
c_{phos}	-0.5	-0.02
c_{dephos}	0.008	0.006

Table 2.4: Sensitivity coefficients of the negative feedback system, determined for for the amplitude Amp and duration Dur of B.

Signal amplitude shows dramatically reduced sensitivity to the rates of A:B complex association ($c_{inhibit}$) and B degradation (c_{bdeg}) which are now balanced by the synthesis of A molecules (c_{adeg}).

2.5 Conclusions

We surveyed complementary representations of signalling networks, including continuous deterministic and discrete stochastic models. These were applied to analyse general types of biochemical control modules such as receptor activation, positive and negative feedback modules. We have shown that negative feedback loops can stabilise a system, keep the signal level bounded, and make the system robust against changes in parameter values. Positive feedback systems have the ability to generate a bistable response and allow the system to memorise effects of transient stimuli. Perturbations of such systems were shown to have a dramatic impact on the behaviour of the system. A stochastic description of the system demonstrated the impact of inherent fluctuations on the regulatory system. Stochastic fluctuations were shown to produce qualitative changes in the dynamics.

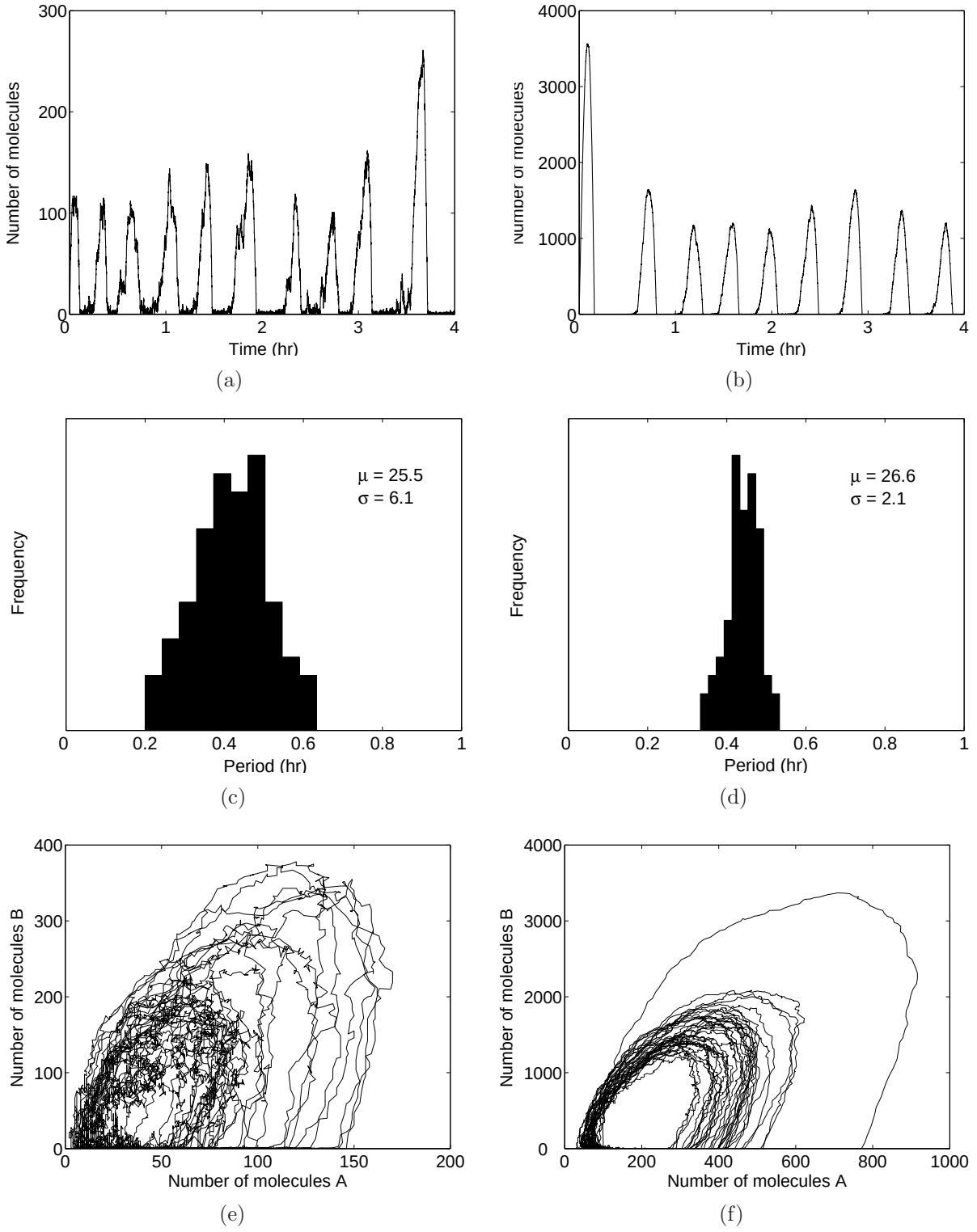


Figure 2.15: Variability of oscillations: (a and b) the number of molecules B, (c and d) the corresponding limit cycle and (e and f) the histogram of the cycle periods for $\Omega = 100$ and $\Omega = 1000$, respectively.

Chapter 3

FGF signalling pathway

3.1 Introduction

In this chapter, we apply the π-calculus approach to model the FGF (Fibroblast growth factor) signalling pathway. We have chosen this pathway because mutations which affect quantitative features of pathway dynamics have been identified both as highly significant in common forms of human cancer [Greenman et al., 2007] and as the underlying cause of developmental skeletal dysmorphology syndromes [Kan et al., 2002]. It is not, however, immediately obvious how these mutations might lead to pathogenic outcomes. Our goal is to use computational modelling supported by reasoning to characterise key parameters that shape pathway dynamics, which, in turn, can influence the prioritisation of biological experiments and development of effective therapeutic interventions.

3.1.1 FGF biology

The binding of FGF ligand to its receptor triggers the sequence of downstream events which culminates in the activation of the Ras-MAPK (ERK) pathway. This proceeds

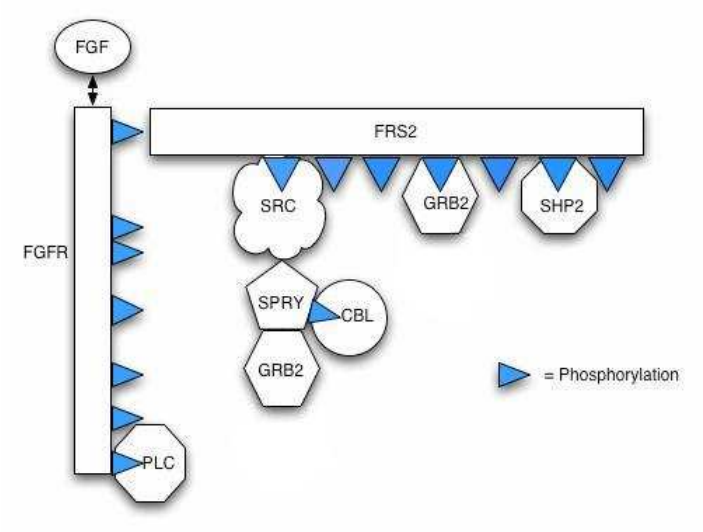


Figure 3.1: Diagram of the possible bindings of the FGF receptor.

via binding of the scaffold protein FRS2 to the activated receptor. FRS2 subsequently phosphorylates and recruits multiple protein molecules, including Grb2, Src, and Shp2. Figure 3.1 shows different communication partners of the FGFR:FRS2 complex. Binding of Grb2 is followed by Sos binding and subsequent activation of the ERK pathway. Other binding partners of FRS2 act to downregulate FGF signal transduction by ubiquitination, internalisation or degradation. Src binding to FRS2 is reported to attenuate the signal through internalisation of the active receptor complex. Spry is a direct physiological substrate of Src which also antagonises FGF signalling through Cbl-dependent ubiquitination and subsequent degradation of the receptor complex. It is also known that Shp2 induces dephosphorylation of the FRS2 docking sites, thus preventing them from binding their respective partner molecules.

In this chapter, we develop and analyse a family of stochastic π-calculus models of the FGF pathway. Simulation results for the derived models closely resemble the known dynamic behaviour of the system under normal physiological conditions. By interrogation

of parameter space and component dependencies (*in silico* genetics) we show that the behaviour of the pathway is dominated by two inter-linked variables: the rate of receptor kinase activation and the rate of signal attenuation by receptor complex internalisation. We confirm these findings in living cells by experimental intervention, which reveals that the interplay between receptor activation and attenuation exhibits counter-intuitive features, and that pathogenic mutations in FGF receptors lead to sustained signalling properties which are accentuated at low levels of receptor occupancy, thereby yielding novel insight into the function of oncogenic mutations.

3.2 Related work

Different subnetworks of the FGF and related signalling networks were the subject of previous modelling studies [Yamada et al., 2004, Schoeberl et al., 2002]. A comprehensive model of the EGF pathway which shares most of its components with FGF was developed in [Schoeberl et al., 2002]. This paper has applied the law of mass action to derive an ODE model of the pathway. The model was then analysed to infer changes over time in pathway components following stimulation with the growth factor. The authors concluded that growth factor concentration, as well as the affinity of ligand-receptor interactions are the main determinants of the strength of the signal activation. The model of [Schoeberl et al., 2002] comprises 103 species and 83 kinetic parameters. The large dimension of the model makes it difficult to manipulate the model and to gain an understanding of signalling dynamics beyond experimentally known and observable effects.

Models of the FGF and EGF signalling pathways, as well as the differences between them, were analysed in [Yamada et al., 2004]. Each model comprises around 40 ODEs. The authors try to avoid an exponential explosion problem by allowing only pre-ordered sequential association of several molecules into a complex. The effect of such simplifications

on the model behaviour is not considered there. Analysis indicates that a scaffold protein FRS2 which participates in the FGF, but not the EGF pathway, plays a key role in amplifying the signal and generating a sustained signal activation. This generally known effect of the scaffold protein is reproduced only by overexpression of the receptor. The increase of the level of FGF ligand is also predicted to result in the rapid increase of the signalling amplitude.

Unlike models of [Yamada et al., 2004, Schoeberl et al., 2002], the FGF pathway model developed in this chapter is easily manipulable and thus allows us to reason about the effect of different pathway components under varying hypothesised scenarios. The model is also suitable for determining the effect of stochasticity and random perturbations on the signalling output.

3.3 FGF model design

3.3.1 Components and interactions of the FGF pathway

We derive the FGF model based upon information from literature on the early steps of FGF signal propagation. In particular, we incorporate several features in the model which have been reported to negatively regulate FGF signal propagation (reviewed [Dikic and Giordano, 2003, Tsang and Dawid, 2004]).

The reactions encoded in the model are as follows (shown in Fig. 3.2):

1. FGF ligand binds to the FGF receptor (FGFRs) creating a complex of two FGFRs and two FGF ligands.

2. The existence of an FGFR dimer leads to phosphorylation of FGFRs on two residues Y653 and Y654 in the activation loop of the receptor. Mutagenesis and structural

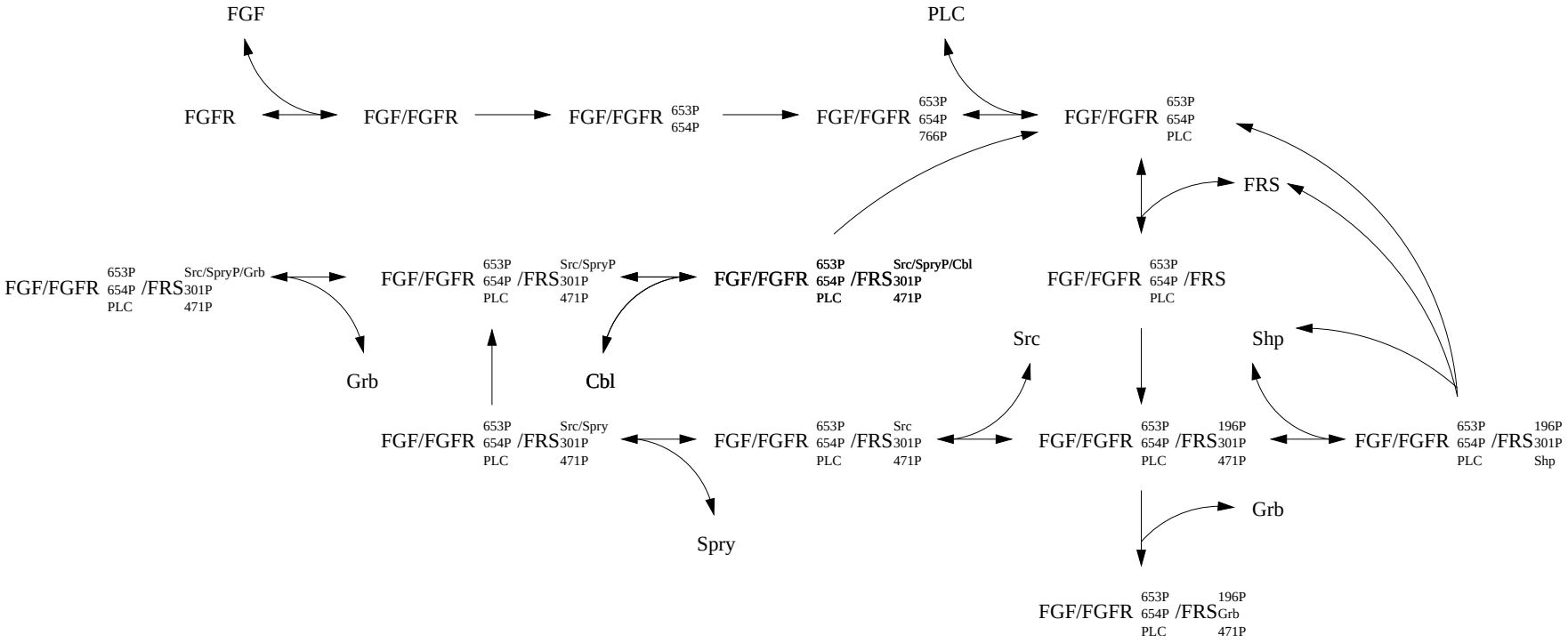


Figure 3.2: Schematic representation of molecular interactions in the FGF pathway

studies [Mohammadi et al., 1997, Furdui et al., 2006] have shown that phosphorylation of these residues is required for activation of FGFR kinase activity and phosphorylation of other substrates.

3. The dual Y653/654 form of the receptor leads to phosphorylation of other FGFR receptor residues (Y663, Y583, Y585, Y766) which have been shown in a number of studies to be required for execution of FGFR dependent signalling functions [Foehr et al., 2001]. In this thesis, we only consider Y766 further.

4. FRS2 binds both phosphorylated and dephosphorylated forms of the FGFR. FRS2 has been shown in multiple studies to be an essential mediator of FGFR functions as a consequence of recruitment of effectors to specific phosphorylated sites on FRS2 [Xu and Goldfarb, 2001].

5. The dual Y653/654 form of the receptor leads to phosphorylation of the FGFR substrate FRS2.

6. We incorporate a step in which FRS2 is dephosphorylated by a phosphatase (denoted Shp2). Shp2 has been shown experimentally to be a negative regulator of FRS2 functions [Hadari et al., 1998].

7. A number of effector proteins interact with the phosphorylated form of FRS2. In this model we include Src binding to residue FRS2 Y196, Grb2:Sos binding to FRS2 Y306, and Shp2 binding to FRS2 Y471 [Schlessinger, 2004, Li et al., 2004].

8. We incorporate signal attenuation by relocation (i.e., endocytosis and/or degradation) of FGFR:FRS2 when Src is activated by association with the phosphorylated FRS2 Y219.

9. Another method of attenuating signal propagation is by PLCgamma bound to Y766 of FGFR leading to relocation/degradation of FGFR [Sorokin et al., 1994].

10. The signal attenuator Sprouty is a known inhibitor of FGFR signalling and is synthesised in response to signal activation [Hanafusa et al., 2002]. Here we include a variable to regulate the arrival of Sprouty into the system in a time dependent manner.

11. We incorporate the association of Sprouty with Src and concomitant phosphorylation of Sprouty residue Y55 [Li et al., 2004].

12. The Y55 phosphorylated form of Sprouty binds Cbl, which leads to ubiquitin modification of FRS2 and a decrease in FRS2 concentration by ubiquitin mediated proteolysis [Fong et al., 2003].

13. Y55P form of Sprouty is dephosphorylated by Shp2 bound to FRS2.

14. Sprouty Y55P competes with FRS2 for binding Grb2 as has been suggested from some studies in the literature [Hanafusa et al., 2002].

We note the difficulty of representing the pathway graphically, due to exponential explosion of the number of molecular states. Molecules FGFR and FRS2 comprise independent residues that interact with different partner molecules. For example, phosphorylation of FRS2 residues enables FRS2 interaction with Src, Grb2, and Shp2, or any combination of these (e.g., binding of Src does not influence the ability of the molecule to bind Grb2 or Shp2). However, since the traditional graphical pathway map depicts the global state of FRS2, exponential explosion occurs when trying to display all combinations of FRS2 association with its partners. Figure 3.2 is therefore not an exact representation of pathway reactions; rather, it helps the reader to understand the textual description provided earlier.

3.3.2 Choice of parameter values

The values of rates of the FGF pathway reactions 1-14 were assembled based on the literature (Table 3.1). The stochastic π-calculus assumes exponentially distributed reaction rates (with the mean value in the Table 3.1); this is justified since, if collision times are small compared to the times between collisions, molecules are moving chaotically, and a constant ratio of overall collisions leads to reactions.

The model is idealised in that it does not take into account variations in composition, affinities or rate constants that might occur in different cell types or physiological conditions. We explore the ability of the computational modelling approach based on π-calculus to accommodate future quantitative or qualitative modifications to the core model. We explicitly address this issue by evaluation of parameter dependencies below.

3.3.3 π-calculus model of FGF signalling

Figure 3.3 shows how the process representing multi-functional molecule FRS2 is implemented in π-calculus. The process `FRS2` contains four independent subprocesses representing the state of:

- association with FGFR;
- activation of residue Y196 and its association with Src;
- activation of residue Y306 and its association with Grb2;
- activation of residue Y417 and its association with Shp2.

State transitions of the FGFR subprocess in Fig. 3.3 correspond to the FRS2 interactions with the FGF receptor. Since both phosphorylated and unphosphorylated forms of receptor bind FRS2, the transition on channel `bindFGFR` does not have any preconditions. Once

bound to FGFR, FRS2 can disassociate by transition on channel `relFGFR`. If the bound receptor is phosphorylated, it informs its FRS2 partner about this state change on channel `preFGFR`. Binding to the phosphorylated form of the receptor enables phosphorylation of the rest of FRS2 components: Y196, Y306 and Y471. This is modelled by the message on channel `preFRS` being sent from FGFR-bound residue to others. This message is only delivered when FRS2 residues are not phosphorylated.

Once FRS2 residues are informed about the possibility to phosphorylate, they proceed with phosphorylation on the channel `phFRS`. Phosphorylated forms of FRS2 Y196, Y306 and Y471 can bind their respective partner molecules: Src, Grb2 and Shp2. Whenever FRS2 Y471 is bound to Shp2, it can trigger dephosphorylation of either of three residues Y196, Y306 or Y471. These transitions are modelled by process interactions on channel `dphFRS`. Dephosphorylation is immediately followed by dissociation, if respective residues are bound to their communication partners.

FRS2 association with and activation of Src leads to internalisation of FRS2 together with all bound components. We initially model internalisation as termination of signalling by which the activated complex disappears. This is shown as process interaction on channel `degSrc`, followed by notification message `deg` to other residues of FRS2 after which Src-bound residue degrades. When FRS2 residue receives a message on `deg`, it immediately notifies its communication partner and degrades. Similarly, degradation messages are sent when FRS2 is ubiquitinated through interaction with Cbl.

Notice how the π-calculus model avoids exponential explosion of the state space. Rather than specifying a global state of the FRS2 molecule, we use the parallel composition of molecular residues, each of which derives values from its local state space. For example, binding between FRS2 and Src is not affected by the state of association with either Grb2 or Shp2. The complete stochastic π-calculus model of the FGF pathway is

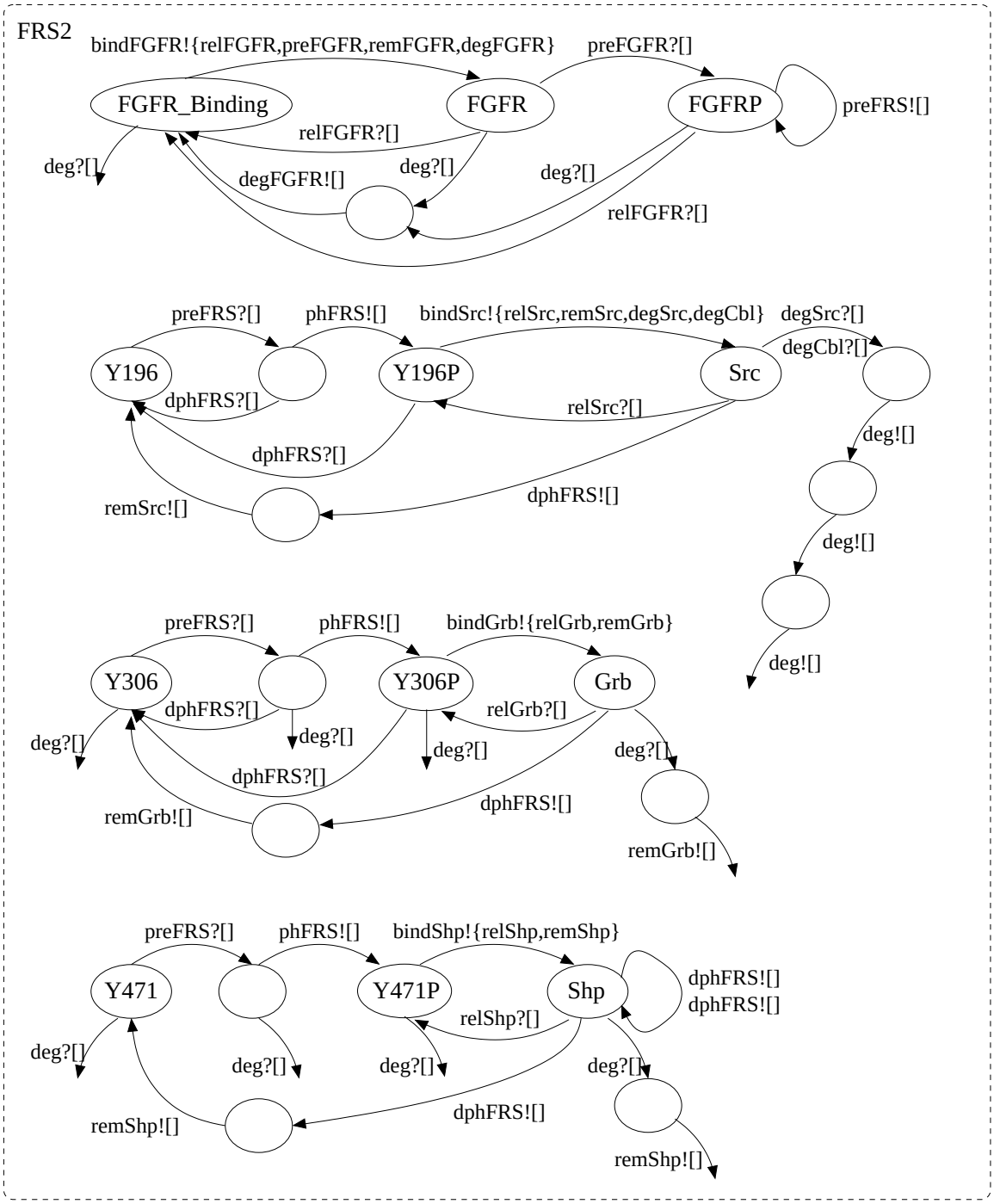


Figure 3.3: Representation of FRS2 molecule in π-calculus.

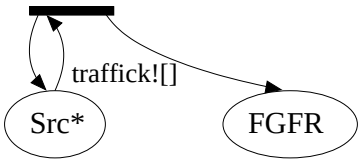


Figure 3.4: Src-induced receptor recruitment in π-calculus.

available in the Appendix B.

Alternative network architectures

The mechanism by which activated FGF receptors are internalised inside the cell is not completely understood. Another important aspect of our modelling study is therefore to explore the behaviour of the pathway under different hypothesised scenarios of internalisation. In our initial model, Src encoded an abstraction of internalisation. Following the conventional view of internalisation as termination of signalling, we assumed that the internalised receptor complex disappears without specifying its subsequent fate.

Recent experimental observations of our group [Sandilands et al., 2007] and others (reviewed [Clague and Urbe, 2001, Miaczynska et al., 2004]) suggest both negative and positive effects of receptor internalisation on signalling. Src exerts additional positive control on the initial phase of signalling by increasing the rate at which FGFR is recycled back to the membrane to re-engage with ligand. Initially, a low number of receptors is capable of transducing a low-level signal. After the signalling network becomes activated, activated Src increases the number of receptors available for signalling. In Fig. 3.4, we show the modification of the signalling network architecture that incorporates a Src-mediated positive feedback loop (FGFR $\rightarrow$ Src $\rightarrow$ FGFR).

3.3.4 ODE model of the FGF pathway

In addition, we develop an ODE model of the FGF signalling pathway. The key difficulty with using an ODE framework for representing the FGF signalling pathway is combinatorial explosion of the number of equations. This results in more that 370 equations required to represent all possible states of the receptor complex in Fig. 3.1. This also makes a detailed simulation study difficult. The average time needed for its simulation varies between 4 and 42 hours, compared to 2-15 min for the stochastic π-calculus model of the same signalling mechanism. The complexity of the deterministic model does not allow one to apply analytic techniques to its analysis (see [Gaffney et al., 2007] for a description of simplification techniques used to reduce the size of the model, not used here). This significantly limits the ability to manipulate and reason about the behaviour of the model.

The ODE model of the FGF signalling pathway is given on the supporting website [FGF, 2008]. Simulations of the model under normal physiological conditions closely resemble results from the stochastic model. Limited by the time and size complexity of the ODE model, we chose to perform systematic analysis of the pathway dynamics using the stochastic framework only.

3.4 Biological implications of the model

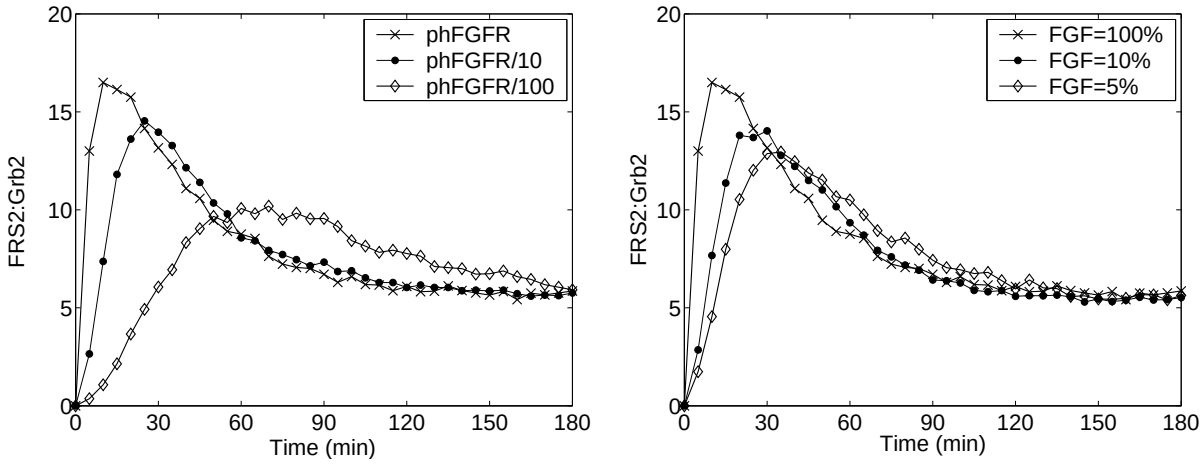
We perform simulations of the derived stochastic π-calculus model starting with the initial concentrations of FGF, unphosphorylated and unbound FGFR, unphosphorylated FRS2, Src, Grb2, Shp2, and Plc equal to 100nM. The model accurately predicts the behaviour of the FGF pathway for known conditions described in the literature. The concentration of the signalling response component FRS2:Grb2 shows a rapid increase shortly after expo-

sure to FGF ligand stimulation, reaching its maximum level at about 10 min. Activation of the negative feedback loops (steps 8, 9, 12, and 14) results in signal downregulation after its successful transduction, thus preventing sustained pathway activation.

We interrogate the model in two ways: by exploration of parameter space to establish the key variables and by removing individual components to study their role in signal propagation. These steps are explicitly designed to build *in silico* experiments, develop new biological hypotheses, and test them by *in vitro* experiments that provide further modelling input.

A detailed approach for understanding the parameter space is to study the effect of parameter variations on the main features of the FRS2:Grb2 concentration profile. We consider the impact of kinetic parameters on two dynamic characteristics of the pathway response: the maximum amplitude and the duration of the FRS2:Grb2 activation. Our sensitivity analysis indicates that the pathway is organised in a way that decouples kinetic parameters that influence the signalling amplitude and duration. We found that the duration was most sensitive to the endocytosis reaction, as shown in Fig. 3.6(a), which is modelled as the decay of the receptor complex bound to Src. In contrast, the amplitude was found to be most sensitive to the rate of FGFR phosphorylation (Fig. 3.5(a)) and the level of FGF ligand (Fig. 3.5(b)). The decoupled control provides more flexibility of the pathway to control its output, which is crucial as the FGF pathway regulates a large variety of cellular processes.

The responses of the average signalling amplitude and duration upon changes in the rates of individual reactions are shown in Table 3.2. In most cases, the resulting deviations of the signalling responses do not exceed 5% of the changes of the respective parameters, proving that the system is robust against parameter perturbations. The only sensitive parameters are those controlling Src-mediated relocation of the receptor complex and



(a) (b)

Figure 3.5: Varying FGFR kinase activity and FGF levels: (a) 10- and 100-fold decrease of FGFR kinase rate compared with the default rate; (b) 100%, 10% and 5% of the FGF level.

FGFR kinase activation.

Reducing the rate constant of FGFR kinase activity leads to the reduction of maximal value of the signal, which now occurs after considerable delay. Fig. 3.5(a) shows the outcome of the 10- and 100-fold reduction of the rate of FGFR phosphorylation. Maximal signal expression reaches 90 and 60% of its initial value and is delayed by 15 and 40 min, respectively. A similar signal reduction and delay occurs in simulations in which the concentration of FGF was reduced to 10 and 5% (Fig. 3.5(b)).

3.4.1 In silico mutagenesis predicts the roles of pathway components

The model reactions 8, 9, 12, and 14 incorporate different negative regulation mechanisms that inhibit response after initial signal activation. The consequences of these signal attenuation mechanisms are evaluated by *in silico* mutagenesis: a 10-fold reduction of the level of various model components before simulation. We observe that inhibition of Sprouty

(Fig. 3.6(a)) does not affect the initial phase of signal upregulation, since it is synthesised after the signal passes its maximum value. Later, Sprouty attenuates signalling, primarily due to the competition for Grb2.

A similar pattern of FRS2:Grb2 expression is generated when Shp2 is not present (Fig. 3.6(a)). When phosphatase Shp2 is inhibited, the initial signalling response reaches a greater peak, but in a longer term the signal stabilises at lower level.

Another key determinant of FGFR signalling dynamics in the model is the rate of internalisation, which we have encoded (perhaps speculatively) by the action of Src. As Table 2.4 and Fig. 3.6(a) demonstrate, the suppression of Src activity is predicted to have a major impact on signalling dynamics; after a fast increase, the signal fails to decrease substantially. This suggests that, in particular, other negative feedback mechanisms are not sufficient to reduce the signal if internalisation is abolished.

In sum, parameter variation and component removal studies of the model indicate that the overall dynamics of this system is dominated by two key variables: the rate of FGFR activation - as judged by varying the concentration of FGF ligand or the rate constants of FGF receptor phosphorylation - and the rate of endocytosis modelled by the recruitment and activation of Src.

3.4.2 Dual role of Src in signal regulation: from model to experiments and back

Prolonged signalling as a result of the suppression of Src activity is an important and previously unrecognised prediction for the regulation of FGF signalling response which emerged from the computational model. To test whether this prediction is valid *in vitro*, we utilised an experimental model of activation of the Ras-MAPK (ERK) pathway in mouse embryonic fibroblasts (MEFs) stimulated with FGF2, using quantitation of phospho-ERK

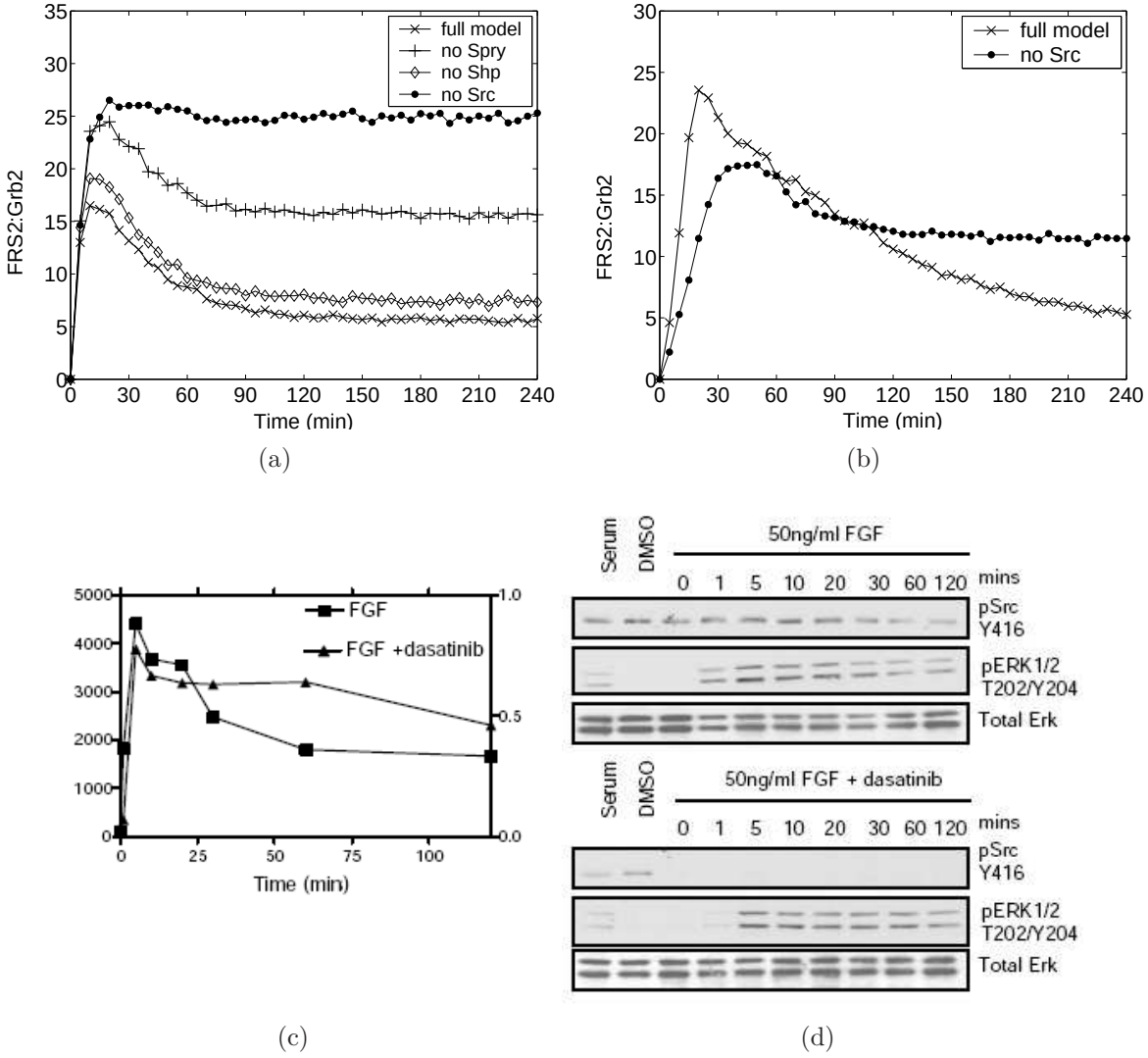


Figure 3.6: In silico mutagenesis: (a) simulations with inhibited Spry, Shp2, and Src compared to the model containing all these components; (b) inhibition of Src in the revised model; (c and d) experimental confirmation of predictions with inhibited Src (experiment was not repeated, see text for explanation).

(pERK) by western blotting as the experimental readout. This approach assumes that the level of experimental pERK is a faithful surrogate for the computed value of the Grb2:FRS2 used in the modelling studies. We take this assumption to hold, as the quantitative behaviour of this output closely resembles the computed value in experimental reports and our own investigations.

The prediction we addressed was that inhibition of Src kinase activity would lead to prolongation of signalling as a consequence of Src acting to remove the activated receptor complex. Quiescent MEFs were stimulated with 50ng/ml FGF2 for varying time points in the presence or absence of the highly specific Src family kinase inhibitor Dasatinib, harvested and examined for the presence of pERK. These results clearly show that untreated MEFs exhibit an FGF stimulus response that conforms to the predicted kinetics, and that pharmacological suppression of Src kinase activity (western blots in Fig. 3.6(d) which are quantified in Fig. 3.6(c)) indeed produces the predicted extended duration of pERK activation (Fig. 3.6(a)). Note that data points in Fig. 3.6(c) are not supplied with error bars because the experiment was not repeated. However, similar results showing Src-dependent extended signalling are presented in [Sandilands et al., 2007].

The outcome of these experimental studies also shows the decreased signal which happens after a delay when Src is inhibited. Therefore, we revise the model by incorporating a positive feedback loop, in which initial low-level activation of FGFR leads to activation of Src which subsequently recruits more receptors to the membrane to be engaged in signalling (FGFR $\rightarrow$ Src $\rightarrow$ FGFR). Simulations of the revised model demonstrate a delayed signal which reaches lower amplitude but fails to attenuate when Src is inhibited (Fig. 3.6(b)), consistent with the experimental data (Fig. 3.6(c)). Through reiteration of modelling and experiments we thus uncovered a positive feedback loop activated at the initial phase of signalling.

3.4.3 Cross-regulation between receptor activation and attenuation

Mutations in FGFR which increase the affinity of receptor-ligand interaction were recently identified as pathogenic *driver* mutations which cause cancer cells to grow [Greenman et al., 2007]. Thus we utilise the computational model of the FGF pathway to determine whether changing the strength of ligand binding to the receptor affects signalling dynamics. The initial model, which does not account for a positive Src regulation, is not able to produce a phenotypic difference between normal and mutant signalling (Fig. 3.7(a) and also shown as the related sensitivity coefficients $S << 1$ in Table 3.2). Surprisingly, the revised model predicts that the pathway amplitude is upregulated in the long run due to recycling of more receptors to the membrane in the mutant (Fig. 3.7(b)). Increased ligand binding shifts the equilibrium between activation and attenuation roles of Src which results in delayed signal attenuation.

We addressed the question whether increasing the rate of receptor activation would lead to accelerated activation and delayed attenuation of signalling *in vitro*. For this purpose we employed a matched pair of MEFs: one derived from normal mouse embryos and the second derived from embryos harbouring the mutant Pro252Arg form of FGFR1 [Hajihosseini et al., 2004]. The MEFs were rendered quiescent by serum starvation and then stimulated for varying time points with 0.1 ng/ml FGF2 and harvested for analysis as in the previous experiment. The result (Fig. 3.7(c) and 3.7(d)) reveals that the introduction of a single gene copy of the mutant driver form of FGFR1 has a marked effect on signalling upon exposure to limiting concentrations of ligand. In the mutant cells peak amplitude is reached rapidly (∼5 min) compared to the wild type (∼30 min) and signal duration is prolonged in the mutant cells, as the model predicts (Fig. 3.7(b)).

Decreasing the rate of FGFR activation by simulating the effects of FGFR kinase

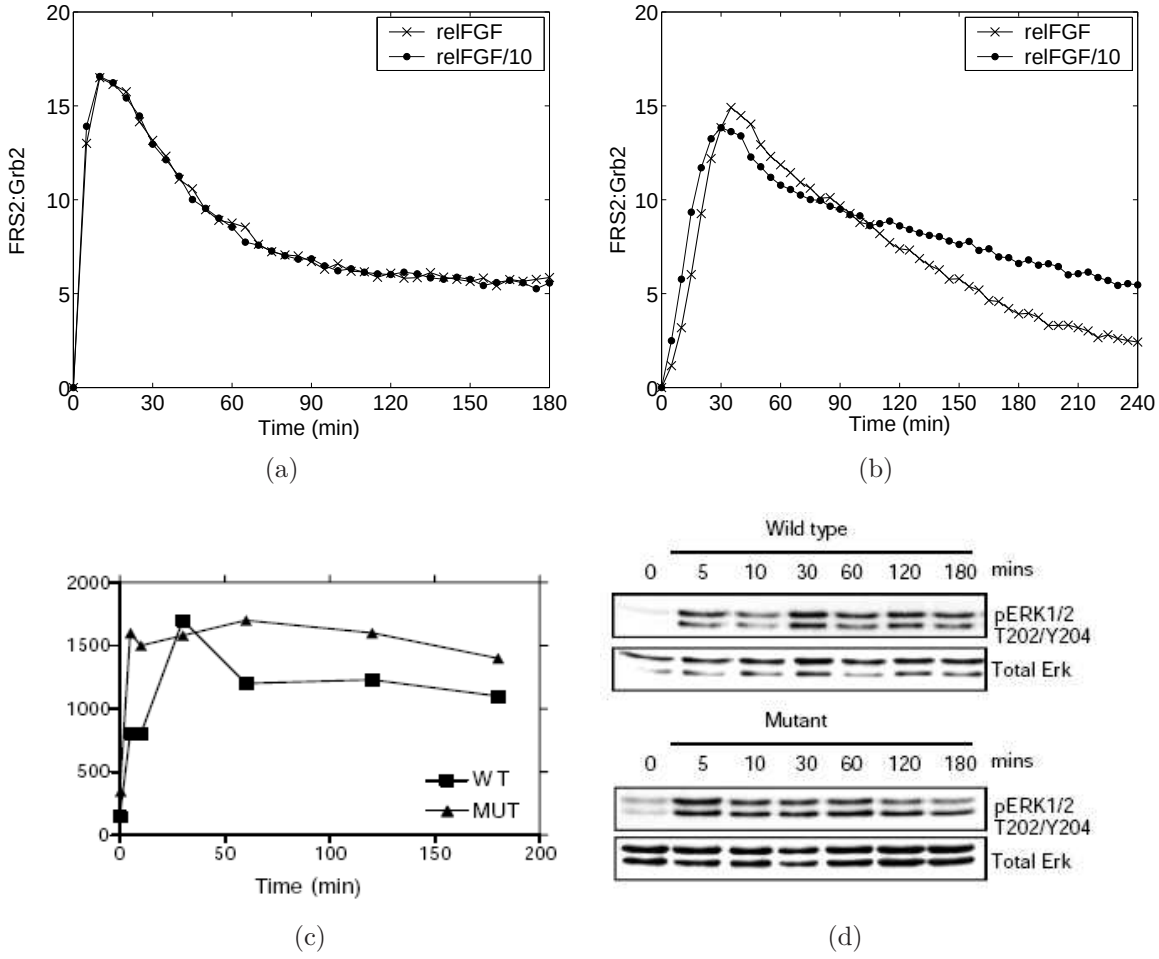


Figure 3.7: Effects of FGFR mutations: simulations with 10-fold inhibition of FGF:FGFR disassociation rate compared to the normal rate in the initial (a) and revised (b) model; (c and d) experimental validation of the phenotypic changes in mutant. Experiments were repeated twice (thus no error bars) with consistent results.

inhibitors accords well with many published studies and should come as no surprise. We showed, however, that increasing the rate of FGFR kinase activation, by which we simulate the effect of pathological mutations in the FGF receptor associated with skeletal development syndromes [Anderson et al., 1998, Hajihosseini et al., 2004] and cancer [Greenman et al., 2007] - leads to an extended duration of signalling which is not overcome by the action of negative regulators such as Sprouty or degradation of the receptor complex.

3.5 Conclusions

In this chapter, we developed a model of the FGF signalling pathway and carried out quantitative analysis of interactions between pathway components. We identified key parameters that shape pathway dynamics and investigated the roles of several proposed negative regulators of signalling. Informed by these *in silico* studies, we experimentally tested and confirmed two counter-intuitive predictions of the model: that suppression of Src kinase prolongs the duration of FGF signalling and that driver mutations in FGFR exhibit fast activation and slow attenuation in response to low level stimulation.

Reaction	Parameter	Value	References
1	FGF binding/release	$5 \times 10^6\,\mathrm{M^{-1}s^{-1}}$ $5 \times 10^{-3}\,\mathrm{s^{-1}}$	[Felder et al., 1993, Mohammadi et al., 2005],
2	FGFR Y653/654 phosphorylation	$0.013\,\mathrm{s^{-1}}$	[Furdui et al., 2006]
3	FGFR Y766 phosphorylation	$0.004\,\mathrm{s^{-1}}$	[Furdui et al., 2006]
4,7,9	FRS2, Src, Grb2, Shp2, Plc binding/release	$2.5 \times 10^6\,\mathrm{M^{-1}}s^{-1}$ $5 \times 10^{-2}\,\mathrm{s^{-1}}$	[Panayotou et al., 1993, Skolnik et al., 1993]
5	FRS2 phosphorylation	$0.005\,\mathrm{s^{-1}}$	[Furdui et al., 2006]
6,13	FRS2, Spry dephosphorylation	$12\,\mathrm{s^{-1}}$	[Montalibet et al., 2005]
7,11,14,12	Sos, Spry, Cbl binding/release	$10^5\,\mathrm{M^{-1}s^{-1}}$ $10^{-4}\,\mathrm{s^{-1}}$	[Sastry et al., 1995]
8	FRS2:Src relocation ($t_{1/2}$)	15 min	[Ware et al., 1997]
9	FGFR:Plc relocation ($t_{1/2}$)	60 min	[Sorokin et al., 1994]
10	Spry induction	$0.083\,n\mathrm{Ms^{-1}}$	[Hanafusa et al., 2002]
11	Spry phosphorylation	$10\,\mathrm{s^{-1}}$	estimated
12	FRS2 ubiquitination and proteolysis ($t_{1/2}$)	25 min	[Wong et al., 2001]

Table 3.1: Kinetic parameter values of the FGF pathway

Parameter	S_{Amp}	S_{Dur}
FGF binding	0.003	0
FGF release	-0.002	0
FGFR Y653/654 phosphorylation	**0.12**	-0.02
FGFR Y766 phosphorylation	$\sim$0	0
FRS2 binding	-2×10^{-5}	0
FRS2 release	0.008	-0.001
Src binding	-0.001	-0.03
Src release	4×10^{-4}	0.02
Grb2 binding	0.008	0.004
Grb2 release	-0.005	-0.003
Shp2 binding	-3×10^{-4}	0
Shp2 release	$\sim$0	0
Plc binding	$\sim$0	0
Plc release	-2×10^{-5}	0
FRS2 phosphorylation	0.01	-0.01
FRS2 dephosphorylation	2×10^{-4}	0
Spry binding	-0.006	-0.006
Spry release	$\sim$0	0
Cbl binding	0.0003	-0.002
Cbl release	$\sim$0	0
Spry phosphorylation	-0.0007	0
Spry dephosphorylation	0.013	0.011
Src-mediated degradation	**-0.14**	**-0.97**
Plc-mediated degradation	0.0005	0
Spry synthesis	-0.002	-0.004
Spry-induced degradation	-0.02	-0.04

Table 3.2: Sensitivity coefficients for the amplitude *Amp* and duration *Dur* of the signalling response component $FRS2 : Grb2$. **Bold** indicates sensitive parameters (in this particular example, sensitivity threshold of 0.05 is chosen). Values smaller than 10^{-5} are annotated as $\sim$0.

Chapter 4

Wnt signalling

4.1 Introduction

It is now recognised that Wnt signalling is involved in many aspects of embryonic development and homeostatic self-renewal in a number of adult tissues. Germline mutations in the Wnt pathway cause several hereditary diseases, and somatic mutations are associated with cancer of the intestine and a variety of other tissues [Kinzler et al., 1991, Gaspar and Fodde, 2004]. In addition to the involvement of this pathway in cancers, it is also implicated in neurodegenerative diseases, tooth agenesis, and regulation of bone density [Moon et al., 2004]. Additional information is available at the Wnt homepage `http://www.stanford.edu/~rnusse/wntwindow.html`.

4.1.1 Wnt biology

The canonical Wnt signalling pathway includes more than 50 known components, exhibiting a complexity that is still to be appreciated by researchers. In the absence of Wnt signalling, a cytoplasmic destruction complex consisting of Axin and APC proteins, and

glycogen synthase kinase-3 (GSK3), recruits and phosphorylates β-catenin. This leads to interaction of β-catenin with β-transducin-repeat-containing protein (β-TrCP), resulting in the ubiquitination of β-catenin and its degradation by the proteasome. Thus, at a steady state without Wnt signalling, β-catenin is rapidly degraded in the cytoplasm.

The pathway is activated by secreted Wnts which interact with Frizzled receptors and with co-receptors such as low-density lipoprotein receptor-related protein-5 and -6 (LRP5 and 6). Receptor complex activation leads to phosphorylation of the cytoplasmic protein Dishevelled (Dvl). Activated Dvl has been reported to function through binding components of the destruction complex to reduce the function of GSK3 and Axin. This in turn reduces the phosphorylation and degradation of β-catenin, leading to its accumulation in the nucleus. Elevation of β-catenin levels by Wnt signalling leads to binding of β-catenin to the TCF/LEF transcription factor. This starts a transcriptional program that is involved in cell cycle progression, and possibly self-renewal [Reya et al., 2003, Lee et al., 2004]. Constitutive activation of the Wnt/β-catenin pathway has been observed in transformed cells due to inactivating mutations in APC and Axin that reduce β-catenin degradation [Kinzler et al., 1991, Gaspar and Fodde, 2004].

Wnt signalling is auto-regulated at different levels by negative feedback loops, demonstrating the importance of limiting the induction of the target genes (reviewed in [Logan and Nusse, 2004]). By stimulating β-catenin degradation, Axin plays a key role in maintaining β-catenin levels low in normal cells. Interestingly, Axin transcription is induced by β-catenin itself. The resulting negative feedback loop (Axin $\dashv$ β-catenin $\rightarrow$ Axin) guarantees that β-catenin is downregulated after successful signal transduction. Other negative feedback loops include β-catenin-dependent transcription of Naked (Nkd) that directly binds and inhibits Dvl [Rousset et al., 2001], and induction of Dickkopf (Dkk) which is able to bind LRP and thus act as extracellular inhibitor of Wnt

[Gonzalez-Sancho et al., 2005].

The apparent complexity of the Wnt pathway architecture emphasises the need to study molecular interactions and time-dependent changes in response to Wnt, using computational modelling. In this chapter, we employ a number of modelling techniques to discover the properties of pathway dynamic behaviour that emerge from a specific pathway architecture. We analyse the robustness of the system against stochastic noise and mutations. We show that, depending on the stimulus level, the system can exhibit complex oscillatory behaviour that is further amplified by random fluctuations in concentrations of system components. We further validate our predictions with appropriate experiments.

4.2 Related models

We have analysed a model of the Wnt signal transduction proposed previously by [Lee et al., 2003] and its further extensions by [Cho et al., 2006, van Leeuwen et al., 2007, Wawra et al., 2007]. The initial model of [Lee et al., 2003] implements and validates the mechanisms of the destruction complex assembly and activation that culminates in β-catenin degradation in unstimulated cells. Due to unavailability of the experimental data at the time, however, the specifics of Wnt signal transduction across the cell membrane is not included in the model. In particular, the recruitment of Axin to the membrane is omitted. Furthermore, the impact of negative control elements is not analysed. In the current modelling effort, we encode the updated biological knowledge of the Wnt pathway components. We then study the response of this pathway to extracellular stimulus gradient in the tissue, as well as its optimality in the presence of stochastic noise.

Results derived from the differential equation models [Lee et al., 2003, Cho et al., 2006, van Leeuwen et al., 2007, Wawra et al., 2007] rely heavily on the assumption that molecules of each species are present in large numbers and treat their populations as variables vary-

ing on a continuous scale. It has, however, been suggested that members of the β-catenin destruction complex such as Axin may be present at low quantities, varying from a few hundreds down to only a few tens of molecules in each cell [Lee et al., 2003]. The low numbers of reacting species may be responsible for random fluctuations that can destabilise the behaviour of the signalling network.

We use a combination of both continuous deterministic and discrete stochastic approaches to study the molecular basis of cellular response to Wnt. We focus on the oscillatory phenomenon that arises from integration of multiple feedback loops in the Wnt architecture. We then examine qualitative changes of the pathway response which arise due to the inclusion of fluctuations in the description of the system.

4.3 Model design

4.3.1 Reactions included in the model

To build a formal model of the Wnt signalling pathway we identify the set of elementary reactions. The rate constants for these reaction steps, given in Table 4.1, were obtained from literature.

1. Gsk binds Axin leading to its phosphorylation [Ikeda et al., 1998, Willert et al., 1999].

2. Phosphorylated Axin forms a destruction complex with APC which is subsequently phosphorylated [Hart et al., 1998, Kishida et al., 1998].

3. β-catenin binds the destruction complex and is phosphorylated by GSK3 [Choi et al., 2006, Rubinfeld et al., 1996].

4. Phosphorylation of β-catenin leads to its ubiquitination, followed by degradation in the proteasome [Aberle et al., 1997].

5. β-catenin binds the non-active form of APC at a low rate [Choi et al., 2006].

6. Wnt signals by activating the receptor complex containing Frizzled and LRP [Cong et al., 2004].

7. Activated receptor complex recruits Axin and GSK3 resulting in the sequences of phosphorylation steps. As the result, Axin is inhibited and degraded [Mao et al., 2001, Willert et al., 1999, Kikuchi, 1999].

8. Increased cytoplasmic concentration of β-catenin leads to its translocation into the nucleus where it binds TCF to activate target genes [Clevers and van de Wetering, 1997].

9. In a negative feedback loop, Axin is transcribed by the β-catenin:TCF complex [Lustig et al., 2002, Leung et al., 2002].

10. β-catenin is synthesised at a constant rate and is degraded by additional mechanisms independent of the APC:Axin complex.

11. Axin is degraded at a rate dependent on its phosphorylation status [Yamamoto et al., 1999].

4.3.2 Wnt model in π-calculus

We employed the stochastic π-calculus formalism to model the Wnt signalling pathway, and used simulation engine BioSPI to perform analysis of this model. Following the previously described scheme, pathway molecules are mapped into π-calculus processes. Molecular processes transduce a signal by undergoing state transitions. The state of the signalling molecule is characterised by phosphorylation and association with its partner molecules. State transitions are implemented as pair-wise communication of processes

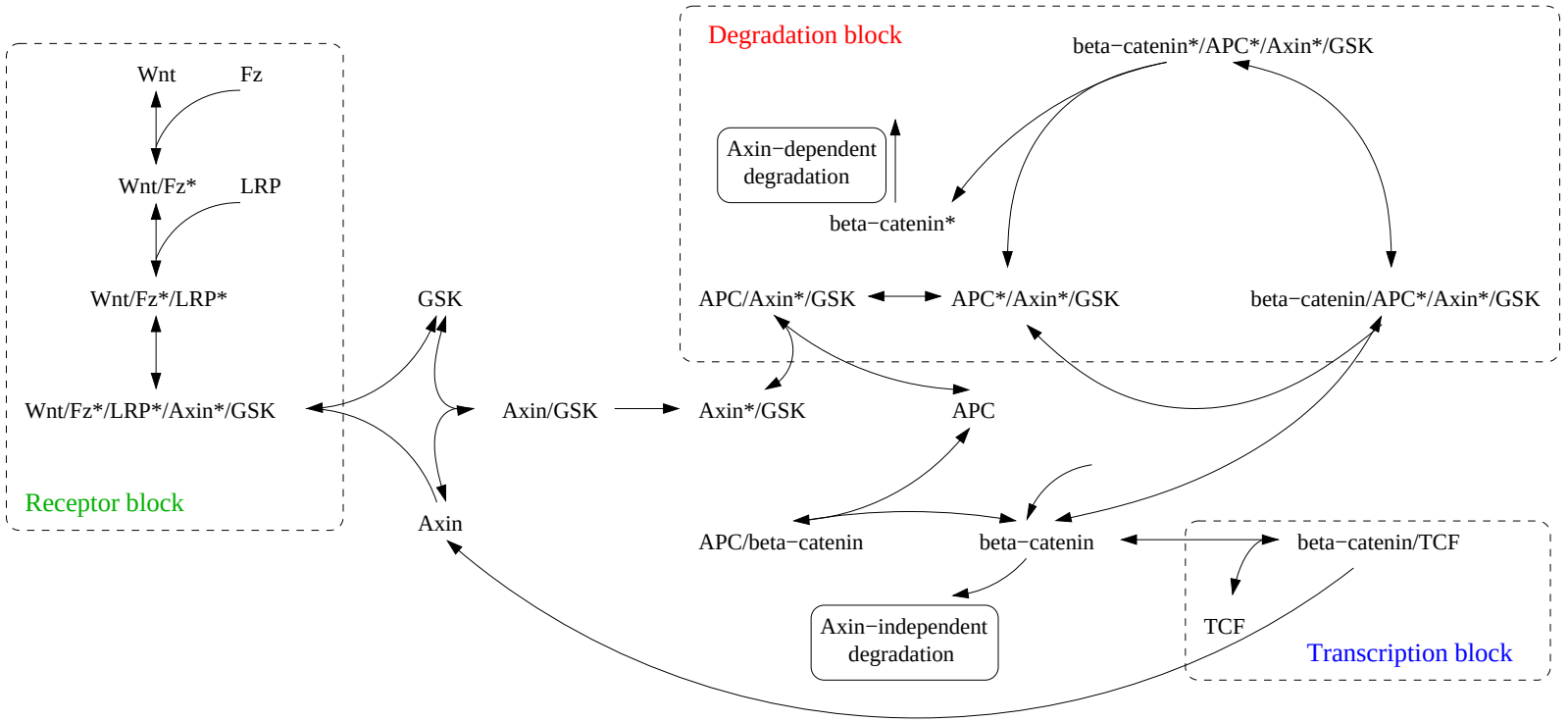


Figure 4.1: Schematic description of molecular reactions of the Wnt pathway

Reaction	Parameter	Value	References
1	Axin binding to Gsk	$65\,n$M	[Ikeda et al., 1998]
1	Axin phosphorylation	$0.03 - 0.2\,\text{min}^{-1}$	[Ikeda et al., 1998]
1	Axin dephosphorylation	$0.1\,\text{min}^{-1}$	[Lee et al., 2003]^{u}
2	APC binding to Axin	$100\,n$M	[Kishida et al., 1998]
2	APC phosphorylation	$0.2\,\text{min}^{-1}$	[Ha et al., 2004]
2	APC dephosphorylation	$0.1\,\text{min}^{-1}$	same as 1
3	β-catenin binding to phosphorylated APC	$10\,n$M	[Ha et al., 2004, Choi et al., 2006]
3	β-catenin phosphorylation	$5\,\text{min}^{-1}$	[Aberle et al., 1997]^{e}
4	β-catenin degradation	$5\,\text{min}^{-1}$	[Aberle et al., 1997]^{e}
5	β-catenin binding to non-active form of APC	$3.1\,\mu$M	[Ha et al., 2004, Choi et al., 2006]
6	Wnt binding to the receptor	$5\,n$M	[Rulifson et al., 2000]
7	Receptor complex binding to Axin:Gsk	$0.5 - 1\,n\text{M}^{-1}\text{min}^{-1}$	[Mao et al., 2001]
7	Wnt-dependent Axin degradation	$0.04\,\text{min}^{-1}$	[Willert et al., 1999, Lustig et al., 2002]^{e}
8	β-catenin:TCF binding	$35\,n$M	[Knapp et al., 2001, Choi et al., 2006]
9	Transcriptional/translational delay	90 min	[Lustig et al., 2002]^{e}
9	Axin transcription	$0.048\,\text{min}^{-1}$	[Lustig et al., 2002]^{e}
10	β-catenin synthesis	$0.423\,n\text{Mmin}^{-1}$	[Lee et al., 2003]^{u}
10	APC independent β-catenin degradation	$0.001\,\text{min}^{-1}$	[Lee et al., 2003]^{u}
11	Axin degradation	$0.005\,\text{min}^{-1}$	[Yamamoto et al., 1999]^{e}
11	AxinP degradation	$0.002\,\text{min}^{-1}$	[Yamamoto et al., 1999]^{e}

Table 4.1: Kinetic parameter values of the Wnt pathway. Superscript$^{\text{e}}$ indicates that values were determined based on time courses presented in respective papers. Superscript$^{\text{u}}$ indicates that we remain uncertain about respective values because the measurements were not reproduced by independent studies.

over channels using a scheme described in Chapter 2. A molecule might contain several independent binding and phosphorylation sites. These are modelled as independent subprocesses within a molecular process.

As a basis for the Wnt model, we use a negative feedback module described in Chapter 2, in which molecules A and B correspond to the Axin:APC destruction complex and β-catenin, respectively. Figure 4.2 depicts the representation of Axin and APC molecules in the calculus. Process `Axin` contains two independent sites which facilitate Axin binding to Gsk and APC, respectively. After the molecule binds Gsk (transition labelled by channel `bindGsk`), it undergoes phosphorylation (channel `phAxin`), which in turn enables the binding of the second Axin site to APC (channel `bindAPC`). Axin:Gsk complex can also be recruited to the membrane in response to Wnt signalling. We represent this as Axin:Gsk binding with Wnt (transition on channel `bindWnt`), followed by either disassociation of the newly created complex (channel `relWnt`), or Wnt-induced Axin degradation (channel `degWnt`). In addition, Axin is degraded at a constant rate (different for phosphorylated and dephosphorylated forms of the molecule) which we represent by transitions on channels `degAxin` and `degAxinP`.

Similarly, we model APC as a process containing independent sites for binding Axin and β-catenin. After association with Axin, APC undergoes phosphorylation which enables its binding to β-catenin at a much higher rate (channels `bindBeta` and `bindBetaP`). When β-catenin is bound to phosphorylated APC, β-catenin itself undergoes phosphorylation (transition on channel `phBeta`) and subsequently degrades (channel `degBetaP`). The complete model also includes processes for Wnt, Fz, LRP, β-catenin, and TCF (available in the Appendix C). The model is simulated using parameter values given in Table 4.1. We analyse stochastic simulation results to determine such statistics as the average amplitude of β-catenin and, in case of oscillations in β-catenin levels, the average amplitude

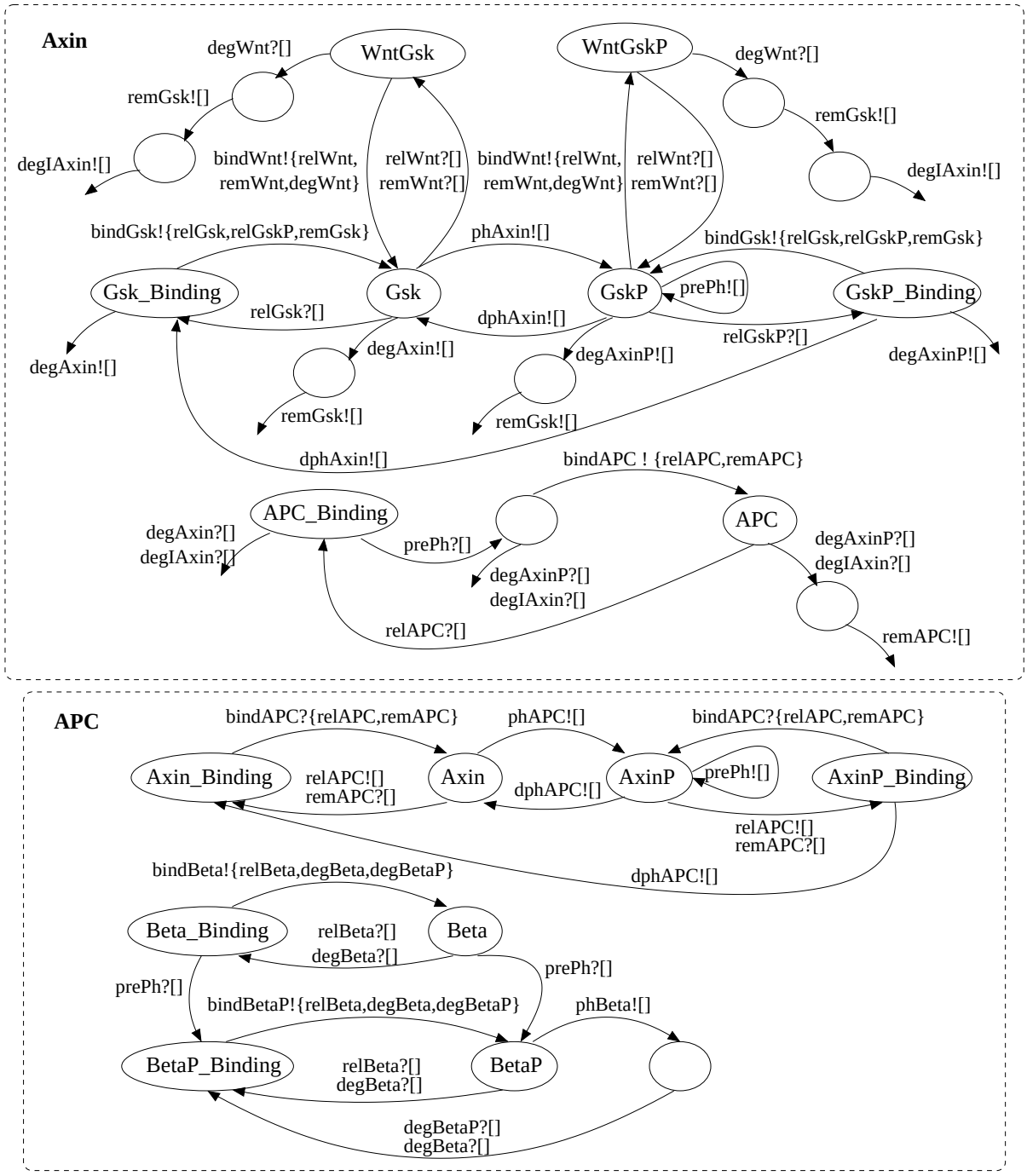


Figure 4.2: Representation of Axin and APC molecules in π-calculus.

and period of oscillations. The period of oscillations is determined as the time interval separating two successive upward crossings of the 100 nM level of β-catenin.

4.3.3 Wnt model in ODE

The deterministic model is constructed from the list of reactions in Table 4.1 using mass action kinetics. The components of the ODE model are the concentrations of β-catenin ([Beta]), elements of the destruction complex ([APC], [Axin], [GSK]), TCF transcription factor ([TCF]), and receptors ([Fz] and [LRP]). Additionally, [Axin], [APC] and [Beta] can change their state by becoming phosphorylated, which we denote as [AxinP], [APCP], and [BetaP], respectively. The concentration of the extracellular Wnt ([Wnt]) is an input parameter of the model.

Based on the law of mass action, we describe the temporal evolution of Axin using the following equation:

$$\begin{aligned}\frac{d[\text{Axin}]}{dt} &= k_{relAPC}([\text{APCAxin}] + [\text{APCPAxin}] + [\text{APCAxinBeta}] + [\text{APCPAxinBeta}] \\ &+ [\text{APCAxinBetaP}] + [\text{APCPAxinBetaP}]) - k_{bindGsk}[\text{Axin}][\text{Gsk}] + k_{relGsk}[\text{AxinGsk}] \\ &+ k_{dphAxin}[\text{AxinP}] + k_{synAxin}[\text{BetaTCF}^*] - k_{degAxin}[\text{Axin}]\end{aligned}$$

This equation specifies that the rate of change of Axin concentration consists of a gain from Axin:GSK3 complex disassociation (at the rate k_{relGsk}), Axin:APC complex disassociation (rate k_{relAPC}), Axin synthesis by activated β-catenin:TCF complex (rate $k_{synAxin}$); and a loss proportional to Axin binding to GSK3 (rate $k_{bindGsk}$), and Axin degradation (constant rate $k_{degAxin}$). Similarly, we define changes of the concentrations of Axin:Gsk,

phosphorylated AxinP:Gsk, and Axin recruited to the membrane in response to Wnt:

$$\frac{d[\text{AxinP}]}{dt} = k_{relAPC}([\text{APCAxinP}] + [\text{APCPAxinP}] + [\text{APCAxinPBeta}]$$
$$+ [\text{APCPAxinPBeta}] + [\text{APCAxinPBetaP}] + [\text{APCPAxinPBetaP}])$$
$$- k_{bindGsk}[\text{AxinP}][\text{Gsk}] + k_{relGsk}[\text{AxinPGsk}]$$
$$- k_{dphAxin}[\text{AxinP}] - k_{degAxinP}[\text{AxinP}]$$

$$\frac{d[\text{AxinGsk}]}{dt} = k_{relAPC}([\text{APCAxinGsk}] + [\text{APCPAxinGsk}] + [\text{APCAxinGskBeta}]$$
$$+ [\text{APCPAxinGskBeta}] + [\text{APCAxinGskBetaP}] + [\text{APCPAxinGskBetaP}])$$
$$k_{bindGsk}[\text{Axin}][\text{Gsk}] - k_{relGsk}[\text{AxinGsk}] - k_{phAxin}[\text{AxinGsk}]$$
$$+ k_{dphAxin}[\text{AxinPGsk}] - k_{degAxin}[\text{AxinGsk}]$$

$$\frac{d[\text{AxinPGsk}]}{dt} = -k_{bindAPC}([\text{APC}][\text{AxinPGsk}] + [\text{APCP}][\text{AxinPGsk}]$$
$$+ [\text{APCBeta}][\text{AxinPGsk}] + [\text{APCPBeta}][\text{AxinPGsk}] + [\text{APCBetaP}][\text{AxinPGsk}]$$
$$+ [\text{APCPBetaP}][\text{AxinPGsk}])$$
$$+ k_{relAPC}([\text{APCAxinPGsk}] + [\text{APCPAxinPGsk}] + [\text{APCAxinPGskBeta}]$$
$$+ [\text{APCPAxinPGskBeta}] + [\text{APCAxinPGskBetaP}] + [\text{APCPAxinPGskBetaP}])$$
$$+ k_{bindGsk}[\text{AxinP}][\text{Gsk}] - k_{relGskP}[\text{AxinPGsk}]$$
$$- k_{bindWnt}[\text{AxinPGsk}][\text{WntFzLRP}] + k_{relWnt}[\text{WntFzLRPAxinPGsk}])$$
$$+ k_{phAxin}[\text{AxinGsk}] - k_{dphAxin}[\text{AxinPGsk}] - k_{degAxinP}[\text{AxinPGsk}]$$

$$\frac{d[\text{WntFzLRPAxinPGsk}]}{dt} = k_{bindWnt}[\text{AxinPGsk}][\text{WntFzLRP}]$$
$$- k_{relWnt}[\text{WntFzLRPAxinPGsk}] - k_{degWnt}[\text{WntFzLRPAxinPGsk}]$$

Note that only Gsk-bound and phosphorylated form of Axin can bind APC.

We again run into a combinatorial explosion problem at the stage of model construction. The complete ODE model contains a separate equation for every molecular state of Axin. Axin interaction with APC then requires additional ten equations (for all possible combinations of Axin and APC):

$$\frac{d[\text{APC}]}{dt} = -k_{bindBeta}[\text{APC}][\text{Beta}] + k_{relBeta}[\text{APCBeta}] - k_{bindAPC}[\text{APC}][\text{AxinPGsk}])$$
$$+ k_{relAPC}([\text{APCAxin}] + [\text{APCAxinP}] + [\text{APCAxinPGsk}] + [\text{APCAxinGsk}])$$
$$+ k_{dphAPC}[\text{APCP}] + k_{degBeta}[\text{APCBeta}] + k_{degBetaP}[\text{APCBetaP}]$$
$$+ k_{degAxinP}([\text{APCAxinP}] + [\text{APCAxinPGsk}]) + k_{degAxin}([\text{APCAxin}] + [\text{APCAxinGsk}])$$

$$\frac{d[\text{APCAxin}]}{dt} = -k_{bindBeta}[\text{APCAxin}][\text{Beta}] + k_{relBeta}[\text{APCAxinBeta}]$$
$$- k_{relAPC}[\text{APCAxin}] - k_{bindGsk}[\text{APCAxin}][\text{Gsk}] + k_{relGsk}[\text{APCAxinGsk}]$$
$$+ k_{dphAxin}[\text{APCAxinP}] - k_{phAPC}[\text{APCAxin}] + k_{dphAPC}[\text{APCPAxin}]$$
$$+ k_{degBeta}[\text{APCAxinBeta}] + k_{degBetaP}[\text{APCAxinBetaP}] - k_{degAxin}[\text{APCAxin}]$$

$$\frac{d[\text{APCAxinP}]}{dt} = -k_{bindBeta}[\text{APCAxinP}][\text{Beta}] + k_{relBeta}[\text{APCAxinPBeta}]$$
$$- k_{relAPC}[\text{APCAxinP}] - k_{bindGsk}[\text{APCAxinP}][\text{Gsk}] + k_{relGskP}[\text{APCAxinPGsk}]$$
$$- k_{dphAxin}[\text{APCAxinP}] - k_{phAPC}[\text{APCAxinP}] + k_{dphAPC}[\text{APCPAxinP}]$$
$$+ k_{degBeta}[\text{APCAxinPBeta}] + k_{degBetaP}[\text{APCAxinPBetaP}] - k_{degAxinP}[\text{APCAxinP}]$$

$$\frac{d[\text{APCAxinGsk}]}{dt} = -k_{bindBeta}[\text{APCAxinGsk}][\text{Beta}] + k_{relBeta}[\text{APCAxinGskBeta}]$$
$$- k_{relAPC}[\text{APCAxinGsk}] + k_{bindGsk}[\text{APCAxin}][\text{Gsk}] - k_{relGsk}[\text{APCAxinGsk}]$$
$$- k_{phAxin}[\text{APCAxinGsk}] + k_{dphAxin}[\text{APCAxinPGsk}] - k_{phAPC}[\text{APCAxinGsk}]$$
$$+ k_{dphAPC}[\text{APCPAxinGsk}] + k_{degBeta}[\text{APCAxinGskBeta}]$$
$$+ k_{degBetaP}[\text{APCAxinGskBetaP}] - k_{degAxin}[\text{APCAxinGsk}]$$

$$\frac{d[\text{APCAxinPGsk}]}{dt} = -k_{bindBeta}[\text{APCAxinPGsk}][\text{Beta}] + k_{relBeta}[\text{APCAxinPGskBeta}]$$
$$+ k_{bindAPC}[\text{APC}][\text{AxinPGsk}] - k_{relAPC}[\text{APCAxinPGsk}] + k_{bindGsk}[\text{APCAxinP}][\text{Gsk}]$$
$$- k_{relGskP}[\text{APCAxinPGsk}] + k_{phAxin}[\text{APCAxinGsk}] - k_{dphAxin}[\text{APCAxinPGsk}]$$
$$- k_{phAPC}[\text{APCAxinPGsk}] + k_{dphAPC}[\text{APCPAxinPGsk}] + k_{degBeta}[\text{APCAxinPGskBeta}]$$
$$+ k_{degBetaP}[\text{APCAxinPGskBetaP}] - k_{degAxinP}[\text{APCAxinPGsk}]$$

$$\frac{d[\text{APCP}]}{dt} = -k_{bindBetaP}[\text{APCP}][\text{Beta}] + k_{relBeta}[\text{APCPBeta}]$$
$$+ k_{relAPC}([\text{APCPAxin}] + [\text{APCPAxinP}] + [\text{APCPAxinGsk}] + [\text{APCPAxinPGsk}])$$
$$- k_{dphAPC}[\text{APCP}] + k_{degBeta}[\text{APCPBeta}] + k_{degBetaP}[\text{APCPBetaP}]$$
$$+ k_{degAxin}([\text{APCPAxin}] + [\text{APCPAxinGsk}]) + k_{degAxinP}([\text{APCPAxinP}] + [\text{APCPAxinPGsk}])$$

$$\frac{d[\text{APCPAxin}]}{dt} = -k_{bindBetaP}[\text{APCPAxin}][\text{Beta}] + k_{relBeta}[\text{APCPAxinBeta}]$$
$$- k_{relAPC}[\text{APCPAxin}] - k_{bindGsk}[\text{APCPAxin}][\text{Gsk}] - k_{relGsk}[\text{APCPAxinGsk}]$$
$$+ k_{phAPC}[\text{APCAxin}] - k_{dphAPC}[\text{APCPAxin}] + k_{degBeta}[\text{APCPAxinBeta}]$$
$$+ k_{degBetaP}[\text{APCPAxinBetaP}] - k_{degAxin}[\text{APCPAxin}]$$

$$\frac{d[\text{APCPAxinP}]}{dt} = -k_{bindBetaP}[\text{APCPAxinP}][\text{Beta}] + k_{relBetaP}[\text{APCPAxinPBeta}]$$
$$- k_{relAPC}[\text{APCPAxinP}] - k_{bindGsk}[\text{APCPAxinP}][\text{Gsk}] + k_{relGsk}[\text{APCPAxinPGsk}]$$
$$- k_{dphAxin}[\text{APCPAxinP}] + k_{phAPC}[\text{APCAxinP}] - k_{dphAPC}[\text{APCPAxinP}]$$
$$+ k_{degBeta}[\text{APCPAxinPBeta}] + k_{degBetaP}[\text{APCPAxinPBetaP}] - k_{degAxinP}[\text{APCAxinP}]$$

$$\frac{d[\text{APCPAxinGsk}]}{dt} = -k_{bindBetaP}[\text{APCPAxinGsk}][\text{Beta}] + k_{relBeta}[\text{APCPAxinGskBeta}]$$
$$- k_{relAPC}[\text{APCPAxinGsk}] + k_{bindGsk}[\text{APCPAxin}] - k_{relGsk}[\text{APCPAxinGsk}]$$
$$- k_{phAxin}[\text{APCPAxinGsk}] + k_{dphAxin}[\text{APCPAxinPGsk}] + k_{phAPC}[\text{APCAxinGsk}]$$
$$- k_{dphAPC}[\text{APCPAxinGsk}] + k_{degBeta}[\text{APCPAxinGskBeta}]$$
$$+ k_{degBetaP}[\text{APCPAxinGskBetaP}] - k_{degAxin}[\text{APCPAxinGsk}]$$

$$\frac{d[\text{APCPAxinPGsk}]}{dt} = -k_{bindBetaP}[\text{APCPAxinPGsk}][\text{Beta}] + k_{relBeta}[\text{APCPAxinPGskBeta}]$$
$$+ k_{binAPC}[\text{APCP}][\text{AxinPGsk}] - k_{relAPC}[\text{APCPAxinPGsk}] + k_{bindGsk}[\text{APCPAxinP}][\text{Gsk}]$$
$$- k_{relGskP}[\text{APCPAxinPGsk}] + k_{phAxin}[\text{APCAxinGsk}] - k_{dphAxin}[\text{APCPAxinPGsk}]$$
$$+ k_{phAPC}[\text{APCAxinPGsk}] - k_{dphAPC}[\text{APCPAxinPGsk}] + k_{degBeta}[\text{APCPAxinPGskBeta}]$$
$$+ k_{degBetaP}[\text{APCPAxinPGskBetaP}] - k_{degAxinP}[\text{APCPAxinPGsk}]$$

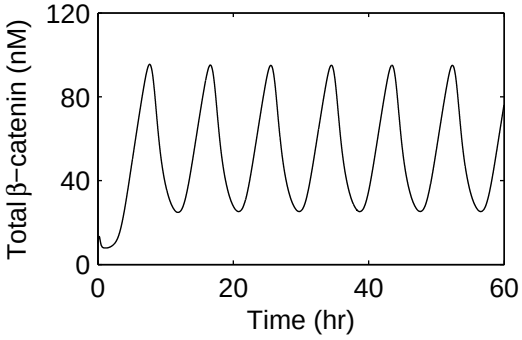
Additionally, APC binding to β-catenin on the second binding site leads to further model growth which now must include equations for all combinations between β-catenin and APC:Axin complex (adding 20 more equations to the model). The complete model containing 46 rate equations is available on the supporting website [Wnt, 2008]. Model simulations were performed with the continuous presence of Wnt stimulus in the system.

4.4 Model predictions

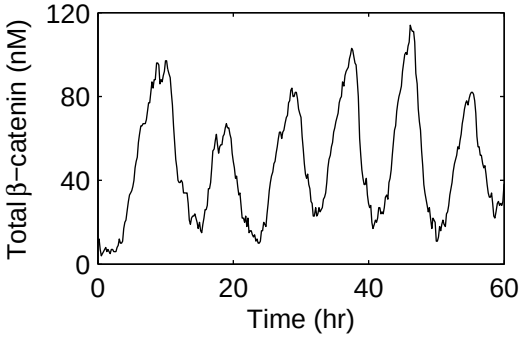
In the absence of Wnt signal, model results are in agreement with previously published experimental data on the rate of β-catenin degradation (half-life 60 min) [Aberle et al., 1997, Salic et al., 2000], and Axin degradation (half-life 4-8 hr) [Yamamoto et al., 1999]. The temporal profile of the response to Wnt stimulation by β-catenin accumulation and Axin transcription which suppresses β-catenin matches closely that reported in [Lustig et al., 2002]. The effect of changing the values of the parameters on the output of the system is analysed in subsequent sections.

4.4.1 Two dynamic oscillatory regimes

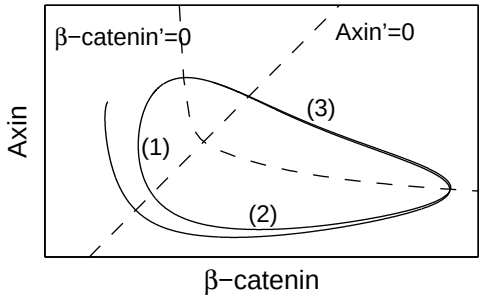
We observe that inclusion of the Axin feedback loop in the deterministic model leads to stable oscillations with a period of $\sim$10 hr at persistent Wnt stimulation (Fig. 4.3(a)). The stochastic model gives a similar result (Fig. 4.3(b)). The mechanism responsible for oscillations is illustrated in Fig. 4.3(c) by a schematic representation of the phase portrait of the system. In Fig. 4.3(c), dashed lines represent β-catenin and Axin nullclines $\beta\text{-catenin}' = 0$ and $\text{Axin}' = 0$. The solid line corresponds to the trajectory of the system (a limit cycle). Starting with low numbers of initial β-catenin molecules and high numbers of Axin, the trajectory (1) quickly moves downwards, corresponding to inhibition of Axin until it hits the nullcline where Axin production and degradation rates are equal. Here, low levels of Axin allow β-catenin to accumulate, corresponding to a slow move of the trajectory to the right (2). The trajectory overshoots the stable point due to a delay in the transcription of Axin, and β-catenin increases further, which gradually turns on Axin production. When sufficient levels of Axin remove β-catenin from the system, the trajectory moves left (3), reaching nullcline, where it turns downward to initiate a new cycle.



(a)



(b)



(c)

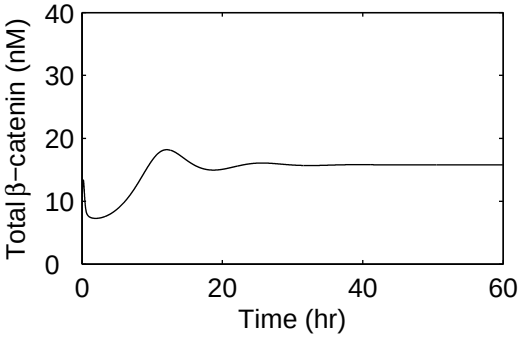
Figure 4.3: Oscillations in β-catenin concentration predicted by deterministic (a) and stochastic (b) versions of the model. (c) shows a schematic phase portrait of the deterministic oscillator.

The deterministic model predicts that, if the Wnt level exceeds a certain threshold, the signalling network undergoes sustained oscillations of the limit cycle type, whereas, at low Wnt stimulus, the network operates in a stable steady state regime with low β-catenin activity (Fig. 4.4(a)). The stochastic system predicts a different time course if Wnt stimulus is kept low. The presence of random fluctuations causes stochastic oscillations of β-catenin activity (Fig. 4.4(b)). The system always evolves toward a stable fixed point, as sketched in Fig. 4.4(c). In Fig. 4.4(c), the grey line illustrates how fluctuations initiate a single excursion in the phase space. However, due to low numbers of molecules[1], the intrinsic fluctuations near the fixed point are significant enough to continually send the system in the fast phase of Axin degradation which is necessarily followed by a transient activation of β-catenin. This behaviour results in sustained noise-induced oscillations with a significantly longer period of $\sim$26 hr. This is contrary to the deterministic system, where there are no perturbations to initiate a new cycle; therefore the trajectory stays at a fixed point.

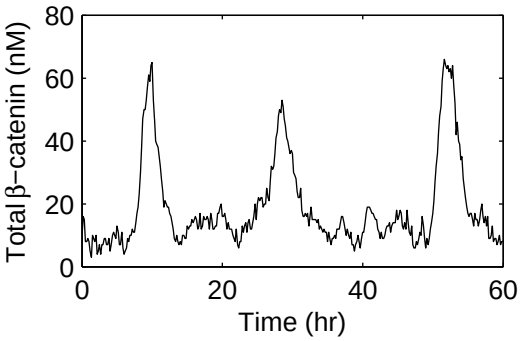
Cell proliferation is regulated by Wnt signals in various tissues. In the intestine, Wnt is required for proliferation of progenitor cells [Korinek et al., 1998, Pinto et al., 2003]. A strong expression of β-catenin coincides with a rapid proliferation of transit partially-differentiated cells. This is consistent with the model prediction of limit cycle regime of β-catenin at high Wnt concentrations (Fig. 4.3).

The model also implies that, in the low Wnt region, the system switches from the deterministic limit cycle solution to autonomous noise-induced oscillations of β-catenin with a markedly longer period (26 hr versus 10 hr). This would provide an explanation as to how the same molecular mechanism can drive both rapid proliferation of cells exposed to the maximal Wnt levels and continually push quiescent cells into the new cycle at low

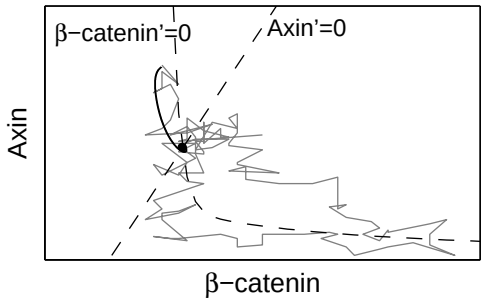
[1]Indeed, in at least one experimental system the number of Axin molecules per cell was found to be as low as 100 [Lee et al., 2003].



(a)



(b)



(c)

Figure 4.4: Time evolution of β-catenin concentration in deterministic (a) and stochastic (b) systems. (c) shows a phase portrait of the deterministic systems that falls into a stable steady state ($\bullet$) and a noise-induced excursion through a phase space of the stochastic system.

Wnt.

4.4.2 Experimental validation

Theoretically, it has been proposed that Wnt oscillations form the basis of the boundary formation in developing tissues, in particular, control segmentation process in somitogenesis (see experimental results in [Aulehla et al., 2003] and computational analysis in [Wawra et al., 2007]). However, the question whether oscillatory response appears in cell cultures responsive to Wnt has not yet been addressed.

Thus, we addressed the question whether Wnt stimulation triggers oscillatory expression of known β-catenin targets *in vitro*. The cyclical nature of the response to Wnt would not be apparent in many experiments, since Wnt signalling is followed routinely using a TOPFLASH reporter [Korinek et al., 1997], where the stable nature of the luciferase protein produced and measured integrates the transcription over time.

HeLa cells were treated with a Wnt3A-conditioned medium. Cells were harvested every two hours post treatment for 24 hours. Real-time PCR quantitation of relative Axin2 and NKD1 mRNA expression was calculated for each time point against time zero (no treatment) according to the comparative CT method. The first step in the calculations is the normalisation of the Axin2 and NKD1 gene to the epithelial cell specific KRT8 gene in order to control for quantity and quality of the cDNA samples.

Figures 4.5(a) and 4.5(c) demonstrate a synchronised cyclic expression of Axin2 and NKD1 determined by *in vitro* experiments. The first cycle persists for about 12 hours post treatment and peaks at 6 hours in both genes. After 12 hours of treatment, both genes show a further cycle of induction, with NKD1 peaking at 16 hours and Axin2 peaking at 20 hours, and finally dropping off after 24 hours of treatment.

We analyse the results to detect whether there is a periodic component in observed

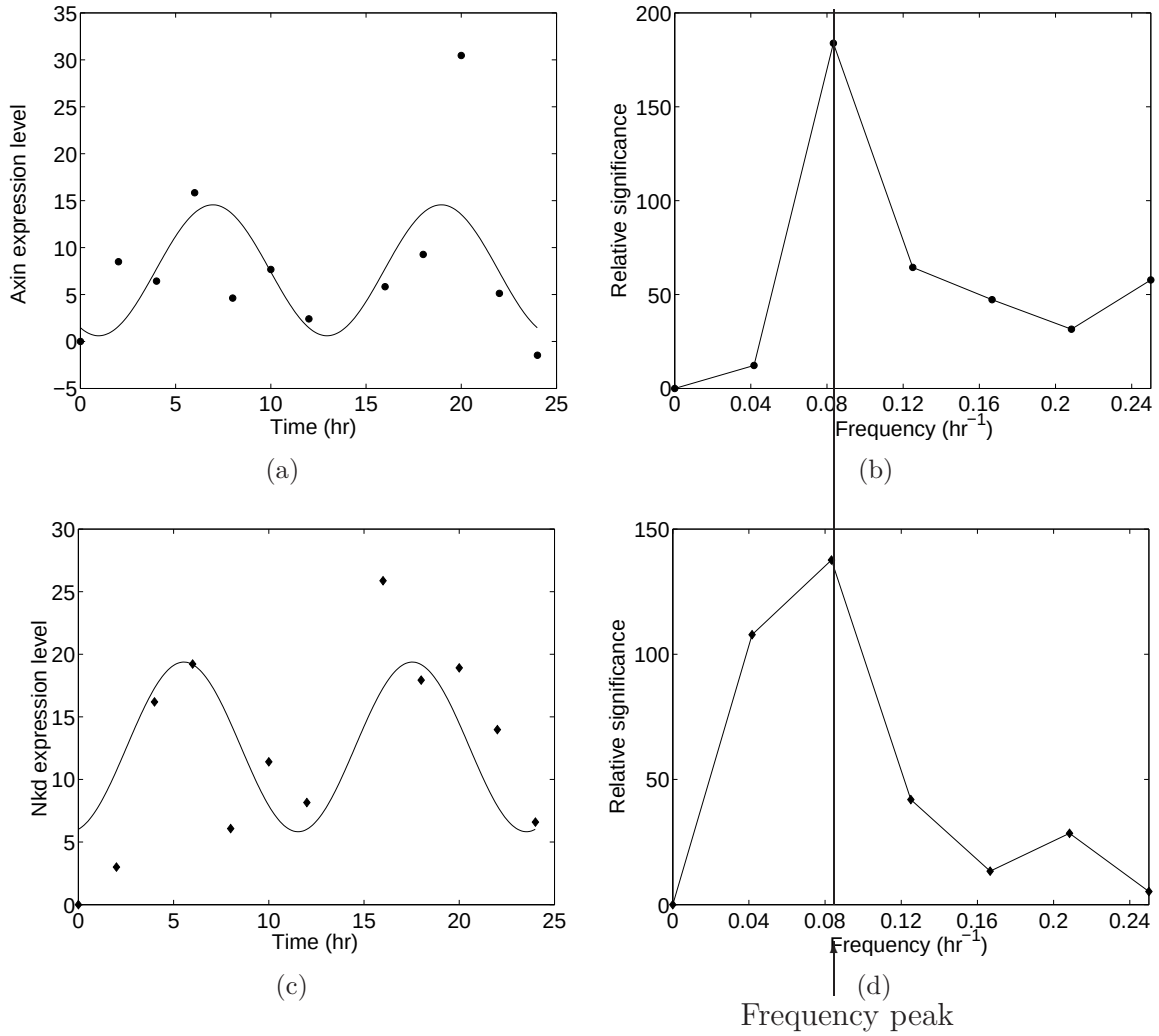


Figure 4.5: Experimental validation of the model predictions: real-time analysis of Axin2 (a) and NKD1 (c) expression in HeLa cells following treatment with Wnt3A conditioned medium; periodogram computed for Axin2 (b) and NKD1 (d) time series.

time series data $\{y_1, ..., y_N\}$, where N is the sample size. We build a frequency domain representation of the time series, known as periodogram and defined as follows:

$$I(\omega) = \frac{1}{N}|\sum_{n=1}^{N} y_n e^{-i\omega n}|^2, \omega \in [0, \pi].$$

The periodogram is further analysed at Fourier frequencies $\omega_k = 2\pi k/N, \quad k = 1, ..., [(N-1)/2]$. If a time series has a significant sinusoidal component with frequency ω_0, then the periodogram exhibits a peak at that frequency with high probability. Conversely, if the time series is a purely random process, then the plot of the periodogram $I(\omega)$ against the Fourier frequencies approaches a straight line [Wichert et al., 2004].

The periodograms computed for the Axin2 and NKD1 time series are shown in Fig. 4.5(b) and 4.5(d), respectively. These graphs exhibit a peak at frequency 0.083 hr^{-1} (corresponding to a period of $\sim$12 hr), indicating that both time series do have a cyclic component. Despite the fact that the length of the analysed time series is relatively short and there is a high level of noise inherent to such an experimental system, the periodogram can clearly detect the presence of a dominant frequency. Formal statistical significance tests that produce the p-value of 0.2 are not able to detect a dominant frequency due to the short sample size. The experimental frequency estimate is consistent with the model predictions of an oscillation period of $\sim$10 hr. It should be noted that the true frequency can slightly differ due to the sampling size used, and requires further experimentation. In Fig. 4.5(a) and 4.5(c) we used the estimated oscillatory frequency to fit amplitude and phase to data, using a least squares estimator.

4.4.3 Inherent robustness and adaptability

Negative feedback systems generally exhibit reduced sensitivity to variations of their components. We perform formal sensitivity analysis of the Wnt system to determine which rate constants in the model have the greatest influence on the overall system behaviour. The sensitivity coefficients, shown in Table 4.2, are defined as the relative change in the objective function Z upon a relative change in the parameter p, as defined by relations (2.5) and (2.6). In our case, the objective function Z relates to the level of β-catenin. When Wnt stimulus is low, we calculate the sensitivity of the β-catenin concentration at steady state (S_{AVG}).

To determine sensitivity of β-catenin levels in the oscillatory domain at higher Wnt, the periodic signal is expanded into a series of sine and cosine functions:

$$z(t) = A_0 + \sum_{n=1}^{\infty}(a_n \cos n\omega t + b_n \sin n\omega t),$$

where A_0 is the mean value of the signal, a_n and b_n, $n \geq 1$, are the amplitudes of the cosine and sine of the nth component, ω is the frequency of the signal. The amplitude can then be studied in terms of $A_n = \sqrt{a_n^2 + b_n^2}$. In the oscillatory domain at higher Wnt, sensitivity of frequency (S_{FREQ}) and amplitude of oscillations (mean amplitude S_{A_0}, absolute amplitude of the first and second components of oscillating trajectory, S_{A_1} and S_{A_2}) is calculated, using Fourier transform of Z.

Our sensitivity analysis predicts that β-catenin dynamics in the system is primarily controlled by the strength of β-catenin binding to the nuclear transcription factor TCF, by the synthesis rate of β-catenin and its inhibitor Axin, and the rate of Axin inhibition in response to Wnt stimulation.

Interestingly, the model predicts that APC mutations affect dynamic oscillatory be-

Parameter	S_{AVG}	S_{FREQ}	S_{A_0}	S_{A_1}	S_{A_2}
Wnt binds the receptor	0.11	-0.01	**0.2**	0.12	-0.005
Wnt release	-0.08	-0.004	**-0.18**	**-0.18**	-0.15
Axin:Gsk binding	-0.09	-0.05	-0.1	-0.06	-0.014
Axin:Gsk release	0.03	0.02	0.03	0.006	-0.02
Axin*:Gsk release	0.03	-0.005	0.03	-0.005	-0.05
Axin*:APC binding	-0.04	-0.07	-0.06	-0.08	-0.1
Axin*:APC release	0.03	0.005	0.06	0.06	-0.009
APC:β-catenin binding	-0.0003	0.003	-0.002	-0.009	-0.02
APC*:β-catenin binding	**-0.17**	-0.09	**-0.13**	0.03	0.1
APC:β-catenin release	0.004	0	0.005	0.009	0.02
TCF binding	**-0.82**	-0.02	**-0.8**	**-0.72**	**-0.4**
TCF release	**1.16**	**-0.26**	**1.27**	**0.68**	**-0.19**
Axin phosphorylation	-0.03	-0.06	-0.05	-0.08	-0.12
Axin dephosphorylation	0.01	-0.004	0.008	0.001	-0.01
APC phosphorylation	-0.08	-0.05	-0.09	-0.05	0.02
APC dephosphorylation	0.09	0.01	0.1	0.02	-0.1
β-catenin phosphorylation	-0.007	-0.009	-0.008	-0.009	-0.008
Wnt degradation	**-0.79**	0.06	**-1.22**	**-0.79**	-0.12
β-catenin synthesis	**0.19**	-0.06	**0.28**	**0.73**	**1.44**
β-catenin degradation	0.11	-0.04	0.09	-0.08	**-0.35**
β-catenin* degradation	-0.001	-0.007	-0.004	-0.01	-0.01
Axin synthesis	**-0.92**	0.0003	**-1.29**	**-0.69**	**0.14**
Transcriptional delay	**-0.46**	-0.11	**-0.66**	**-0.45**	-0.12
Axin degradation	0.1	-0.03	0.08	0.005	-0.08
Axin* degradation	0.09	-0.02	0.02	-0.008	-0.04
Wnt:LRP-Axin binding	0.01	0.02	0.05	**0.14**	**0.25**
Wnt:LRP-dependent Axin degradation	**0.75**	**-0.14**	**1.15**	**0.66**	**-0.16**

Table 4.2: Sensitivity coefficients that characterise the average value, frequency, and amplitude of β-catenin. **Bold** indicates highly sensitive parameters (in this particular example, sensitivity threshold of 0.12 is chosen).

haviour in a non-intuitive fashion. Mutations that decrease the rate of APC*:β-catenin binding lead to upregulation of the mean amplitude of the signal (AVG and A_0). At the same time, these mutations decrease the amplitude of the oscillatory components (A_1 and A_2), thereby reducing the region of dynamic oscillations of β-catenin. The interval of stochastic oscillations also diminishes when such mutations occur.

4.5 Biological significance of Wnt oscillations

In recent years, broad effects of Wnt signalling have emerged in a variety of different systems. A large proportion of experimental systems for studying the effects of Wnt are based on the segmentation process in somitogenesis. These studies have suggested that Wnt/β-catenin signalling is involved in a segmentation clock [Aulehla et al., 2003].

Another context in which the role of Wnt is important is proliferation, differentiation, and self-renewal of stem cells and their progeny. In the intestine, Wnt inhibition by dominant negative mutation of Tcf4 or by Dkk1 overexpression results in a loss of tissue [Korinek et al., 1998]. Mutations that stabilise β-catenin lead to cell population expansion and intestinal tumours [Kinzler et al., 1991]. In the hematopoietic system, stem and progenitor cells show enhanced proliferation in the presence of purified Wnt [Reya et al., 2003, Willert et al., 2003]. Contrary to these studies, active Wnt signalling promotes cell differentiation rather than self-renewal in the neural crest and human embryonic cells [Lee et al., 2004, Dravid et al., 2005].

These disparate observations underscore the importance of preserving a balance between proliferation and differentiation cell fate, despite naturally occurring perturbations and mutations. It also underscores the complexity of the cell fate determination mechanism mediated by Wnt. However, the only model discussed by biologists in this context is an intuitive linear threshold model which associates high Wnt with the increased ampli-

tude of β-catenin followed by subsequent cell fate change. The cellular context in which Wnt signalling occurs, possibly through involvement of other pathways such as Notch and BMP, is then responsible for cell-specific response to Wnt.

Formal modelling allows us to address a question: whether or not these different responses to Wnt could be generated by a common fate-determination mechanism that maps cell-specific signalling activation events into cell-specific responses. We demonstrate that, depending on the stimulus strength, Wnt/β-catenin pathway induces two distinct cellular responses. Low-level Wnt stimulation is sufficient to upregulate pathway targets to a significant amplitude, but only for a limited duration. After such an outbreak of activity, the cell returns to a quiescent state that helps it to eliminate a secondary response and preserve its fate. At moderate to high Wnt levels, cells are exposed to continuous oscillations of β-catenin levels of high amplitude and duration, resulting in a faster induction of the early targets which are sufficient to trigger a secondary response. Such secondary targets may be responsible for the onset of cell differentiation, observed in rapidly proliferating transit cells that are direct descendants of stem cells.

Our model also predicts that the stochastically induced short activation cycles of β-catenin are limited to the low Wnt region. Mutations that incapacitate the Axin:APC destruction complex and increase the average concentration of β-catenin also gradually decrease the domain of stochastic oscillations. While the number of cells responding to Wnt increases in mutants, the fraction of cells triggering primary but not secondary response decreases. If stem cell properties are generated by the initial cell response phase only, such architecture would therefore protect the tissue from exponential growth when mutations occur.

4.6 Conclusions

In this chapter, we derived deterministic and stochastic models of the Wnt signalling pathway. Model analysis has been performed to infer the properties of the pathway dynamic behaviour. We concluded that, depending on stimulation levels, the proposed negative feedback might lead to complex oscillatory behaviour of the components of the Wnt pathway. We were able to confirm this behaviour in cell cultures. We later suggested how the oscillatory behaviour can have functional significance in Wnt-dependent cell decisions.

Chapter 5

Cellular decisions

5.1 Introduction

Recent advances in Wnt signalling suggest that this pathway may be involved in the control of tissue homeostasis [van de Wetering et al., 2002, Reya et al., 2003, Sato et al., 2004, Lowry et al., 2005, Dravid et al., 2005, He et al., 2007]. In particular, it has been proposed that Wnt signalling regulates cell proliferation and renewal in the large intestinal epithelium. In this chapter, we study the problem of how Wnt signalling ensures that a fine balance between cell proliferation and differentiation is maintained in the large intestinal epithelium. We formally test the feasibility of the mechanisms proposed by different groups of experimentalists. Our approach is based on designing a family of models that couple cellular decisions with the state of the intracellular Wnt cascade and cell microenvironment. We then analyse these models to reveal signalling characteristics that ensure accuracy and robustness of Wnt-mediated determination of proliferative cell fate and lead to tissue architecture which is resistant to mutations.

5.1.1 Crypt biology

The large intestinal tract is built of geometric tubular structures called *crypts*. Intestinal homeostasis involves cell generation by division at the crypt base, progressive cell differentiation while they migrate to the top of the crypt, and cell death followed by extrusion when they reach the top. *Stem* cells, believed to reside at the crypt bottom, have the unique ability to maintain the entire epithelium. As they divide and move up, stem cells must constantly adjust their behaviour by entering partially differentiated population (called *transit*) prior to terminally differentiating. Simultaneously, the proliferative capability of transit cells is the highest and decreases as cells move upward. Cancer is a genetic disorder that arises through the loss of strict control over cell proliferation and differentiation. The question that arises is which factors control the ability of intestinal cells to keep a fine-tuned balance between cell division and differentiation.

In this chapter, we test the feasibility of different biological hypotheses about the influence of Wnt signalling on cell fate and the emergence of the robust regulation of cell numbers in the tissue. The function of Wnt cannot be measured experimentally; rather, only the average behaviour of the collection of cells in response to Wnt factors can be observed. We therefore employ computational modelling to examine whether the specific properties of the Wnt pathway architecture can provide the conditions necessary for the emergence of the robust regulation mechanism that ensures homeostasis in the intestine.

We base our approach on a conceptual extension of the stochastic π-calculus for spanning multiple scales. We describe in detail how to build a multi-scale model that couples signal transduction network to cellular decisions to proliferate and differentiate. We then analyse the model to reveal how a population of cells interacts and develops into a tissue under the influence of the environment.

5.2 Related work

Several modelling approaches for studying the self-renewal process in the intestine exist [Johnston et al., 2007, Loeffler et al., 1997, Gerike et al., 1998]. A recent model by [Johnston et al., 2007] is representative of the class of deterministic spatially-uniform models. The authors describe the evolution of cell numbers in stem, transit, and differentiated compartments, assuming constant compartment-dependent rates of renewal, differentiation and death. The model is shown to be very sensitive to changes in these macroscopic rate constants. The authors subsequently investigate the impact of the hypothetical negative feedback mechanism that, based on regulation of the rate at which cells differentiate, allows the crypt to maintain an equilibrium in cell numbers.

In a similar compartment-based but stochastic approach [Loeffler et al., 1997], crypt growth is described by a Markov process that models a stem cell population in which each stem cell produces zero, one, or two stem cells, according to a fixed probability distribution that does not vary from individual to individual. In the same manner as [Johnston et al., 2007], the probability of self-renewal versus differentiation is assumed to be pre-programmed and independent of the conditions, except in the case of stem cells knowing their number. Both models, however, do not give an indication of how the knowledge of stem cell numbers can be acquired and propagated between physically separated cells. No experimental evidence exists that supports this assumption.

The incorporation of a spatial cell fate control mechanism is achieved in [Gerike et al., 1998], where a deterministic model for crypt proliferation regulated by diffusible growth factor is presented. The epithelium is modelled as a one-dimensional array of cells. Each cell enters a cell cycle only if the growth factor concentration in the respective cell exceeds a certain threshold. The growth factor is spread by diffusion starting from the bottom of the crypt, but the concentration of the growth factor in the tissue

is constant. The model mechanism ensures the dynamic regulation of cell proliferation without the need to impose a static type-dependent program executed by every cell. However, under more realistic conditions of a stochastic time-varying growth-factor field, the accuracy of this mechanism would collapse, resulting in a high variability in the numbers of proliferative cells and crypt size.

5.3 Extending π-calculus to model cells

We are interested in testing possible hypotheses about Wnt signalling-based control of cell proliferation and differentiation in the intestine. To test the feasibility of the mechanisms proposed by different research groups, we build a π-calculus model that couples cellular decisions with the state of the Wnt signalling network embedded in every cell. Next, we explain how an extension of the π-calculus can be used for spanning different scales of the biological system.

5.3.1 Cells as mobile ambients

In order to extend the model with cell-level dynamics, we first acquire a mechanism for embedding molecules into cells. We use ambients [Cardelli and Gordon, 2006] to define a bounded place where interactions between local agents happen. Enrichment of the stochastic π-calculus with ambients, introduced in [Regev et al., 2004] and described below, provides the ability to specify communication between π-calculus processes based on their location within a common boundary.

An ambient is a location where computation happens: `cell[X]` stands for the process `X` running at the location `cell` (i.e., in an ambient `cell`). Locations may reside within locations: in `cell[mol[A] | mol[B]]` two ambients `mol` are incorporated into

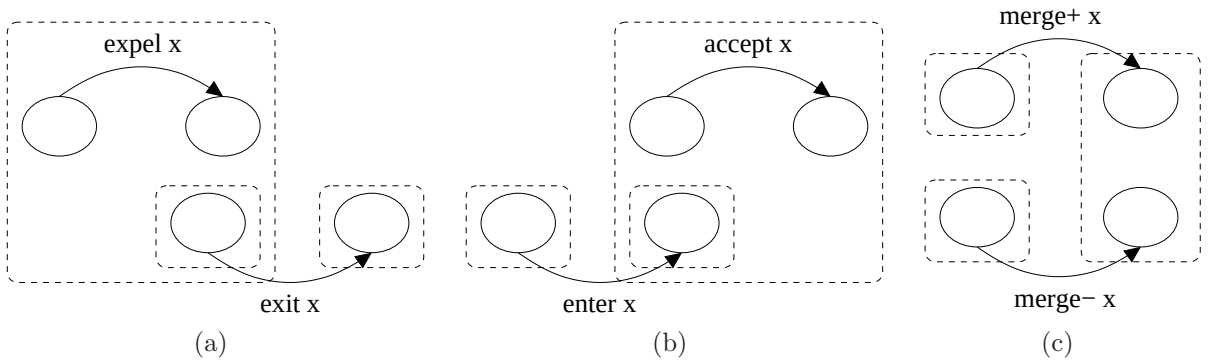


Figure 5.1: Ambient capabilities

the ambient `cell`. In ambient calculus, a computation may contain the reconfiguration of a hierarchy of locations. In the following, we graphically represent an ambient as a dashed rectangle around the processes and the sub-ambients it contains, possibly labelled with the ambient name. The derived models are simulated using the BioSPI platform [Regev and Shapiro, 2004] which supports ambients without modifying the semantics of the Stochastic Simulation Algorithm [Gillespie, 1976].

Spatial configuration of the ambient system can be changed using capabilities such as `exit/expel` from the ambient (graphically shown in Fig. 5.1(a)), `accept/enter` to the ambient (Fig. 5.1(b)), or `merge` with the ambient (Fig. 5.1(c)). Communication abstraction is extended to represent compartment restriction on interactions based on their locations. Three types of communication restrictions are: `local` (between processes in the same ambient, as graphically represented in Fig. 5.2(a)), `s2s` (between processes in sibling ambients in Fig. 5.2(b)), and `p2c/c2p` (between processes in parent and child ambients in Fig. 5.2(c)).

To represent cells, allowing molecules to be assigned and re-assigned to specific cells, we abstract cells as ambients. Consequently, molecular interaction within one cell is abstracted by the `s2s` communication direction. For communicating the state of the intracellular molecular network to the cell decision-making process, we use the `p2c/c2p`

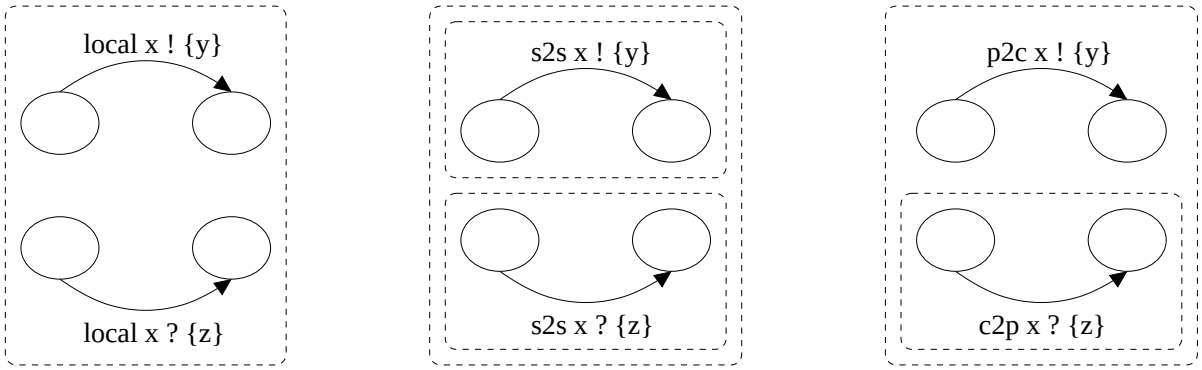


Figure 5.2: Communication directions between ambients

direction.

5.3.2 Cell division

Here, we consider the simplest possible model of cell division in which the time a cell spends to complete one division cycle follows an exponential distribution with the rate of the channel `cycle`. Upon completion of communication on the channel `cycle`, process `Cycle` creates another instance of a cell which is immediately extruded from the mother cell (Fig. 5.3). If we choose to simulate the re-distribution of molecules between mother and daughter cells, the process `Cycle` has to send a message (marked by channel `move` in Fig. 5.3) to every molecule instructing it to exit a mother cell and enter a daughter cell. The competition between channels `move` and `die` ensures that only a certain proportion of molecules will be re-assigned to the daughter cell.

5.3.3 Modelling space and cell movement

Most biochemical models do not incorporate explicit spatial information with the exception of the system volume Ω which is treated as a uniformly mixed solution. Although

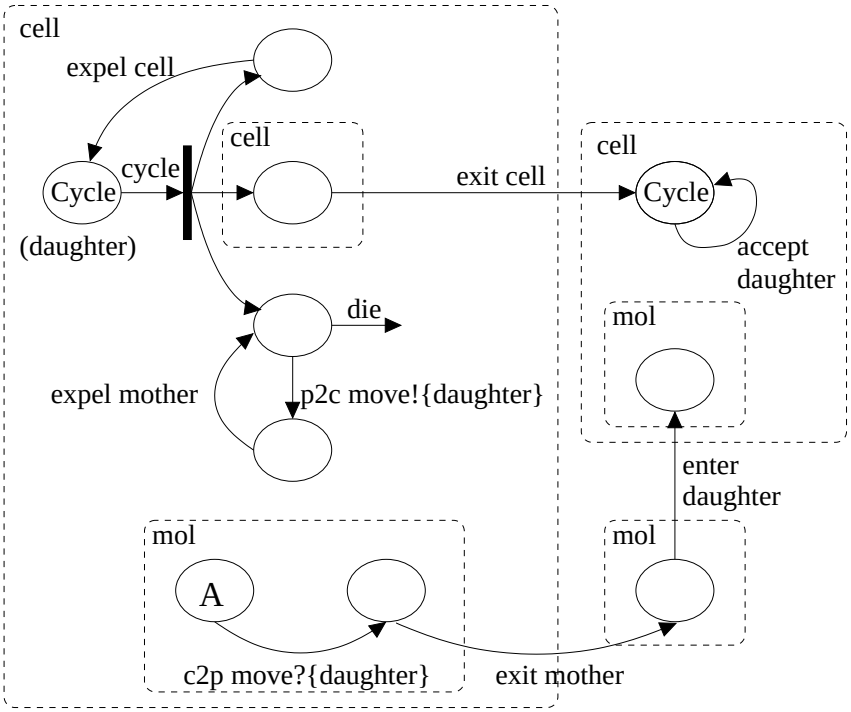


Figure 5.3: Cell division in π-calculus. The probability of reassigning a molecule to a daughter cell converges to 1/2, if the rates of channels `move` and `die` are equal.

this is clearly not how molecules are arranged within living cells, the omission of spatial heterogeneity has been the norm in biochemical simulations because it greatly facilitates modelling and reduces the computational complexity of simulations.

Here we present the spatial abstraction that describes the diffusion of the extracellular morphogene in one direction in the tissue. The pressure exerted by cell division due to higher amounts of the morphogene directs the cells to move away from the morphogene source. This would accommodate the scenario of Wnt morphogene distribution along the crypt length and cell movement to the top of the crypt. An analogous extension of the π-calculus framework with spatial information is necessary when the desired objective is to simulate diffusion of extracellular growth or inhibitory factors, competition for space between different cells, or cell adhesion.

To model spatial abstraction, we define a neighbourhood relationship between cells. Two cells are neighbours if they share a private channel which is used to send instructions from one neighbour to another. In one-dimensional space, it is sufficient for each cell to keep the reference to its upper neighbour (channel `next` in Fig. 5.4(a)). An extracellular signal (as modelled by interactions on channel `pos`) and cell movement (interactions on channel `move`) are functions of the neighbourhood. Following cell division, the upper neighbour is requested to free its position by moving upwards (input on channel `next`). A daughter cell is then inserted in the neighbouring position, as illustrated in Fig. 5.4(b). The extension of the lattice abstraction in two-dimensional space is straightforward; each cell contains references to its upper, left, and right neighbours (refer to the Attachment for details).

Diffusion of the external morphogene is simulated by calculating the concentration of the external factor field at each cell position rather than simulating the movement of factor molecules within the spatial lattice. The channel `pos` with an appropriate rate is carried

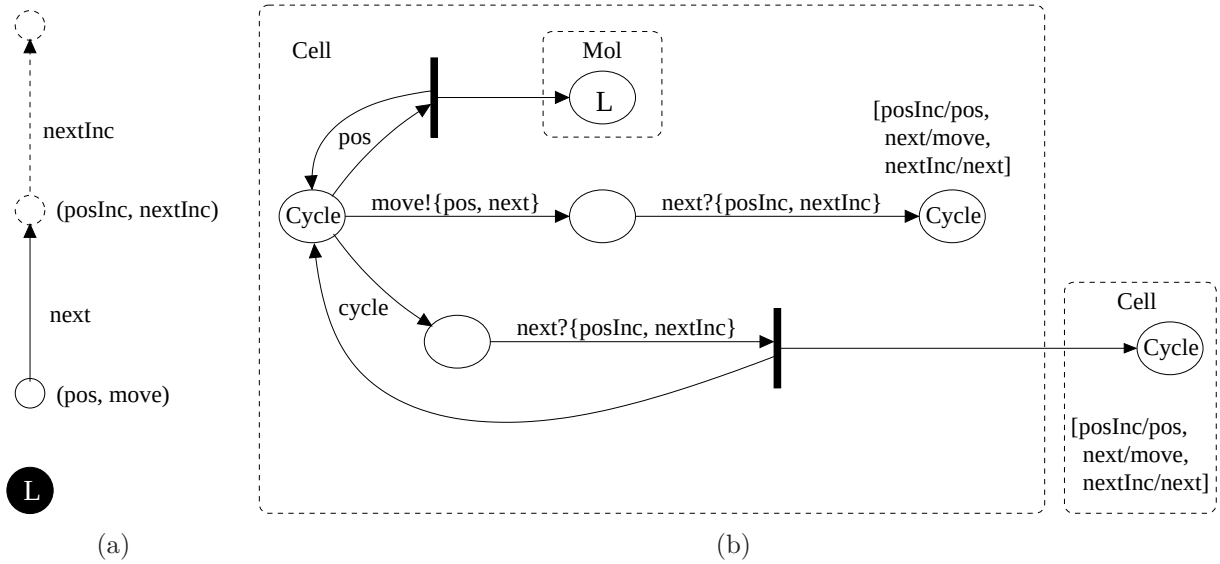


Figure 5.4: Cell organisation: (a) linear array of cells referencing upper neighbour; (b) lattice representation in π-calculus.

by each cell to indicate its distance to the morphogene source. The rate of the channel `pos` decays exponentially with distance from the source.

5.4 A model of intra- and inter-cellular dynamics of the crypt

Numerous and often inconsistent evidence exists suggesting that Wnt signalling controls the balance between cell proliferation and differentiation in the intestinal crypts and other tissues. Wnt is suggested to influence cell advance or withdrawal from the cell cycle [van de Wetering et al., 2002, Reya et al., 2003, Sato et al., 2004], and cell ability to maintain its stem-cell phenotype or differentiate [Dravid et al., 2005, Lowry et al., 2005, He et al., 2007]. To test different hypotheses about the Wnt-based regulatory mechanisms involved in the intestinal homeostasis, we build a multi-scale model that couples the state of the intracellular network to different decisions that the cell might make. The extracel-

lular diffusible Wnt triggers changes in the intracellular state and thus influences cellular behaviour. We examine how these mechanisms influence the robust turnover of cells in the intestinal crypt and its disregulation in cancer.

5.4.1 Intracellular signal transduction network

We adapt a model of the Wnt signalling pathway from Chapter 4 (Fig. 5.5). Pathway molecules are represented as processes incorporated into ambients. Molecular complex formation is modelled as ambient merge (e.g., transitions `merge- bindBeta` and `merge+ bindBeta` of processes `AxinP` and `Beta`). These can later disassociate by exiting the complex ambient (transitions `exit breakAxin` and `expel breakAxin`). Communications of molecules located in sibling ambients are marked by a direction `s2s`, while interactions within one ambient (e.g., molecular complex) are assigned a `local` direction. The coupling between subcellular and cellular scales is modelled by dependency of cellular decisions on the state of the intracellular network embedded in every cell ambient. `c2p/p2c` directions are used to communicate the state of the molecular network to the cell ambient.

5.4.2 Proliferative and differentiated cell fate

In our model, we adopt two threshold mechanisms in order to decide whether the cell undergoes proliferation, differentiation, or stays quiescent. Increased β-catenin activity influences the initiation of a new cell cycle. The time to complete the cycle is assumed to follow an exponential distribution. Variability of the cycle length is thus incorporated into the delay needed for the cell to make a decision to proliferate.

In addition, the β-catenin expression is linked to the ability of a stem cell to preserve its phenotype. We assume that once the cell starts expressing differentiation markers, differentiation is irreversible. While a stem cell divides to produce two cells with an equal

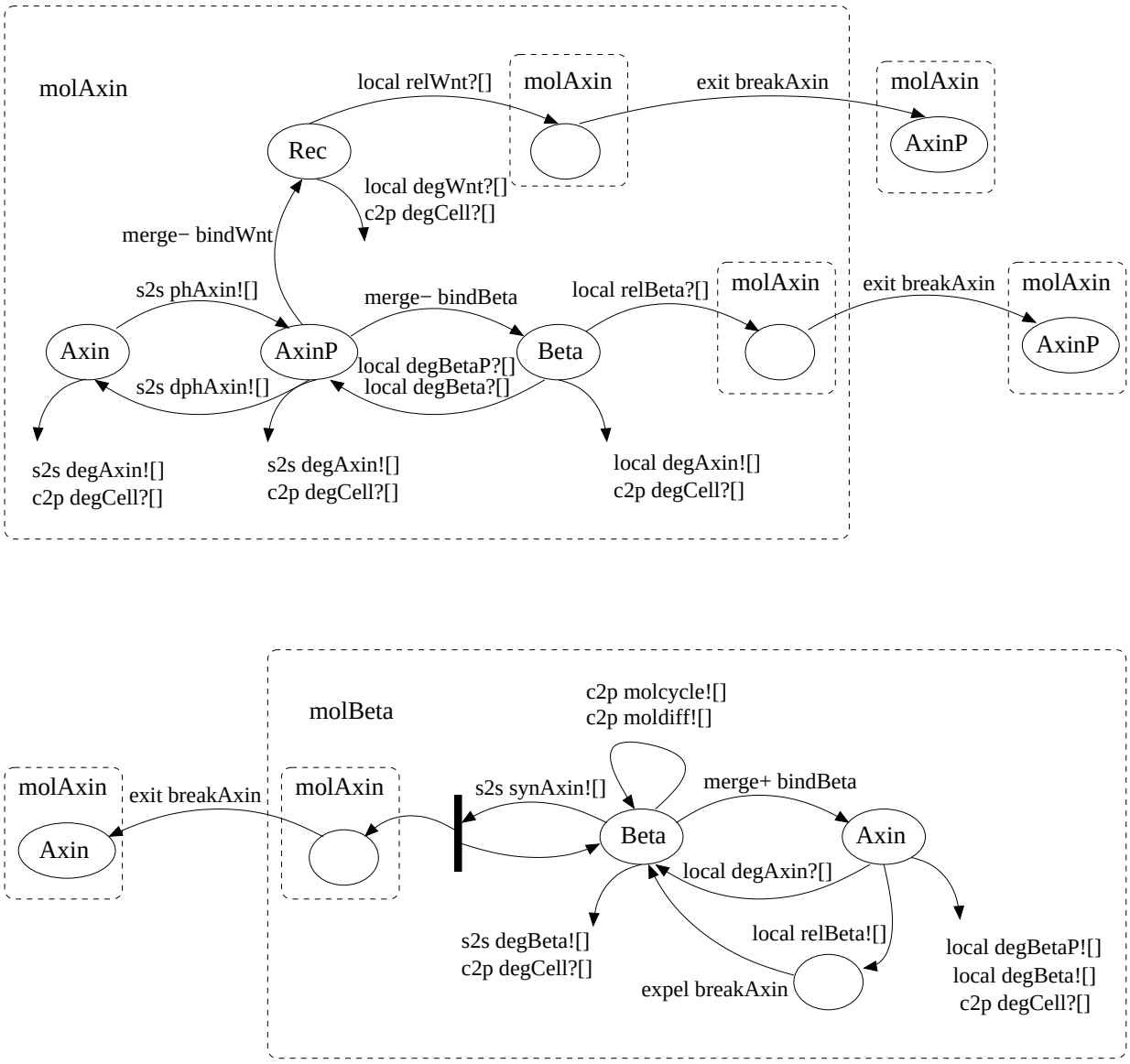


Figure 5.5: Adaptation of the Wnt subsystem in a multi-scale model.

stem-cell capability, a differentiated cell divides to produce two identical differentiated cells. Differentiated cells are also assumed to have a limited life span, as opposed to stem cells which are subjected to only a low-level apoptosis.

The alternative hypothesised scenarios of cell-fate decisions which we compare are:

Hypothesis 1. Transient activation of β-catenin in the cell triggers initiation of a new cell cycle. High levels of β-catenin are required to preserve stem cell properties.

Hypothesis 2. Transient activation of β-catenin is sufficient to push the cell into a new cycle while prolonged β-catenin signalling causes the stem cell to start expressing differentiation markers.

Each threshold mechanism is associated with a channel which transmits a signal to the cell once the level of the intracellular β-catenin exceeds a threshold. In Fig. 5.6, channel `molcycle` is used to instruct the cell to enter a new cycle. Another threshold `moldiff` blocks (Hypothesis 1) or triggers (Hypothesis 2) cell differentiation (Fig. 5.6). Complete model of the multi-scale cellular system is available in the Appendix D.

5.4.3 Wnt gradient in the tissue

Because Wnt targets are widely expressed in the stem and proliferative cell compartment, it is widely accepted that Wnt factors are produced at the bottom of the crypt and are then transported by diffusion [Brittan and Wright, 2004]. However, it has recently been suggested that the Wnt gradient follows a more complex pattern due to a surprisingly strong expression of the extracellular Wnt inhibitors at the bottom of the crypt [Gregorieff et al., 2005]. We approximate this by additionally decreasing the rate at which Wnt is received by cells located at the bottom of the crypt spatial lattice (channel `pos` in Fig. 5.6).

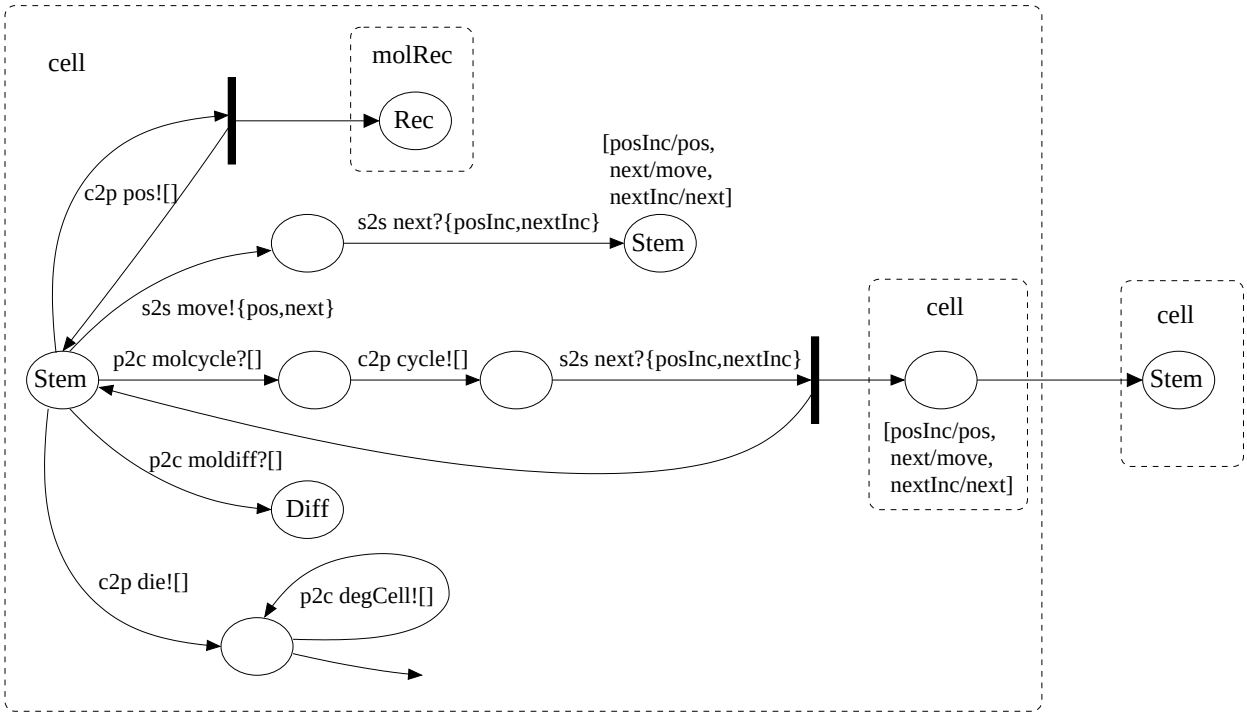


Figure 5.6: Stem cell evolution: a cell undergoes proliferation, differentiation, or death. Additionally, the cell is constantly receiving information about the environment, and adjusts its position within the spatial lattice to accommodate newly born cells.

5.5 Robust cell fate determination by Wnt signalling

When a cell senses Wnt in its neighbourhood, changes in its intracellular signalling network occur, which in turn determines cellular decision to proliferate or differentiate. Cellular behaviour is assumed to be stochastic and discrete: a cell enters cell cycle or differentiates with the probability proportional to the concentrations of the signalling response component β-catenin (as determined by the rates of channels `molcycle` and `moldiff`).

Using the BioSPI platform, we perform extensive simulations of the described scenarios in order to derive the properties of the multi-scale cellular system whose regulatory control is the extracellular diffusible factor Wnt. The derived models are subsequently analysed with respect to the number of cell divisions as a function of cell position along the crypt axis (i.e., distance to the Wnt source), the total number of cells in the crypt, and the influence of stochasticity and random parameter perturbations on the tissue response.

The first family of models implement the cellular decision mechanism described by Hypothesis 1. Our analysis shows that under these assumptions the fate that the cell assumes is very sensitive to the level of Wnt to which it is exposed. The distribution of proliferating cells mimics the distribution of the Wnt factors along the crypt axis. The result is a high variability in the size of the proliferative cell compartment and crypt size, which is inconsistent with the experimental observations. Moreover, activating mutations in the Wnt pathway, which increase the level of intracellular β-catenin, lead to significant expansion of the stem cell compartment. Consequently, the number of cells in the crypt becomes unstable and results in exponential growth. We conclude that Hypothesis 1 is unable to reproduce the tissue response observed experimentally.

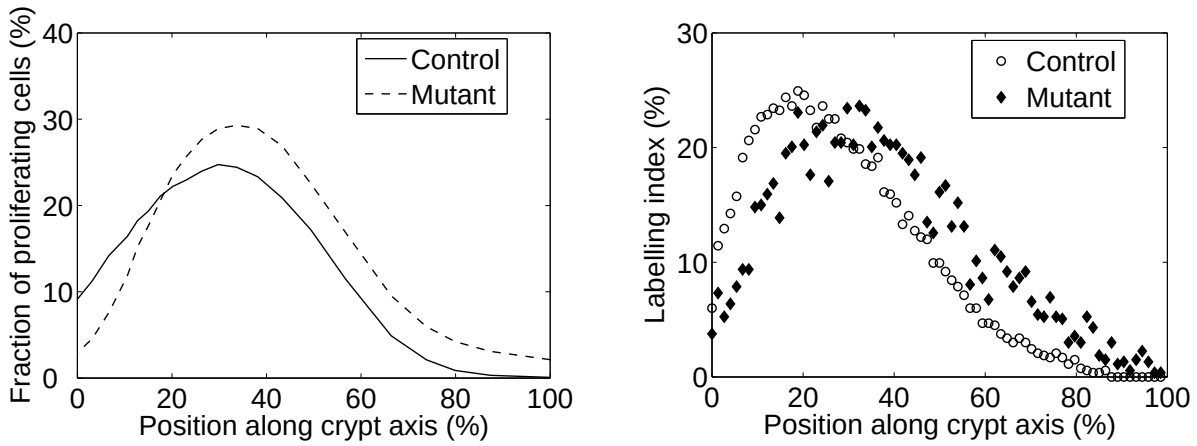
Simulations of the model based on Hypothesis 2 reveal that this combination of intracellular and cellular dynamics ensures robust tissue response mediated by Wnt (Fig. 5.7). Rather than being scattered throughout the crypt length, proliferative cells are confined

to the restricted compartment at the bottom of the crypt. This is consistent with the experimental data [Potten et al., 1992, Wong et al., 2002] (Fig. 5.7(b)). The number of proliferative cells as well as the total number of cells in the crypt shows little variability, despite random noise and stochastic perturbations present in the model. This is consistent with the reports of a surprisingly narrow distribution of crypt sizes, a fact that has not yet been reproduced in modelling studies.

We next investigate the effects of the mutations in the Wnt pathway which were identified in concrete cancer models: Familial Adenomatous Polyposis [Potten et al., 1992], and hyperplastic and adenomatous polyps [Wong et al., 2002, Sansom et al., 2004]. To simulate the effect of mutations, we decrease the rate of the β-catenin inhibition by the active APC/Axin destruction complex (channel `bindBetaP` in Fig. 5.5). An up to 5-fold decrease of the β-catenin inhibition rate results primarily in a shift of the proliferative cells toward the top of the crypt (Fig. 5.7(a)). The size of the crypt is increased only slightly. These predictions are in good agreement with the experimental evidence [Potten et al., 1992, Wong et al., 2002]. Figure 5.7(b), adapted from [Potten et al., 1992], shows experimental evidence of changes in the structure of the proliferative compartment resulting from mutations that decrease the activity of β-catenin inhibitor complex.

Further inhibition of the destruction complex leads to more advanced forms of intestinal cancer: colorectal adenomas [Wong et al., 2002, Sansom et al., 2004]. While cell proliferation shifts upwards at the initial stage, the model predicts the break up of the mechanism that confines proliferative cells to the bottom of the crypt. This is consistent with the experimental observations of proliferation in adenomas being almost evenly distributed throughout crypt length [Wong et al., 2002, Sansom et al., 2004].

Our model provides an explanation of the observed phenomena. Cell proliferation is triggered by even modest increase of the Wnt levels which is sufficient to upregulate



(a) (b)

Figure 5.7: Cell fate control by the Wnt pathway: (a) model predictions of the proliferative cell distribution in both healthy and mutant tissues agree well with (b) the experimental data.

β-catenin to high amplitude. As Wnt increases, stochastic oscillations in β-catenin expression become deterministic and their frequency increases along with the cell proliferation rate. Analysis of the model also shows that stem cell fate which is preserved only at the region of stochastic oscillations of β-catenin activity, is limited to low Wnt region and decreases under mutant conditions. Thus, the model is not only consistent with the reports of low β-catenin activity in stem cells [Lowry et al., 2005, Dravid et al., 2005] and the reduced proliferation rate of stem cells caused by rare outbreaks of β-catenin, but it also suggests the protection mechanism against the stem cell expansion that would immediately lead to the exponential growth of tumours [Johnston et al., 2007].

5.6 Conclusions

In this chapter, we employed formal modelling techniques based on the stochastic π-calculus to examine different hypotheses about the influence of the Wnt pathway on homeostasis of the intestinal epithelium, and its role in tumourigenesis. We proposed

that possible function of Wnt is to ensure robust cell fate determination. The model of Wnt signalling was subsequently coupled to the cellular behaviour and the environment to test its role in maintaining a fine-tuned balance between cell division and differentiation. The result of the model is consistent with different characteristics of the distribution of cells in the crypt. The model can explain both the stability of healthy regulation and the changes seen in mutant phenotypes. The model also suggests which characteristics of tissue architecture can protect it from unbounded growth. Increased predictive power of the multi-scale model demonstrates its usefulness to the understanding of the cellular machinery underlying robustness and adaptability.

Chapter 6

Conclusions and future work

In this thesis, we have applied modelling approaches based on process algebras, in particular stochastic π-calculus, to study problems relevant to systems biology. We have built models of various molecular signalling networks that offer a comprehensive and exact description of both static and dynamic information about pathway elements. We have analysed these models to derive properties of the collective action of molecular components, emphasising the importance of detecting feedback loops and non-linearities present in these systems. We have identified possible hypotheses about the structure and dynamics of systems under study, and have evaluated these hypotheses according to qualitative changes of possible system behaviour under both physiologically normal and perturbed conditions.

Two real-life signalling pathways which we analysed using this approach are FGF and Wnt. In both cases, by deriving exact descriptions of reaction networks, we were able to uncover previously unknown qualities of the overall behaviour of these pathways. We characterised pathway components according to their ability to amplify and downregulate signalling, and induce short- and long-term changes in intracellular network dynamics. We also demonstrated how malignant mutations in these pathways induce phenotypic changes

in their dynamic behaviour.

In the case of the FGF pathway, we identified the rate of receptor kinase activation and the rate of receptor complex internalisation as the key variables controlling signalling dynamics. The model predicted that inhibition of receptor internalisation would result in prolonged signalling. This hypothesis was later validated by *in vitro* experiments which also added further details to the model. The refined model showed the complex interplay between receptor activation and attenuation which gave a new insight into the nature of pathogenic mutations in FGF receptors.

The model of the Wnt signalling pathway also predicted previously unknown dynamic behaviour of the pathway. Depending on the level of extracellular stimulus, the cell exhibits slow noise-induced or fast regular oscillations in the β-catenin levels. Existence of β-catenin oscillations in a population of cells was later confirmed experimentally. An open question that remains is whether low levels of extracellular Wnt cause stochastic oscillations at the single cell level that become more regular as Wnt increases. In future work, we aim to validate this prediction experimentally.

Additionally, we have incorporated the intracellular dynamics into a higher-level model that includes cells and their interactions with each other and with the environment. The proposed scheme is based on the extension of stochastic π-calculus with ambients, which has allowed us to perform simulations without changing the underlying stochastic computational engine. Based on this extension, we performed a study of the cellular interactions and decision-making which are involved in intestinal homeostasis and cancer initiation. We hypothesised that cell decision to divide or differentiate depends on the level of intracellular β-catenin. We showed which threshold model of cellular decisions is consistent with the experimental evidence of proliferative cell distributions in the intestinal epithelial tissue. Moreover, the hypothesised model ensures robustness of cell-fate determination and

leads to tissue architecture robust to mutations. In the future it remains to be determined whether the proposed scenario is valid in cellular systems.

Our studies have proved the utility of the π-calculus modelling approach for deriving new insights into normal and pathological dynamics of processes occurring in multicellular systems at different scales. π-calculus offers a concise way to encode biological knowledge, formulate hypotheses, evaluate, and select those scenarios that are consistent with existing experimental evidence. Simulations of π-calculus models do not build the global state of the system and therefore avoid exponential explosion of the state space. Finally, the stochastic nature of process algebra makes it suitable for analysing the impact of random fluctuations on the behaviour of the system.

Among the limitations of the stochastic π-calculus modelling approach is the insufficient tool support for analysing models, including visualisation of the model results. This is particularly important in case of multi-level and compartmental models. Additionally, it would be greatly beneficial to extend the simulation algorithm to support more complex reaction rates governed by general probability distributions. Future work has to account for these extensions of the π-calculus simulation platform.

Bibliography

[Aberle et al., 1997] Aberle, H., Bauer, A., Stappert, J., Kispert, A., and Kemler, R. (1997). Beta-catenin is a target for the ubiquitin-proteasome pathway. *EMBO J*, 16(13):3797–3804.

[Albeck et al., 2006] Albeck, J. G., MacBeath, G., White, F. M., Sorger, P. K., Lauffenburger, D. A., and Gaudet, S. (2006). Collecting and organizing systematic sets of protein data. *Nat Rev Mol Cell Biol*, 7:803–812.

[Anderson et al., 1998] Anderson, J., Burns, H., Enriquez-Harris, P., Wilkie, A., and Heath, J. (1998). Apert syndrome mutations in fibroblast growth factor receptor 2 exhibit increased affinity for FGF ligand. *Hum Mol Genet*, 7(9):1475–1483.

[Aulehla et al., 2003] Aulehla, A., Wehrle, C., Brand-Saberi, B., Kemler, R., Gossler, A., Kanzler, B., and Herrmann, B. (2003). Wnt3a plays a major role in the segmentation clock controlling somitogenesis. *Dev Cell*, 4(3):395–406.

[Bower and Beeman, 1998] Bower, J. and Beeman, D. (1998). *The book of Genesis*. Springer.

[Brittan and Wright, 2004] Brittan, M. and Wright, N. (2004). The gastrointestinal stem cells. *Cell Prolif*, 37:35–53.

[Calder et al., 2006] Calder, M., Vyshemirsky, V., Gilbert, D., and Orton, R. (2006). Analysis of signalling pathways using Continuous Time Markov Chains. *T Comp Sys Biology*, 4220:44–67.

[Cardelli and Gordon, 2006] Cardelli, L. and Gordon, A. (2006). Mobile ambients. *Theor Comput Sci*, 240:177–213.

[Cho et al., 2006] Cho, K., Baek, S., and Sung, M. (2006). Wnt pathway mutations selected by optimal beta-catenin signaling for tumorigenesis. *FEBS Lett*, 580(15):3665–3670.

[Choi et al., 2006] Choi, H., Huber, A., and Weis, W. (2006). Thermodynamics of beta-catenin-ligand interactions: the roles of the N- and C-terminal tails in modulating binding affinity. *J Biol Chem*, 281(2):1027–1038.

[Clague and Urbe, 2001] Clague, M. and Urbe, S. (2001). The interface of receptor trafficking and signalling. *J Cell Sci*, 114(Pt 17):3075–3081.

[Clevers and van de Wetering, 1997] Clevers, H. and van de Wetering, M. (1997). TCF/LEF factor earn their wings. *Trends Genet*, 13(12):485–489.

[Cong et al., 2004] Cong, F., Schweizer, L., and Varmus, H. (2004). Wnt signals across the plasma membrane to activate the beta-catenin pathway by forming oligomers containing its receptors, Frizzled and LRP. *Development*, 131(20):5103–5115.

[Dikic and Giordano, 2003] Dikic, I. and Giordano, S. (2003). Negative receptor signalling. *Curr Opin Cell Biol*, 15(2):128–135.

[Dravid et al., 2005] Dravid, G., Ye, Z., Hammond, H., Chen, G., Pyle, A., Donovan, P., Yu, X., and Cheng, L. (2005). Defining the role of Wnt/beta-catenin signaling in the survival, proliferation, and self-renewal of human embryonic stem cells. *Stem Cells*, 23(10):1489–1501.

[Fages et al., 2004] Fages, F., Soliman, S., and Chabrier-Rivier, N. (2004). Modelling and querying interaction networks in the biochemical abstract machine BIOCHAM. *J Biol Phys Chem*, 4:64–73.

[Felder et al., 1993] Felder, S., Zhou, M., Hu, P., Urena, J., Ullrich, A., Chaudhuri, M., White, M., Shoelson, S., and Schlessinger, J. (1993). SH2 domains exhibit high-affinity binding to tyrosine-phosphorylated peptides yet also exhibit rapid dissociation and exchange. *Mol Cell Biol*, 13(3):1449–1455.

[FGF, 2008] FGF (2008). `http://www.cs.bham.ac.uk/~oxt/fgfmap.html`.

[Foehr et al., 2001] Foehr, E., Raffioni, S., Murray-Rust, J., and Bradshaw, R. (2001). The role of tyrosine residues in fibroblast growth factor receptor 1 signaling in PC12 cells. systematic site-directed mutagenesis in the endodomain. *J Biol Chem*, 276(40):37529–37536.

[Fong et al., 2003] Fong, C., Leong, H., Wong, E., Lim, J., Yusoff, P., and Guy, G. (2003). Tyrosine phosphorylation of Sprouty2 enhances its interaction with c-Cbl and is crucial for its function. *J Biol Chem*, 278(35):33456–33464.

[Furdui et al., 2006] Furdui, C., Lew, E., Schlessinger, J., and Anderson, K. (2006). Autophosphorylation of FGFR1 kinase is mediated by a sequential and precisely ordered reaction. *Mol Cell*, 21(5):711–717.

[Gaffney et al., 2007] Gaffney, E., Heath, J., and Kwiatkowska, M. (2007). A mass action model of a Fibroblast Growth Factor signalling pathway and its simplification. *Bull Math Biol*, page to appear.

[Gaspar and Fodde, 2004] Gaspar, C. and Fodde, R. (2004). APC dosage effects in tumorigenesis and stem cell differentiation. *Int J Dev Biol*, 48(5-6):377–386.

[Gerike et al., 1998] Gerike, T., Paulus, U., Potten, C., and Loeffler, M. (1998). A dynamic model of proliferation and differentiation in the intestinal crypt based on a hypothetical intraepithelial growth factor. *Cell Prolif*, 31:93–110.

[Gillespie, 1976] Gillespie, D. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J Comp Phys*, 22:403–434.

[Gonzalez-Sancho et al., 2005] Gonzalez-Sancho, J., Aguilera, O., Garcia, J., Pendas-Franco, N., Pena, C., Cal, S., Garcia de Herreros, A., Bonilla, F., and Munoz, A. (2005). The wnt antagonist DICKKOPF-1 gene is a downstream target of beta-catenin/TCF and is downregulated in human colon cancer. *Oncogene*, 24(6):1098–1103.

[Greenman et al., 2007] Greenman, C., Stephens, P., Smith, R., Dalgliesh, G., Hunter, C., Bignell, G., Davies, H., Teague, J., Butler, A., Stevens, C., Edkins, S., O'Meara, S., Vastrik, I., Schmidt, E., Avis, T., Barthorpe, S., Bhamra, G., Buck, G., Choudhury, B., and Clements, J, e. a. (2007). Patterns of somatic mutation in human cancer genomes. *Nature*, 446(7132):153–158.

[Gregorieff et al., 2005] Gregorieff, A., Pinto, D., Begthel, H., Destree, O., Kielman, M., and Clevers, H. (2005). Expression pattern of wnt signaling components in the adult intestine. *Gastroenterology*, 129(2):626–638.

[Ha et al., 2004] Ha, N., Tonozuka, T., Stamos, J., Choi, H., and Weis, W. (2004). Mechanism of phosphorylation-dependent binding of APC to beta-catenin and its role in beta-catenin degradation. *Mol Cell*, 15(4):511–521.

[Hadari et al., 1998] Hadari, Y., Kouhara, H., Lax, I., and Schlessinger, J. (1998). Binding of Shp2 tyrosine phosphatase to FRS2 is essential for fibroblast growth factor-induced PC12 cell differentiation. *Mol Cell Biol*, 18(7):3966–3973.

[Hajihosseini et al., 2004] Hajihosseini, M., Lalioti, M., Arthaud, S., Burgar, H., Brown, J., Twigg, S., Wilkie, A., and Heath, J. (2004). Skeletal development is regulated by fibroblast growth factor receptor 1 signalling dynamics. *Development*, 131(2):325–335.

[Hanafusa et al., 2002] Hanafusa, H., Torii, S., Yasunaga, T., and Nishida, E. (2002). Sprouty1 and Sprouty2 provide a control mechanism for the Ras/MAPK signalling pathway. *Nat Cell Biol*, 4(11):850–858.

[Hart et al., 1998] Hart, M., de los Santos, R., Albert, I., Rubinfeld, B., and Polakis, P. (1998). Downregulation of beta-catenin by human Axin and its association with the APC tumor suppressor, beta-catenin and GSK3 beta. *Curr Biol*, 8(10):573–581.

[He et al., 2007] He, X., Yin, T., Grindley, J., Tian, Q., Sato, T., Tao, W., Dirisina, R., Porter-Westpfahl, K., Hembree, M., Johnson, T., Wiedemann, L., Barrett, T., Hood, L., Wu, H., and Li, L. (2007). PTEN-deficient intestinal stem cells initiate intestinal polyposis. *Nat Genet*, 39(2):189–198.

[Heath et al., 2007] Heath, J., Kwiatkowska, M., Norman, G., Parker, D., and Tymchyshyn, O. (2007). Probabilistic model checking of complex biological pathways. *Theor Comput Sci (Special Issue on Converging Sciences: Informatics and Biology)*, 391:239–257.

[Hlavacek et al., 2006] Hlavacek, W., Faeder, J., Blinov, M., Posner, R., Hucka, M., and Fontana, W. (2006). Rules for modeling signal-transduction systems. *Sci STKE*, 2006(344):re6.

[Ikeda et al., 1998] Ikeda, S., Kishida, S., Yamamoto, H., Murai, H., Koyama, S., and Kikuchi, A. (1998). Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3beta and beta-catenin and promotes GSK-3beta-dependent phosphorylation of beta-catenin. *EMBO J*, 17(5):1371–1384.

[Johnston et al., 2007] Johnston, M., Edwards, C., Bodmer, W., Maini, P., and Chapman, S. (2007). Mathematical modeling of cell population dynamics in the colonic crypt and in colorectal cancer. *Proc Natl Acad Sci USA*, 104(10):4008–4013.

[Kan et al., 2002] Kan, S., Elanko, N., Johnson, D., Cornejo-Roldan, L., Cook, J., Reich, E., Tomkins, S., Verloes, A., Twigg, S., Rannan-Eliya, S., McDonald-McGinn, D., Zackai, E., Wall, S., Muenke, M., and Wilkie, A. (2002). Genomic screening of fibroblast growth-factor receptor 2 reveals a wide spectrum of mutations in patients with syndromic craniosynostosis. *Am J Hum Genet*, 70(2):472–486.

[KEGG, 1995] KEGG (1995). `http://www.kegg.com/`.

[Kikuchi, 1999] Kikuchi, A. (1999). Roles of Axin in the Wnt signalling pathway. *Cell Signal*, 11(11):777–788.

[Kingsmore, 2006] Kingsmore, S. (2006). Multiplexed protein measurement: technologies and applications of protein and antibody arrays. *Nat Rev Drug Discov*, 5:310–321.

[Kinzler et al., 1991] Kinzler, K., Nilbert, M., Vogelstein, B., Bryan, T., Levy, D., Smith, K., Preisinger, A., Hamilton, S., Hedge, P., Markham, A., and et al. (1991). Identification of a gene located at chromosome 5q21 that is mutated in colorectal cancers. *Science*, 251(4999):1366–1370.

[Kishida et al., 1998] Kishida, S., Yamamoto, H., Ikeda, S., Kishida, M., Sakamoto, I., Koyama, S., and Kikuchi, A. (1998). Axin, a negative regulator of the Wnt signaling pathway, directly interacts with adenomatous polyposis coli and regulates the stabilization of beta-catenin. *J Biol Chem*, 273(18):10823–10826.

[Kitano et al., 2005] Kitano, H., Funahashi, A., Matsuoka, Y., and Oda, K. (2005). Using process diagrams for the graphical representation of biological networks. *Nat Biotechnol*, 23:961–966.

[Knapp et al., 2001] Knapp, S., Zamai, M., Volpi, D., Nardese, V., Avanzi, N., Breton, J., Plyte, S., Flocco, M., Marconi, M., Isacchi, A., and Caiolfa, V. (2001). Thermodynamics of the high-affinity interaction of TCF4 with beta-catenin. *J Mol Biol*, 306(5):1179–1189.

[Kohn et al., 2006] Kohn, K., Aladjem, M., Kim, S., Weinstein, J., and Pommier, Y. (2006). Depicting combinatorial complexity with the molecular interaction map notation. *Mol Syst Biol*, 2:51.

[Korinek et al., 1998] Korinek, V., Barker, N., Moerer, P., van Donselaar, E., Huls, G., Peters, P., and Clevers, H. (1998). Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nat Genet*, 19:379–83.

[Korinek et al., 1997] Korinek, V., Barker, N., Morin, P., van Wichen, D., deWeger, R., Kinzler, K., Vogelstein, B., and Clevers, H. (1997). Constitutive transcriptional activation by a beta-cateninTcf complex in APC-/- colon carcinoma. *Science*, 275(5307):1784–1787.

[Kwiatkowska et al., 2007] Kwiatkowska, M., Norman, G., and Parker, D. (2007). *Lecture Notes in Computer Science*, pages 220–270. Stochastic Model Checking. Springer.

[Kwiatkowska et al., 2006] Kwiatkowska, M., Norman, G., Parker, D., Tymchyshyn, O., Heath, J., and Gaffney, E. (2006). Simulation and verification for computational modelling of signalling pathways. In *Proc of Winter Simul Conf*, pages 1666–1674.

[Le Novere and Shimizu, 2001] Le Novere, N. and Shimizu, T. (2001). StochSim: Modelling of stochastic biomolecular processes. *Bioinformatics*, 17:575–576.

[Lee et al., 2003] Lee, E., Salic, A., Kruger, R., Heinrich, R., and Kirschner, M. (2003). The roles of APC and Axin derived from experimental and theoretical analysis of the Wnt pathway. *PLoS Biol*, 1(1):E10.

[Lee et al., 2004] Lee, H., Kleber, M., Hari, L., Brault, V., Suter, U., Taketo, M., Kemler, R., and Sommer, L. (2004). Instructive role of Wnt/beta-catenin in sensory fate specification in neural crest stem cells. *Science*, 303(5660):1020–1023.

[Leung et al., 2002] Leung, J., Kolligs, F., Wu, R., Zhai, Y., Kuick, R., Hanash, S., Cho, K., and Fearon, E. (2002). Activation of AXIN2 expression by beta-catenin-T cell factor. A feedback repressor pathway regulating Wnt signaling. *J Biol Chem*, 277(24):21657–21665.

[Li et al., 2004] Li, X., Brunton, V., Burgar, H., Wheldon, L., and Heath, J. (2004). FRS2-dependent SRC activation is required for fibroblast growth factor receptor-induced phosphorylation of Sprouty and suppression of ERK activity. *J Cell Sci*, 117(Pt 25):6007–6017.

[Loeffler et al., 1997] Loeffler, M., Bratke, T., Paulus, U., Li, Y., and Potten, C. (1997). Clonality and life cycles of intestinal crypts explained by a state dependent stochastic model of epithelial stem cell organization. *J Theor Biol*, 186(1):41–54.

[Logan and Nusse, 2004] Logan, C. and Nusse, R. (2004). The Wnt signaling pathway in development and disease. *Annu Rev Cell Dev Biol*, 20:781–810.

[Lowry et al., 2005] Lowry, W., Blanpain, C., Nowak, J., Guasch, G., Lewis, L., and Fuchs, E. (2005). Defining the impact of beta-catenin/Tcf transactivation on epithelial stem cells. *Genes Dev*, 19(13):1596–1611.

[Lustig et al., 2002] Lustig, B., Jerchow, B., Sachs, M., Weiler, S., Pietsch, T., Karsten, U., van de Wetering, M., Clevers, H., Schlag, P., Birchmeier, W., and Behrens, J. (2002). Negative feedback loop of Wnt signaling through upregulation of conductin/axin2 in colorectal and liver tumors. *Mol Cell Biol*, 22(4):1184–1193.

[Mao et al., 2001] Mao, J., Wang, J., Liu, B., Pan, W., Farr, G. r., Flynn, C., Yuan, H., Takada, S., Kimelman, D., Li, L., and Wu, D. (2001). Low-density lipoprotein receptor-related protein-5 binds to Axin and regulates the canonical Wnt signaling pathway. *Mol Cell*, 7(4):801–809.

[Mendes, 1997] Mendes, P. (1997). Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3. *Trends Biochem Sci*, 22(9):361–363.

[Miaczynska et al., 2004] Miaczynska, M., Pelkmans, L., and Zerial, M. (2004). Not just a sink: endosomes in control of signal transduction. *Curr Opin Cell Biol*, 16(4):400–406.

[Mohammadi et al., 1997] Mohammadi, M., McMahon, G., Sun, L., Tang, C., Hirth, P., Yeh, B., Hubbard, S., and Schlessinger, J. (1997). Structures of the tyrosine kinase domain of fibroblast growth factor receptorin complex with inhibitors. *Science*, 276(5314):955–960.

[Mohammadi et al., 2005] Mohammadi, M., Olsen, S., and Ibrahimi, O. (2005). Structural basis for fibroblast growth factor receptor activation. *Cytokine Growth Factor Rev*, 16(2):107–137.

[Montalibet et al., 2005] Montalibet, J., Skorey, K., and Kennedy, B. (2005). Protein tyrosine phosphatase: enzymatic assays. *Methods*, 35(1):2–8.

[Moon et al., 2004] Moon, R., Kohn, A., De Ferrari, G., and Kaykas, A. (2004). Wnt and beta-catenin signalling: disease and therapies. *Nat Rev Genet*, 5(9):691–701.

[Panayotou et al., 1993] Panayotou, G., Gish, G., End, P., Truong, O., Gout, I., Dhand, R., Fry, M., Hiles, I., Pawson, T., and Waterfield, M. (1993). Interactions between SH2 domains and tyrosine-phosphorylated platelet-derived growth factor beta-receptor sequences: analysis of kinetic parameters by a novel biosensor-based approach. *Mol Cell Biol*, 13(6):3567–3576.

[Phillips and Cardelli, 2005] Phillips, A. and Cardelli, L. (2005). A graphical representation for the stochastic pi-calculus. In *Proc Bioconcur'05*.

[PID, 2006] PID (2006). `http://pid.nci.nih.gov/`.

[Pinto et al., 2003] Pinto, D., Gregorieff, A., Begthel, H., and Clevers, H. (2003). Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. *Genes Dev*, 17:1709–13.

[Potten et al., 1992] Potten, C., Kellett, M., Rew, D., and Roberts, S. (1992). Proliferation in human gastrointestinal epithelium using bromodeoxyuridine in vivo: data for different sites, proximity to a tumour, and polyposis coli. *Gut*, 33(4):524–529.

[Priami, 1995] Priami, C. (1995). Stochastic pi-calculus. *The Computer Journal*, 38(7):578–589.

[Regev et al., 2004] Regev, A., Panina, E. M., Silverman, W., Cardelli, L., and Shapiro, E. (2004). Bioambients: an abstraction for biological compartments. *Theor Comput Sci*, 325(1):141–167.

[Regev and Shapiro, 2002] Regev, A. and Shapiro, E. (2002). Cellular abstractions: Cells as computation. *Nature*, 419(6905):343.

[Regev and Shapiro, 2004] Regev, A. and Shapiro, E. (2004). *Modelling in Molecular Biology*. The pi-calculus as an abstraction for biomolecular systems. Springer.

[Reya et al., 2003] Reya, T., Duncan, A., Ailles, L., Domen, J., Scherer, D., Willert, K., Hintz, L., Nusse, R., and Weissman, I. (2003). A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature*, 423(6938):409–414.

[Rousset et al., 2001] Rousset, R., Mack, J., Wharton, K., Axelrod, J., Cadigan, K., Fish, M., Nusse, R., and Scott, M. (2001). Naked cuticle targets dishevelled to antagonize Wnt signal transduction. *Genes Dev*, 15(6):658–671.

[Rubinfeld et al., 1996] Rubinfeld, B., Albert, I., Porfiri, E., Fiol, C., Munemitsu, S., and Polakis, P. (1996). Binding of GSK3beta to the APC-beta-catenin complex and regulation of complex assembly. *Science*, 272(5264):1023–1026.

[Rulifson et al., 2000] Rulifson, E., Wu, C., and Nusse, R. (2000). Pathway specificity by the bifunctional receptor frizzled is determined by affinity for wingless. *Mol Cell*, 6(1):117–126.

[Salic et al., 2000] Salic, A., Lee, E., Mayer, L., and Kirschner, M. (2000). Control of beta-catenin stability: reconstitution of the cytoplasmic steps of the wnt pathway in Xenopus egg extracts. *Mol Cell*, 5(3):523–532.

[Sandilands et al., 2007] Sandilands, E., Akbazardeh, S., Vechionne, A., McEwan, D., Frame, M., and Heath, J. (2007). Src kinase dictates activation, trafficking and signalling dynamics of Fibroblast Growth Factor Receptors. *Curr Biol*, page in press.

[Sansom et al., 2004] Sansom, O., Reed, K., Hayes, A., Ireland, H., Brinkmann, H., Newton, I., Batlle, E., Simon-Assmann, P., Clevers, H., Nathke, I., Clarke, A., and Winton, D. (2004). Loss of Apc in vivo immediately perturbs Wnt signaling, differentiation, and migration. *Genes Dev*, 18(12):1385–1390.

[Sastry et al., 1995] Sastry, L., Lin, W., Wong, W., Di Fiore, P., Scoppa, C., and King, C. (1995). Quantitative analysis of Grb2-Sos1 interaction: the N-terminal SH3 domain of Grb2 mediates affinity. *Oncogene*, 11(6):1107–1112.

[Sato et al., 2004] Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nat Med*, 10(1):55–63.

[SBML, 2000] SBML (2000). `http://sbml.org/`.

[Schlessinger, 2004] Schlessinger, J. (2004). Fibroblast growth factor receptor pathway. *Sci STKE (Connections Map)*, page http://stke.sciencemag.org/cgi/cm/stkecm;CMP_15049.

[Schoeberl et al., 2002] Schoeberl, B., Eichler-Jonsson, C., Gilles, E., and Muller, G. (2002). Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat Biotechnol*, 20(4):370–375.

[Science STKE, 1996] Science STKE (1996). `http://stke.sciencemag.org/`.

[Shapiro et al., 2002] Shapiro, B., Levchenko, A., Meyerowitz, E., Wold, B., and Mjolsness, E. (2002). Cellerator: extending a computer algebra system to include biochemical arrows for signal transduction simulations. *Bioinformatics*, 19(5):677–678.

[Skolnik et al., 1993] Skolnik, E., Lee, C., Batzer, A., Vicentini, L., Zhou, M., Daly, R., Myers, M., Backer, J., Ullrich, A., White, M., and Schlessinger, J. (1993). The SH2/SH3 domain-containing protein GRB2 interacts with tyrosine-phosphorylated IRS1 and Shc: implications for insulin control of ras signalling. *EMBO J*, 12(5):1929–1936.

[Sorokin et al., 1994] Sorokin, A., Mohammadi, M., Huang, J., and Schlessinger, J. (1994). Internalization of fibroblast growth factor receptor is inhibited by a point mutation at tyrosine 766. *J Biol Chem*, 269(25):17056–17061.

[Tsang and Dawid, 2004] Tsang, M. and Dawid, I. (2004). Promotion and attenuation of FGF signaling through the Ras-MAPK pathway. *Sci STKE*, 2004(228):pe17.

[van de Wetering et al., 2002] van de Wetering, M., Sancho, E., Verweij, C., de Lau, W., Oving, I., Hurlstone, A., Batlle, E., Coudreuse, D., Haramis, A.-P., Tjorn-Pon-Fong, M., Moerer, P., van den Born, M., Soete, G., Pals, S., Eilers, M., Medema, R., and Clevers, H. (2002). The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell*, 111:241–250.

[van Leeuwen et al., 2007] van Leeuwen, I., Byrne, H., Jensen, O., and King, J. (2007). Elucidating the interactions between the adhesive and transcriptional functions of beta-catenin in normal and cancerous cells. *J Theor Biol*, 247(1):77–102.

[Varma et al., 1999] Varma, A., Morbidelli, M., and Wu, H. (1999). *Parametric Sensitivity in Chemical Systems*. Cambridge University Press.

[Ware et al., 1997] Ware, M., Tice, D., Parsons, S., and Lauffenburger, D. (1997). Over-expression of cellular Src in fibroblasts enhances endocytic internalization of Epidermal Growth Factor receptor. *J Biol Chem*, 272(48):30185–30190.

[Wawra et al., 2007] Wawra, C., Kuhl, M., and Kestler, H. (2007). Extended analyses of the Wnt/beta-catenin pathway: Robustness and oscillatory behaviour. *FEBS Lett*, 581(21):4043–4048.

[Wichert et al., 2004] Wichert, S., Fokianos, K., and Stimmer, K. (2004). Identifying periodically expressed transcripts in microarray time series data. *Bioinformatics*, 20(1):5–20.

[Willert et al., 2003] Willert, K., Brown, J., Danenberg, E., Duncan, A., Weissman, I., Reya, T., Yates, J. r., and Nusse, R. (2003). Wnt proteins are lipid-modified and can act as stem cell growth factors. *Nature*, 423(6938):448–452.

[Willert et al., 1999] Willert, K., Shibamoto, S., and Nusse, R. (1999). Wnt-induced dephosphorylation of axin releases beta-catenin from the axin complex. *Genes Dev*, 13(14):1768–1773.

[Wnt, 2008] Wnt (2008). `http://www.cs.bham.ac.uk/~oxt/wntmap.html`.

[Wong et al., 2001] Wong, E., Lim, J., Low, B., Chen, Q., and Guy, G. (2001). Evidence for direct interaction between Sprouty and Cbl. *J Biol Chem*, 276(8):5866–5875.

[Wong et al., 2002] Wong, W., Mandir, N., Goodlad, R., Wong, B., Garcia, S., Lam, S., and Wright, N. (2002). Histogenesis of human colorectal adenomas and hyperplastic polyps: the role of cell proliferation and crypt fission. *Gut*, 50(2):212–217.

[Xu and Goldfarb, 2001] Xu, H. and Goldfarb, M. (2001). Multiple effector domains within SNT1 coordinate ERK activation and neuronal differentiation of PC12 cells. *J Biol Chem*, 276(16):13049–13056.

[Yamada et al., 2004] Yamada, S., Taketomi, T., and Yoshimura, A. (2004). Model analysis of difference between EGF pathway and FGF pathway. *Biochem Biophys Res Commun*, 314(4):1113–1120.

[Yamamoto et al., 1999] Yamamoto, H., Kishida, S., Kishida, M., Ikeda, S., Takada, S., and Kikuchi, A. (1999). Phosphorylation of axin, a Wnt signal negative regulator, by glycogen synthase kinase-3beta regulates its stability. *J Biol Chem*, 274(16):10681–10684.

Appendix A

Implementation of elementary signalling modules

A.1 Receptor module

```
-language(spifcp).

%public channels
public(bind(0.001), rel(0.001)) .
public(basalphos(0.0005), phos(0.1)) .
public(syn(0.5), deg(0.005)) .

System(N1, N2) ::= <<
        CREATE_A(N1) | CREATE_S(N2) | Syn | Clock .

CREATE_A(N)  ::= {N =< 0}, true ; {N > 0}, {N--} | A | self .
CREATE_S(N)  ::= {N =< 0}, true ; {N > 0}, {N--} | S | self >> .

S   ::= bind ! [], 0 . %stimulus binding to the receptor

A   ::= bind ? [], AS ;
basalphos ! [], AP . %low-level phosphorylation
AS  ::= rel ! [], (A | S) ; %receptor complex dissociation
        phos ! [], APS ; %phosphorylation of the bound receptor
        basalphos ! [], APS .
APS ::= rel ! [], (AP | S) ;
        deg ! [], 0 . %degradation of the phosphorylated receptor complex
AP  ::= bind ? [], APS ;
        deg ! [], 0 . %degradation of the phosphorylated receptor

Syn ::= syn ! [], (A | Syn) . %constant synthesis of non-active receptor
```

```
Clock ::= rel ? [], Clock ;
        syn ? [], Clock ;
        deg ? [], Clock ;
        phos ? [], Clock ;
        basalphos ? [], Clock .
```

A.2 Positive feedback module

```
-language(spifcp).

%public channels
public(bind(0.1), rel(0.001)) .
public(basalphos(0.0005), phos(0.1) .
public(syn(0.005), deg(0.005)).

System(N1, N2) ::= <<
        CREATE_A(N1) | CREATE_S(N2) | Clock .

CREATE_A(N)  ::= {N =< 0}, true ; {N > 0}, {N--} | A | self .
CREATE_S(N)  ::= {N =< 0}, true ; {N > 0}, {N--} | S | self >> .

S   ::= bind ! [], 0 .

A   ::= bind ? [], AS ;
basalphos ! [], AP . %low-level phosphorylation
AS  ::= rel ! [], (A | S) ; %receptor complex dissociation
        phos ! [], APS ; %phosphorylation of the bound receptor
        basalphos ! [], APS .
APS ::= rel ! [], (AP | S) ;
        syn ! [], (APS | A) ; %receptor synthesis in a feedback loop
        deg ! [], 0 . %degradation of the phosphorylated receptor complex
AP  ::= bind ? [], APS ;
        syn ! [], (AP | A) ;
        deg ! [], 0 .

Clock ::= rel ? [], Clock ;
        syn ? [], Clock ;
        deg ? [], Clock ;
        phos ? [], Clock ;
        basalphos ? [], Clock .
```

A.3 Negative feedback module

```
-language(spifcp).

%public channel declaration
public(inhibit(10.0), phos(0.05), dephos(0.001)).
public(asyn(0.2), adeg(0.1)).
%vary bsyn to reflect stimulus induction
public(bsyn(100.0), bdeg(0.01)).

System(N1, N2) ::= <<
        CREATE_A(N1) | CREATE_B(N2) | Syn | Clock .

CREATE_A(N) ::= {N=<0}, 0 ; {N>0}, {N--} | A | self .
CREATE_B(N) ::= {N=<0}, 0 ; {N>0}, {N--} | B | self >> .

A ::= << rel(0.1), deg(5.0), a(infinite), b(infinite) .
        phos ! [], AP ; %inhibitor activation
        adeg ! [], true  . %low-level degradation
AP ::= inhibit ! {rel, deg, a, b}, AP_Bound ; %inhibition of B
        dephos ! [], A ; %deactivation of the inhibitor
        adeg ! [], true .
AP_Bound ::= rel ? [], AP ; %dissociation from B
        deg ? [], AP ;
        b ? [], AP ;
        adeg ! [], a ! [], true >> .

B ::= << rel(0.1), deg(5.0), a(infinite), b(infinite).
        inhibit ? {rel, deg, a, b}, BA ;
        asyn ! [], (B | A) ; %inhibitor synthesis in a feedback loop
        bdeg ! [], true . %low-level degradation
BA ::= rel ! [], B ;
        a ? [], B ;
        deg ! [], true ;
        bdeg ! [], b ! [], true >> .

Syn ::= bsyn ! [], (Syn | B) . %constant synthesis of activator molecules

Clock ::= phos ? [], Clock ;
        dephos ? [], Clock ;
        asyn ? [], Clock ;
        adeg ? [], Clock ;
        bsyn ? [], Clock ;
        bdeg ? [], Clock .
```

Appendix B

Modelling FGF pathway

```
-language(spifcp).

FAST=>10000.0.

%public channels
public(bindFGF(0.005)).
public(bindFGFR(0.0025), bindPlc(0.0025)).
public(bindSrc(0.0025), bindGrb(0.0025), bindShp(0.0025), bindGSpry(0.0025)).
public(bindSpry(0.01), bindCbl(0.01), bindSos(0.01)).
public(ph653(0.013), ph766(0.004), phFRS(0.005), phSpry(10), dphFRS(12)).
public(create_spry(0.083)).

%Initial process accepts three arguments:
%      #1: the initial number of FGFR, Src, Grb2, Shp, Plc, and Cbl molecules
%              and the maximum number of Spry molecules
%      #2: the initial number of FGF molecules
%      #3: the initial number of FRS2 molecules
System(N1,N2,N3) ::= << CREATE_FGFR(N1) | CREATE_FGF(N2)   | CREATE_FRS(N3) |
       CREATE_SRC(N1) | CREATE_SHP(N1)  | CREATE_GRB(N1)   | CREATE_PLC(N1) |
       CREATE_CBL(N1) | CREATE_SOS(N1)  | CREATE_DSPRY(N1) | Clock .

CREATE_FGF(N)   ::= {N =< 0}, true ; {N > 0}, {N--} | FGF  | self .
CREATE_FGFR(N)  ::= {N =< 0}, true ; {N > 0}, {N--} | FGFR | self .
CREATE_FRS(N)   ::= {N =< 0}, true ; {N > 0}, {N--} | FRS2 | self .
CREATE_SRC(N)   ::= {N =< 0}, true ; {N > 0}, {N--} | Src  | self .
CREATE_GRB(N)   ::= {N =< 0}, true ; {N > 0}, {N--} | Grb2 | self .
CREATE_SHP(N)   ::= {N =< 0}, true ; {N > 0}, {N--} | Shp  | self .
CREATE_CBL(N)   ::= {N =< 0}, true ; {N > 0}, {N--} | Cbl  | self .
CREATE_PLC(N)   ::= {N =< 0}, true ; {N > 0}, {N--} | Plc  | self .
CREATE_SOS(N)   ::= {N =< 0}, true ; {N > 0}, {N--} | Sos  | self .
CREATE_DSPRY(N) ::= {N =< 0}, true ; {N > 0}, {N--} | Spry | Dself(N) .
```

```
Dself(N) ::= create_spry![], CREATE_DSPRY(N) >> .

FGF ::= << relFGF(0.005), remFGF(infinite), degFGF(infinite) . %private channels
        bindFGF ? {relFGF, remFGF, degFGF}, FGF_Bound . %binding to the receptor

FGF_Bound ::= relFGF![], FGF ; %dissociation from FGFR
        remFGF?[], FGF ;  %immediate dissociation
        degFGF?[], true >> . %degradation

FGFR ::= << relFGF(0.005), remFGF(infinite), degFGF(infinite),
        relFGFR(0.05), remFGFR(infinite), degFGFR(infinite),
        preFRS(FAST),
        relPlc(0.05), degPlc(0.00028), degPlcI(infinite),
        pre653(infinite), pre766(infinite), deg1(infinite), deg2(infinite) .

        %contains four sites that engage in independent activities
        FGFR_Ligand_Binding | FGFR_FRS2_Binding | FGFR_Y653 | FGFR_Y766 .

FGFR_Ligand_Binding ::= bindFGF ! {relFGF, remFGF, degFGF}, FGFR_Ligand; %binds FGF
        deg1?[], true; deg2?[], true. %synchronises with other sites on degradation
FGFR_Ligand ::= relFGF?[], FGFR_Ligand_Binding ; %dissociaties from FGF
        pre653![], FGFR_Ligand ; %allows phosphorylation of Y653 while bound
        deg1?[], remFGF![], true ; deg2?[], degFGF![], true .

FGFR_FRS2_Binding ::=
        bindFGFR ! {relFGFR, preFRS, remFGFR, degFGFR}, FGFR_FRS2 ; %binds FRS2
        deg1?[], true .
FGFR_FRS2 ::= relFGFR?[], FGFR_FRS2_Binding ; %FRS2 dissociation
        remFGFR?[], FGFR_FRS2_Binding ; %immediate dissociation from FRS2
        degFGFR?[], deg2![], deg2![], deg2![], true ; %removes all bound partners
        deg1?[], (remFGFR![], true ; remFGFR?[], true) .

FGFR_Y653 ::= pre653?[], (ph653![], FGFR_Y653P ; %phosphorylates after binding FGF
                deg1?[], true ; deg2?[], true ) ;
        deg1?[], true ; deg2?[], true .
FGFR_Y653P ::= preFRS![], FGFR_Y653P ; %synchronises with FRS2 for phosphorylation
        pre766![], FGFR_Y653P ; %synchronises with Y766 for phosphorylation
        deg1?[], true ; deg2?[], true .

%residue Y766 phosphorylates after phosphorylation on Y653 is complete
FGFR_Y766 ::= pre766?[], (ph766![], FGFR_Y766P ; deg2?[], true ) ;
        deg2?[], true .
FGFR_Y766P ::= bindPlc ! {relPlc, degPlc, degPlcI}, FGFR_Plc_Bound ;  %binds Plc
        deg2?[], true .
```

```
FGFR_Plc_Bound ::= relPlc?[], FGFR_Y766P ; %dissociates from Plc
        degPlc?[], ( deg1![], deg1![], deg1![], true ; %degrades leaving partners
                deg2?[], true ) ;
        deg2?[], degPlcI![], true >> .

FRS2 ::= << relFGFR(0.05), degFGFR(infinite), remFGFR(infinite),
        preFGFR(FAST), preFRS(FAST),
        relSrc(0.05), remSrc(infinite), degSrc(0.00111),
        relGrb(0.05), remGrb(infinite), degGrb(infinite),
        relShp(0.05), remShp(infinite), degShp(infinite),
        postFRS(infinite), degCbl(infinite), deg1(infinite), deg2(infinite) .

        FRS2_FGFR_Binding | FRS2_Y196 | FRS2_Y306 | FRS2_Y471. %4 independent sites

FRS2_FGFR_Binding ::=
        bindFGFR ? {relFGFR, preFGFR, remFGFR, degFGFR}, FRS2_FGFR ; %binds FGF
        deg1?[], true ; deg2?[], true . %synchronises with others for degradation
FRS2_FGFR ::= preFGFR?[], FRS2_FGFRP ; %notified of FGFR phosphorylation
        relFGFR![], FRS2_FGFR_Binding ; %dissociates from FGFR
        remFGFR?[], FRS2_FGFR_Binding ;
        deg1?[], remFGFR![], true ;
        deg2?[], (degFGFR![], true ; remFGFR?[], true) .
FRS2_FGFRP ::= relFGFR![], FRS2_FGFR_Binding ;
        remFGFR?[], FRS2_FGFR_Binding ;
        preFRS![], FRS2_FGFRP ;  %synchronises with other sites for phosphorylation
        deg1?[], remFGFR![], true ;
        deg2?[], (degFGFR![], true ;  remFGFR?[], true) .

FRS2_Y196 ::= preFRS?[], (phFRS![], FRS2_Y196P;  %phosphorylates if bound to FGFRP
                postFRS?[], FRS2_Y196 ) ;
        postFRS?[], FRS2_Y196 .
FRS2_Y196P ::= bindSrc ! {relSrc, degCbl, remSrc, degSrc}, FRS2_Src ; %binds Src
        postFRS?[], FRS2_Y196 .
FRS2_Src ::= relSrc?[], FRS2_Y196P ; %dissociates from Src
        degCbl?[], deg1![], deg1![], deg1![], true ; %leaves all bound partners
        degSrc?[], deg2![], deg2![], deg2![], true ; %removes all bound partners
        postFRS?[], remSrc![], FRS2_Y196 .

FRS2_Y306 ::= preFRS?[], (phFRS![], FRS2_Y306P ;
                postFRS?[], FRS2_Y306 ;
                deg1?[], true ; deg2?[], true) ;
        postFRS?[], FRS2_Y306 ;
        deg1?[], true ; deg2?[], true .
FRS2_Y306P ::= bindGrb ! {relGrb, remGrb, degGrb}, FRS2_Grb ; %binds Grb2
```

```
        postFRS?[], FRS2_Y306 ;
        deg1?[], true ; deg2?[], true .
FRS2_Grb ::= relGrb?[], FRS2_Y306P ; %dissociates from Grb
        postFRS?[], remGrb![], FRS2_Y306 ;
        deg1?[], remGrb![], true ;
        deg2?[], degGrb![], true .

FRS2_Y471 ::= preFRS?[], (phFRS![], FRS2_Y471P ;
                deg1?[], true ; deg2?[], true) ;
        deg1?[], true ; deg2?[], true .
FRS2_Y471P ::= bindShp ! {relShp, remShp, degShp}, FRS2_Shp ; %binds Shp
        deg1?[], true ; deg2?[], true .
FRS2_Shp ::= relShp?[], FRS2_Y471P ;
        dphFRS![], postFRS![], FRS2_Shp ;  %dephosphorylation of Y196
        dphFRS![], postFRS![], FRS2_Shp ;  %dephosphorylation of Y306
        dphFRS![], remShp![], FRS2_Y471 ;  %dephosphorylation of Y471
        deg1?[], remShp![], true ;
        deg2?[], degShp![], true >> .

Src ::= << relSrc(0.05), remSrc(infinite), degSrc(0.00111),
        relSpry(0.01), relSpryP(0.0001), remSpry(infinite), degSpry(infinite),
        degCbl2(infinite), degCbl3(infinite).

        %Src contains independent FRS2 and Spry binding sites
        Src_FRS_Binding | Src_Spry_Binding .

Src_FRS_Binding ::=
        bindSrc ? {relSrc, degCbl3, remSrc, degSrc}, Src_FRS ;  %binds FRS2 Y196
        degCbl2?[], Src_FRS_Binding .
Src_FRS ::= relSrc![], Src_FRS_Binding ; %dissociation from Src
        remSrc?[], Src_FRS_Binding ;
        degSrc![], degSpry![], true ;
        degCbl2?[], ( degCbl3![], Src_FRS_Binding ; %pass instructions to FRS2
                 remSrc?[], Src_FRS_Binding ;
                 degSrc![], degSpry![], true ) .

Src_Spry_Binding ::= bindSpry ! {relSpry, relSpryP, remSpry, degCbl2},  Src_Spry;
        degSpry?[], true .
Src_Spry ::= relSpry?[], Src_Spry_Binding ;
        relSpryP?[], Src_Spry_Binding ;
        degSpry?[], remSpry![], true >> .

Spry ::= << relSpry(0.01), relSpryP(0.0001), remSpry(infinite),
        relGSpry(0.05), degGrb(infinite), remGrb(infinite),
```

```
        relCbl(0.0001), remCbl(infinite), degCbl1(0.00067),
        degCbl(infinite), degCbl2(infinite), deg(infinite),
        preSpry(infinite), postSpry(infinite) .

        %independent binding to Src, Cbl and Grb
        Spry_Src_Binding | Spry_Cbl_Binding | Spry_Grb_Binding .

Spry_Grb_Binding ::= preSpry?[], SpryP_Grb_Binding ; %binds Grb if phosphorylated
        deg?[], true .
SpryP_Grb_Binding ::= bindGSpry ! {relGSpry, remGrb, degGrb}, SpryP_Grb ;
        postSpry?[], Spry_Grb_Binding ;
        deg?[], true .
SpryP_Grb ::= relGSpry?[], SpryP_Grb_Binding ; %dissociates from Grb
        postSpry?[], ( remGrb![], Spry_Grb_Binding ; deg?[], degGrb![], true) ;
        deg?[], degGrb![], true .

Spry_Cbl_Binding ::= preSpry?[], SpryP_Cbl_Binding ; %binds Cbl if phosphorylated
        deg?[], true .
SpryP_Cbl_Binding ::= bindCbl ! {relCbl, remCbl, degCbl}, SpryP_Cbl ;
        postSpry?[], Spry_Cbl_Binding ;
        deg?[], true .
SpryP_Cbl ::= relCbl?[], SpryP_Cbl_Binding ; %dissociates from Cbl
        postSpry?[], ( remCbl![], Spry_Cbl_Binding ; deg?[], degCbl![], true) ;
        degCbl1![], SpryP_Cbl ;
        deg?[], degCbl![],  true .

Spry_Src_Binding ::= bindSpry ? {relSpry, relSpryP, remSpry, degCbl2}, Spry_Src ;
        postSpry![], Spry_Src_Binding .
Spry_Src ::= phSpry![], SpryP_Src ; %phosphorylates after binding to Src
        relSpry![], Spry_Src_Binding ; %dissociates from Src
        remSpry?[], deg![], deg![], true ;
        postSpry![], Spry_Src .
SpryP_Src ::=
        relSpryP![], SpryP_Src_Binding ; %slower dissociation from Src
        remSpry?[], deg![], deg![], true ;
        degCbl1?[], ( degCbl2![], SpryP_Src ; remSpry?[], deg![], deg![], true) ;
        preSpry![], SpryP_Src . %synchronisation for phosphorylation
SpryP_Src_Binding ::= bindSpry ? {relSpry, relSpryP, remSpry, degCbl2}, SpryP_Src;
        degCbl1?[], SpryP_Src_Binding ;
        preSpry![], SpryP_Src_Binding >> .

Cbl ::= << relCbl(0.0001), remCbl(infinite), degCbl(infinite) .
        bindCbl ? {relCbl, remCbl, degCbl}, Cbl_Bound .
```

```
Cbl_Bound ::= relCbl![], Cbl ; %binds SpryP
        remCbl?[], Cbl ;
        degCbl?[], true >> .

Grb2 ::= << relGrb(0.05), remGrb(infinite), relGSpry(0.05),
        degGrb(infinite), relSos(0.0001), remSos(infinite), degSos(infinite) .

        %competition between FRS2 and SpryP for binding Grb2
        bindGrb ? {relGrb, remGrb, degGrb}, Grb2_FRS ; %binds FRS2 Y306
        bindGSpry ? {relGSpry, remGrb, degGrb}, Grb2_Spry_Bound . %binds SpryP

Grb2_FRS ::= bindSos ! {relSos, remSos, degSos}, Grb2_FRS_Sos ; %binds Sos
        relGrb![], Grb2 ;
        remGrb?[], Grb2 ;
        degGrb?[], true .

Grb2_FRS_Sos ::= relSos?[], Grb2_FRS ;
        relGrb![], remSos![], Grb2 ;
        remGrb?[], remSos![], Grb2 ;
        degGrb?[], degSos![], true .

Grb2_Spry_Bound ::= relGSpry![], Grb2 ;
        degGrb?[], true ;
        remGrb?[], Grb2 >> .

Sos ::= << relSos(0.0001), remSos(infinite), degSos(infinite) .
        bindSos ? {relSos, remSos, degSos}, Sos_Grb . %binds FRS2:Grb2

Sos_Grb ::= relSos![], Sos ; degSos?[], true ; remSos?[], Sos >> .

Shp ::= << relShp(0.05), remShp(infinite), degShp(infinite) .
        bindShp ? {relShp, remShp, degShp}, Shp_FRS . %binds FRS2 Y471

Shp_FRS ::= relShp![], Shp ; remShp?[], Shp; degShp?[], true >> .

Plc ::= << relPlc(0.05), degPlc(0.00028), degPlcI(infinite).
        bindPlc ? {relPlc, degPlc, degPlcI}, Plc_FGFR. %binds FGFR Y766

Plc_FGFR ::= relPlc![], Plc; degPlc![], true ; degPlcI?[], true >>.

%Complements communication on channels that correspond to first-order reactions
Clock ::= ph653?[], Clock ; ph766?[], Clock ;
        phFRS?[], Clock ; dphFRS?[], Clock ;
        phSpry?[], Clock ; create_spry?[], Clock .
```

Appendix C

Implementation of Wnt pathway model

```
-language(spifcp).

public(bindFz(0.02), bindLRP(0.02)).
public(bindWnt(1.0)).
public(bindGsk(0.0015)).
public(phAxin(0.2), dphAxin(0.1)).
public(bindAPC(0.001)).
public(phAPC(0.2), dphAPC(0.1)).
public(bindBeta(0.00003), bindBetaP(0.01)).
public(bindTCF(0.0003)).
public(phBeta(5.0)).
public(degBeta1(0.001)).
public(transcription(0.0111), synAxin(0.048)).
public(synBeta(0.423)).
%stochastic oscillations are at synWnt=0.05, deterministic are at synWnt=0.2
public(synWnt(0.05)).
public(degLig1(0.008)).

%Initial process takes three arguments:
%       #1 is the initial number of Fz, LRP, APC, Gsk, TCF molecules,
%       #2 is the initial number of Axin molecules, and
%       #3 is the initial number of beta-catenin molecules
System(N1,N2,N3) ::= <<
        CREATE_FZ(N1)  | CREATE_LRP(N1)  |  CREATE_APC(N1) | CREATE_GSK(N1) |
        CREATE_TCF(N1) | CREATE_Axin(N2) | CREATE_Beta(N3) | Clock | Syn .
CREATE_FZ(N)   ::= {N =< 0}, true ; {N > 0}, {N--} | Fz    | self .
CREATE_LRP(N)  ::= {N =< 0}, true ; {N > 0}, {N--} | LRP   | self .
CREATE_Axin(N) ::= {N =< 0}, true ; {N > 0}, {N--} | Axin  | self .
CREATE_APC(N)  ::= {N =< 0}, true ; {N > 0}, {N--} | APC   | self .
```

```
CREATE_GSK(N)  ::= {N =< 0}, true ; {N > 0}, {N--} | Gsk   | self .
CREATE_Beta(N) ::= {N =< 0}, true ; {N > 0}, {N--} | Beta  | self .
CREATE_TCF(N)  ::= {N =< 0}, true ; {N > 0}, {N--} | TCF   | self >> .

Wnt ::= << %private channels
        relFz(0.1), degLig(0.008), remFz(0.04),
        relLRP(0.1), remLRP(infinite),
        relWnt(0.001), remWnt(infinite), degWnt(infinite) .

        bindFz ! {relFz, degLig, remFz}, WntFz ; %binding to the receptor
        degLig1![], true . %degradation of free ligand

WntFz ::= bindLRP ! {relLRP, remLRP}, WntFzLRP ; %sequentially binds Fz and LRP
        relFz?[], Wnt ;
        degLig![], true . %ensures termination of signalling

%Activated receptor complex recruits Axin:Gsk
WntFzLRP ::= bindWnt ? {relWnt, remWnt, degWnt}, WntFzLRPAxinGsk ;
        relLRP?[], WntFz ;
        degLig![], remLRP![], true .

WntFzLRPAxinGsk ::= relWnt![], WntFzLRP ; %dissociates from Axin complex
        relFz?[], remLRP![], remWnt![], Wnt ; %dissociates from Fz
        degLig![], remLRP![], remWnt![], true ; %degrades leaving all partners
        remFz![], remLRP![], degWnt![], Wnt >> . %inhibits Axin

Fz ::= << relFz(0.1), degLig(0.008), remFz(0.04) .
        bindFz ? {relFz, degLig, remFz}, FzBound . %binds Wnt

FzBound ::= relFz![], Fz ; %dissociates from Wnt
        degLig?[], Fz ;
        remFz?[], Fz >> .

%LRP is activated by Wnt and recruits and deactivates Axin
LRP ::= << relLRP(0.1), remLRP(infinite) .
        bindLRP ? {relLRP, remLRP}, LRP_Bound . % bind Wnt:Fz complex

LRP_Bound ::= relLRP![], LRP ;
        remLRP?[], LRP >> .

Gsk ::= << relGsk(0.1), relGskP(0.01), remGsk(infinite) .
        bindGsk ? {relGsk, relGskP, remGsk}, Gsk_Bound . %kinase binds Axin
Gsk_Bound ::= relGsk![], Gsk ;
        relGskP![], Gsk ;
```

```
        remGsk?[], Gsk >> .

Axin ::= << prePh(infinite), postPh(infinite),
        degAxin(0.005), degAxinP(0.002), degIAxin(infinite),
        relGsk(0.1), relGskP(0.01), remGsk(infinite),
        relWnt(0.001), remWnt(infinite), degWnt(infinite),
        relAPC(0.1), remAPC(infinite) .

        %Axin molecule contains two sites that bind Gsk and APC independently
        Axin_Gsk_Binding | Axin_APC_Binding .

Axin_Gsk_Binding ::= bindGsk ! {relGsk, relGskP, remGsk}, Axin_Gsk ; %binds Gsk
        degAxin![], true ;
        postPh![], Axin_Gsk_Binding . %synchronises for dephosphorylation
Axin_GskP_Binding ::= bindGsk ! {relGsk, relGskP, remGsk}, Axin_GskP ;
        dphAxin![], Axin_Gsk_Binding ; %Axin dephosphorylation
        degAxinP![], true .
Axin_Gsk ::= phAxin![], Axin_GskP ; %Gsk-bound Axin undergoes phosphorylation
        %Gsk-bound form of Axin is recruited to the activated receptor complex
        bindWnt ! {relWnt, remWnt, degWnt}, Axin_Wnt_Gsk ;
        relGsk?[], Axin_Gsk_Binding ;
        degAxin![], remGsk![], true ;
        postPh![], Axin_Gsk .
Axin_GskP ::= relGskP?[], Axin_GskP_Binding ; %slow dissociation of AxinP
        dphAxin![], Axin_Gsk_Binding ;
        bindWnt ! {relWnt, remWnt, degWnt}, Axin_Wnt_GskP ;
        degAxinP![], remGsk![], true ; %slow degradation when phosphorylated
        prePh![], Axin_GskP . %synchronisation on phosphorylation
Axin_Wnt_Gsk ::= relWnt?[], Axin_Gsk ;
        remWnt?[], Axin_Gsk ;
        degWnt?[], remGsk![], degIAxin![], true ;
        postPh![], Axin_Wnt_Gsk .
Axin_Wnt_GskP ::= relWnt?[], Axin_GskP ;
        remWnt?[], Axin_GskP ;
        degWnt?[], remGsk![], degIAxin![], true .

Axin_APC_Binding ::= prePh?[], AxinP_APC_Binding ; %awaits phosphorylation by Gsk
        degAxin?[], true ; degAxinP?[], true ; %sites synchronise for degradation
        degIAxin?[], true .
AxinP_APC_Binding ::= bindAPC ! {relAPC, remAPC}, AxinP_APC ; %binding to APC
        degAxin?[], true ; degAxinP?[], true ;
        degIAxin?[], true ;
        postPh?[], Axin_APC_Binding . %first site is no longer phosphorylated
Axin_APC ::= relAPC?[], Axin_APC_Binding ; %dissociation from APC
```

```
        degAxin?[], remAPC![], true ; degAxinP?[], remAPC![], true ;
        degIAxin?[], remAPC![], true ;
        prePh?[], AxinP_APC .
AxinP_APC ::= relAPC?[], AxinP_APC_Binding ;
        degAxin?[], remAPC![], true ; degAxinP?[], remAPC![], true ;
        degIAxin?[], remAPC![], true ;
        postPh?[], Axin_APC >> .

APC ::= << prePh(infinite), postPh(infinite),
        relAPC(0.1), remAPC(infinite),
        relBeta(0.1), degBeta(0.001), degBetaP(5.0) .

        %contains two independent binding site to bind Axin and beta-catenin
        APC_Axin_Binding | APC_Beta_Binding .

APC_Axin_Binding ::= bindAPC ? {relAPC, remAPC}, APC_Axin ; %binds AxinP
        postPh![], APC_Axin_Binding .
APC_AxinP_Binding ::= bindAPC ? {relAPC, remAPC}, APC_AxinP ;
        dphAPC![], APC_Axin_Binding .
APC_Axin ::= phAPC![], APC_AxinP ; %phosphorylates when bound to AxinP
        relAPC![], APC_Axin_Binding ; %dissociates from AxinP
        remAPC?[], APC_Axin_Binding ;
        postPh![], APC_Axin . %synchronises with another site on dephosphorylation
APC_AxinP ::= dphAPC![], APC_Axin ; %dephosphorylates
        relAPC?[], APC_AxinP_Binding ;
        remAPC?[], APC_AxinP_Binding ;
        prePh![], APC_AxinP . %synchronises with another site on phosphorylation

%The rate of beta-catenin binding depends on the APC phosphorylation state
APC_Beta_Binding ::= bindBeta ! {relBeta, degBeta, degBetaP}, APC_Beta ;
        prePh?[], APC_BetaP_Binding . %two sites synchronise on phosphorylation

APC_BetaP_Binding ::= bindBetaP ! {relBeta, degBeta, degBetaP}, APC_BetaP ;
        postPh?[], APC_Beta_Binding .
APC_Beta ::= relBeta?[], APC_Beta_Binding ;
        degBeta?[], APC_Beta_Binding ; %slow degradation when unphosphorylated
        prePh?[], APC_BetaP .
APC_BetaP ::= phBeta![], APCP_BetaP ; %phosphorylate when bound to APC
        relBeta?[], APC_BetaP_Binding ;
        degBeta?[], APC_BetaP_Binding ;
        postPh?[], APC_Beta .
APCP_BetaP ::= degBeta?[], APC_BetaP_Binding ;
        degBetaP?[], APC_BetaP_Binding >> .%phospho-beta-catenin rapidly degrades
```

```
Beta ::= << relTCF(0.01), degBeta(0.001), relBeta(0.1), degBetaP(5.0) .
        bindTCF ? {relTCF, degBeta}, Beta_TCF ; %beta-catenin can bind TCF
        bindBeta ? {relBeta, degBeta, degBetaP}, Beta_APC ; %binds APC slowly
        bindBetaP ? {relBeta, degBeta, degBetaP}, Beta_APC ; %binds APCP rapidly
        degBeta1![], true . %low-level degradation of free beta-catenin
Beta_TCF ::= relTCF![], Beta ; %dissociates from TCF
        degBeta![], true .
Beta_APC ::= relBeta![], Beta ; %dissociates from APC
        degBeta![], true ;
        degBetaP![], true >> . %rapidly degrades when phosphorylated

TCF ::= << relTCF(0.01), degBeta(0.001) .
        bindTCF ! {relTCF, degBeta}, TCFBound . %TCF binding to beta-catenin
TCFBound ::= relTCF?[], TCF ;
        degBeta?[], TCF ; %synchronises with beta-catenin for degradation
        transcription![], TCFTrans . %TCF initiates transcription of target genes
TCFTrans ::= synAxin![], (TCFTrans | Axin) ; %target genes include Axin
        relTCF?[], TCF ; degBeta?[], TCF >> .

%allows constant synthesis of beta-catenin and Wnt molecules
Syn ::= synBeta![], (Beta | Syn) ; synWnt![], (Wnt | Syn) .

%Clock complements all transitions corresponding to first-order reactions
Clock ::= degLig1?[], Clock ;
        phAxin?[], Clock ; dphAxin?[], Clock ;
        phAPC?[], Clock ; dphAPC?[], Clock ;
        phBeta?[], Clock ; degBeta1?[], Clock ; transcription?[], Clock ;
        synAxin?[], Clock ; synBeta?[], Clock ; synWnt?[], Clock .
```

Appendix D

Multi-scale model of cellular decisions

```
-language(biospi).

FAST=>1000.0.
SLOW=>60.0.

public(synBeta(0.423), bindBeta(0.01), relBeta(0.1), degBetaP(5), degBeta(0.001)).
public(synAxin(0.012),phAxin(0.2),dphAxin(0.1),degAxin(0.00167),breakAxin(infinite)).
public(bindWnt(1), relWnt(0.008), degAxinWnt(0.008), degWnt(0.005)).
public(molcycle(0.0001), moldiff(0.00001), degCell(infinite), dieCell(SLOW)).
public(cycleStem(0.0014), cycleDiff(0.0014), dieStem(0.0001), dieDiff(0.0005)) .
public(one(15), two(30), three(80), four(50), five(30), six(10), seven(1), eight(0),
       nine(0), ten(0), nil(1), latticeA(infinite), latticeQ(infinite), count(FAST)).

System ::= << move(infinite), next1(infinite), next2(infinite), next3(infinite),
        next4(infinite), next5(infinite), next6(infinite), next7(infinite),
        next8(infinite), next9(infinite) .

        %initially a crypt contains a row of ten cells (four stem cells located
        %at the crypt bottom and six differentiated cells at the crypt top)
        cell(<<SCell(one,move,next1)    | Axin10 | clk(<<CellClock>>) | CellSyn>>) |
        cell(<<SCell(two,next1,next2)   | Axin10 | clk(<<CellClock>>) | CellSyn>>) |
        cell(<<SCell(three,next2,next3) | Axin10 | clk(<<CellClock>>) | CellSyn>>) |
        cell(<<SCell(four,next3,next4)  | Axin10 | clk(<<CellClock>>) | CellSyn>>) |
        cell(<<DCell(five,next4,next5)  | Axin10 | clk(<<CellClock>>) | CellSyn>>) |
        cell(<<DCell(six,next5,next6)   | Axin10 | clk(<<CellClock>>) | CellSyn>>) |
        cell(<<DCell(seven,next6,next7) | Axin10 | clk(<<CellClock>>) | CellSyn>>) |
        cell(<<DCell(eight,next7,next8) | Axin10 | clk(<<CellClock>>) | CellSyn>>) |
        cell(<<DCell(nine,next8,next9)  | Axin10 | clk(<<CellClock>>) | CellSyn>>) |
        cell(<<DCell(ten,next9,nil)     | Axin10 | clk(<<CellClock>>) | CellSyn>>) |
```

```
        Lattice | Clock | CellCounter >> .

%Every cell contains a molecular network which evolves with time depending on the
%level of the environmental signal to which cell is exposed. The intracellular
%signalling network is based on the simplified model of the Wnt pathway.
Rec ::= merge+ bindWnt, Rec_Axin ;
        s2s degWnt![], true ;
        c2p degCell?[], true .
Rec_Axin ::= local relWnt?[], expel breakAxin, Rec ;
        local degAxinWnt?[], Rec ;
        local degWnt![], true ;
        c2p degCell?[], true .

Axin ::= s2s phAxin![], AxinP ;
        s2s degAxin![], true ;
        c2p degCell?[], true .
AxinP ::= merge- bindBeta, AxinP_Beta ;
        merge- bindWnt, AxinP_Rec ;
        s2s dphAxin![], Axin ;
        s2s degAxin![], true ;
        c2p degCell?[], true .
AxinP_Beta ::= local relBeta![], molAxin(<<exit breakAxin, AxinP>>) ;
        local degBetaP![], AxinP ;
        local degBeta?[], AxinP ;
        local degAxin![], true ;
        c2p degCell?[], true .
AxinP_Rec ::= local relWnt![], molAxin(<<exit breakAxin, AxinP>>) ;
        local degAxinWnt![], true ;
        local degWnt?[], AxinP ;
        c2p degCell?[], true .

Beta ::= merge+ bindBeta, Beta_Axin ;
        c2p synAxin![], Beta ;
        s2s degBeta![], true ;
        c2p molcycle![], Beta ;
        c2p moldiff![], Beta ;
        c2p degCell?[], true .
Beta_Axin ::= local relBeta?[], expel breakAxin, Beta ;
        local degBetaP?[], true ;
        local degBeta![], true ;
        local degAxin?[], Beta ;
        c2p degCell?[], true .

CellSyn ::= p2c synBeta![], (CellSyn | molBeta(<<Beta>>)) ;
```

```
        p2c synAxin?[], (CellSyn | molAxin(<<Axin>>)) .

CellClock ::= p2c phAxin?[], CellClock ;
        s2s dphAxin?[], CellClock ;
        s2s degAxin?[], CellClock ;
        s2s degBeta?[], CellClock ;
        s2s degWnt?[], CellClock ;
        c2p synBeta?[], CellClock .

%Every newly created cell contains 10 Axin molecules and no beta-catenin
Axin10 ::= molAxin(<<Axin>>) | molAxin(<<Axin>>) | molAxin(<<Axin>>) |
        molAxin(<<Axin>>) | molAxin(<<Axin>>) | molAxin(<<Axin>>) |
        molAxin(<<Axin>>) | molAxin(<<Axin>>) | molAxin(<<Axin>>) |
        molAxin(<<Axin>>) .

%When beta-catenin level exceeds Th(molcycle), a stem cell divides
%producing two stem cells. When beta-catenin level exceeds Th(moldiff),
%a stem cell differentiates.
SCell(pos,move,next) ::= << next1(infinite) .
        %special case of next==nil corresponds to cell without upper neighbour
        << next=?=nil,
                ( c2p pos![], SDiffuse ;
                  s2s move ! {pos, next1}, SMoveNil ;
                  p2c molcycle?[], SCycleNil(pos, move, next) ;
                  p2c moldiff?[], DCell(pos, move, next) ;
                  c2p dieStem![], SDie ) ;
           otherwise,
                ( c2p pos![], SDiffuse ;
                  s2s move ! {pos, next}, SMove ;
                  p2c molcycle?[], SCycle(pos, move, next) ;
                  p2c moldiff?[], DCell(pos, move, next) ;
                  c2p dieStem![], SDie ) >> .

SDiffuse ::= SCell(pos,move,next) | molRec(<<Rec>>) .

%If a cell has no upper neighbour, it asks the Lattice process about the value of
%channel pos at the next position
SMoveNil ::= c2p latticeQ!{pos}, c2p latticeA?{posNext}, SCell(posNext,next1,nil) .
SMove ::= s2s next ? {posNext, nextNext}, SCell(posNext,next,nextNext) .

SCycleNil(pos, move, next) ::= c2p cycleStem ! {pos,move,next1}, SDivideNil ;
                c2p pos![], SCycleNilDiffuse ;
                s2s move ! {pos, next1}, SCycleNilMove .
SDivideNil ::= c2p latticeQ ! {pos}, c2p latticeA ? {posNext}, c2p count![],
```

```
                SCell(posNext, next1, nil) .
SCycleNilDiffuse ::= SCycleNil(pos, move, next) | molRec(<<Rec>>) .
SCycleNilMove ::= c2p latticeQ ! {pos}, c2p latticeA ? {posNext},
                SCycleNil(posNext, next1, nil).

SCycle(pos, move, next) ::= c2p cycleStem ! {pos,move,next}, SDivide ;
                c2p pos![], SCycleDiffuse ; %still receives environmental signal
                s2s move ! {pos, next}, SCycleMove.
SDivide ::= s2s next ? {posNext, nextNext}, c2p count![],
                SCell(posNext,next,nextNext) .
SCycleDiffuse ::= SCycle(pos, move, next) | molRec(<<Rec>>) .
SCycleMove ::= s2s next ? {posNext, nextNext}, SCycle(posNext, next, nextNext) .

%Dead cell is removed only after another cell is placed at its position
SDie ::= p2c degCell![], SDie ;
                c2p dieCell![], SDead .
SDead ::= s2s move ! {pos, next}, true >> .

%When beta-catenin level exceeds Th(molcycle), a differentiated cell divides
%producing two differentiated cells. Differentiated cell has limited lifespan.
DCell(pos,move,next) ::= << next1(infinite) .
        << next=?=nil,
                ( c2p pos![], DDiffuse ;
                  s2s move ! {pos, next1}, DMoveNil ;
                  p2c molcycle?[], DCycleNil(pos, move, next) ;
                  c2p dieDiff![], DDie ) ;
           otherwise,
                ( c2p pos![], DDiffuse ;
                  s2s move ! {pos, next}, DMove ;
                  p2c molcycle?[], DCycle(pos, move, next) ;
                  c2p dieDiff![], DDie ) >> .

DDiffuse ::= DCell(pos,move,next) | molRec(<<Rec>>) .

DMoveNil ::= c2p latticeQ!{pos}, c2p latticeA?{posNext}, DCell(posNext,next1,nil) .
DMove ::= s2s next ? {posNext, nextNext}, DCell(posNext,next,nextNext) .

DCycleNil(pos, move, next) ::= c2p cycleDiff ! {pos,move,next1}, DDivideNil ;
                c2p pos![], DCycleNilDiffuse ;
                s2s move ! {pos, next1}, DCycleNilMove .
DDivideNil ::= c2p latticeQ ! {pos}, c2p latticeA ? {posNext}, c2p count![],
                DCell(posNext, next1, nil) .
DCycleNilDiffuse ::= DCycleNil(pos, move, next) | molRec(<<Rec>>) .
DCycleNilMove ::= c2p latticeQ ! {pos}, c2p latticeA ? {posNext},
```

```
                DCycleNil(posNext, next1, nil).

DCycle(pos, move, next) ::= c2p cycleDiff ! {pos,move,next}, DDivide ;
                c2p pos![], DCycleDiffuse ;
                s2s move ! {pos, next}, DCycleMove.
DDivide ::= s2s next ? {posNext, nextNext}, c2p count![],
                DCell(posNext,next,nextNext) .
DCycleDiffuse ::= DCycle(pos, move, next) | molRec(<<Rec>>) .
DCycleMove ::= s2s next ? {posNext, nextNext}, DCycle(posNext, next, nextNext) .

DDie ::= p2c degCell![], DDie ;
                c2p dieCell![], DDead .
DDead ::= s2s move ! {pos, next}, true >> .

%Lattice abstraction used by cells to find a neighbouring position.
Lattice ::= p2c latticeQ ? {pos}, << pos=?=one, p2c latticeA ! {two}, Lattice ;
                             pos=?=two, p2c latticeA ! {three}, Lattice ;
                             pos=?=three, p2c latticeA ! {four}, Lattice ;
                             pos=?=four, p2c latticeA ! {five}, Lattice ;
                             pos=?=five, p2c latticeA ! {six}, Lattice ;
                             pos=?=six, p2c latticeA ! {seven}, Lattice ;
                             pos=?=seven, p2c latticeA ! {eight}, Lattice ;
                             pos=?=eight, p2c latticeA ! {nine}, Lattice ;
                             pos=?=nine, p2c latticeA ! {ten}, Lattice ;
                             otherwise, p2c latticeA ! {ten}, Lattice >> .

Clock ::= p2c cycleStem ? {pos,move,next}, ( Clock |
                cell(<<SCell(pos,move,next) | Axin10 | CellClock | CellSyn>>) ) ;
        p2c cycleDiff ? {pos,move,next}, ( Clock |
                cell(<<DCell(pos,move,next) | Axin10 | CellClock | CellSyn>>) ) ;
        p2c dieCell?[], Clock ;
        p2c dieStem?[], Clock ; p2c dieDiff?[], Clock ;
        p2c one?[], Clock ; p2c two?[], Clock ;
        p2c three?[], Clock ; p2c four?[], Clock;
        p2c five?[], Clock; p2c six?[], Clock;
        p2c seven?[], Clock; p2c eight?[], Clock;
        p2c nine?[], Clock; p2c ten?[], Clock .

%Auxiliary process tracking the number of cells created in the system
CellCounter ::= p2c count ? [], CellCounter | CellCounter .
```

Appendix E

Glossary

Catalyst is an element of the biochemical reaction that increases the rate of reaction without being permanently changed.

Data normalization is the adjustment of measured values to account for possible run-to-run and day-to-day variability in the assays.

ELISA (Enzyme-linked-immunosorbent assay) is an experimental technique used to measure amounts of a particular protein in solution, using specific antibodies to identify proteins. ELISA involves adsorbing or coupling capture of antibodies to a microtitre plate. Following protein capture, a target protein is detected, either directly or indirectly, through a labelled detection antibody.

Flow cytometry is a method in which fluorescence-intensity data are recorded from particles in solution as they flow past a detector.

Fluorescence speckle microscopy associates fluorophores with macromolecular structures that are tracked by live-cell imaging. The information in the dynamic behaviour of these speckles is converted into a quantitative spatio-temporal readout of cytoskeleton-polymer transport and turnover.

Gel electrophoresis is a method of using gels and an electric field to pull the sample through the gel. As they migrate through a gel, proteins and nucleic acids separate into bands according to size.

Hill exponent is an exponent (h) that traditionally quantifies the extent of cooperative binding of multiple proteins, but is also used to describe sigmoidal steepness.

Homeostasis defines the conditions of a system when it is able to maintain its essential variables within limits acceptable to its own function in the face of unexpected disturbances.

Immunoblot (also known as a **western blot**) is a standard method to determine protein abundance and state of modification in a given sample. Following gel-based separation by mass, proteins are transferred to a sheet in which they can be probed with target-specific antibodies.

Mass spectrometry is a physical method for determining molecular mass. It contains three general components: an ion source in which gas-phase molecular ions are produced from the analyte molecules, a mass analyser in which electrical or magnetic fields are used to separate the analyte ions by their different mass-to-charge ratios, and a detector for recording the separated ions.

Michaelis-Menten kinetics is an approximation of mass action kinetics used for enzyme-substrate interactions when the concentration of the substrate is in excess of the enzyme.

Nothern blotting is used to study the expression patterns of a specific type of RNA molecule as a relative comparison among a set of different samples of RNA. In this process RNA is separated based on size and is then transferred to a membrane that is then probed with a labeled complement of a sequence of interest.

Null hypothesis is a statement that is tested for possible rejection under the assumption that it is true.

P-value is the probability of obtaining a test-statistic at least as extreme as the one observed, assuming that the null hypothesis is true. It is effectively the probability of wrongly rejecting the null hypothesis when it is actually true.

PCR (Polymerase chain reaction) is an experimental technique which allows the exponential copying of part of a DNA molecule using a DNA polymerase enzyme that is tolerant to elevated temperatures. The amplified DNA can be quantified after a certain number of amplification cycles. Reverse transcription polymerase chain reaction (RT-PCR) refers to a two-step procedure in which the RNA strand is reversly transcribed into cDNA followed by amplification of the resulting DNA using PCR.

Phosphorylation is a modification of a protein state where phosphate groups are added to polypeptide chain. Phosphorylation often affects protein structure and function.