A SYSTEMS BIOLOGY APPROACH TO MULTI-SCALE MODELLING AND ANALYSIS OF PLANAR CELL POLARITY IN *DROSOPHILA MELANOGASTER* WING

A thesis submitted for the degree of
Doctor of Philosophy

By
Qian Gao

School of Information Systems, Computing and Mathematics
Brunel University
February 2013

Table of Contents

List of Figures

Abstract

Systems biology aims to describe and understand biology at a global scale where biological systems function as a result of complex mechanisms that happen at several scales. Modelling and simulation are computational tools that are invaluable for description, understanding and prediction these mechanisms in a quantitative and integrative way. Thus multi-scale methods that couple the design, simulation and analysis of models spanning several spatial and temporal scales is becoming a new emerging focus of systems biology.

This thesis uses an exemplar – Planar cell polarity (PCP) signalling – to illustrate a generic approach to model biological systems at different spatial scales, using the new concept of Hierarchically Coloured Petri Nets (HCPN). PCP signalling refers to the coordinated polarisation of cells within the plane of various epithelial tissues to generate sub-cellular asymmetry along an axis orthogonal to their apical-basal axes. This polarisation is required for many developmental events in both vertebrates and non-vertebrates. Defects in PCP in vertebrates are responsible for developmental abnormalities in multiple tissues including the neural tube, the kidney and the inner ear. In *Drosophila* wing, PCP is seen in the parallel orientation of hairs that protrude from each of the approximately 30,000 epithelial cells to robustly point toward the wing tip.

This work applies HCPN to model a tissue comprising multiple cells hexagonally packed in a honeycomb formation in order to describe the phenomenon of Planar Cell Polarity (PCP) in *Drosophila* wing. HCPN facilitate the construction of mathematically tractable, compact and parameterised large-scale models. Different levels of abstraction that can be used in order to simplify such a complex system are first illustrated. The PCP system is first represented at an abstract level without modelling details of the cell. Each

cell is then sub-divided into seven virtual compartments with adjacent cells being coupled via the formation of intercellular complexes. A more detailed model is later developed, describing the intra- and inter-cellular signalling mechanisms involved in PCP signalling. The initial model is for a wild-type organism, and then a family of related models, permitting different hypotheses to be explored regarding the mechanisms underlying PCP, are constructed. Among them, the largest model consists of 800 cells which when unfolded yields 164,000 places (each of which is described by an ordinary differential equation).

This thesis illustrates the power and validity of the approach by showing how the models can be easily adapted to describe well-documented genetic mutations in the *Drosophila* wing using the proposed approach including clustering and model checking over time series of primary and secondary data, which can be employed to analyse and check such multi-scale models similar to the case of PCP. The HCPN models support the interpretation of biological observations reported in literature and are able to make sensible predictions. As HCPN model multi-scale systems in a compact, parameterised and scalable way, this modelling approach can be applied to other large-scale or multi-scale systems.

Acknowledgements

The work presented in this thesis is the result of the combined effort and support of many individuals, and I have been fortunate to have had such knowledgeable, insightful, and caring friends, colleagues, and collaborators during my time at Brunel University.

I would like to thank my supervisor, Professor David Gilbert, for his guidance and support during my PhD. David has always enthusiastically supported my work and given me incredible freedom to pursue an interdisciplinary subject. I have been very fortunate to have had him as a supervisor, and I have learned a lot under his direction. It has been a privilege and a pleasure to work with him.

I would also like to thank Dr David Tree, who has displayed an infinite amount of patience with me, explaining biological concepts to me and shaping my research to remain focused on biologically relevant and significant questions. I have benefited greatly from his guidance and insight.

I would like to thank Professor Xiaohui Liu, who shows a kindness and a genuine concern for the well-being of his students.

This thesis depends heavily on the support from Professor Monika Heiner and her group, who collaborated with me on this work. Monika's group has been developing Snoopy, a software tool to construct, edit and simulate Petri Nets (PN) and Coloured Petri Nets (CPN) models, which has been the main model construction and simulation environment for this work. I would like thank Dr Fei Liu, who used to work in Monika's group, for his support in tackling technical issues in applying CPN in Snoopy.

Some of the clustering work presented in Chapter 5 of this thesis reflects a close collaboration with Daniele Maccagnola, who was responsible for developing and encoding

the clustering technique of DBScan with PCA during his visit to David's group.

I would like to particularly thank Dr Crina Grosan who have helped me immensely over the years, for her feedback, advice and support towards my research and thesis writing.

I would like to Ovidiu Parvu for discussions and suggestions on the way to speed up the Ordindary Differential Equations (ODEs) simulation using Matlab.

I continue to indebted to all of the research supervisors and mentors that I have been fortunate enough to work with during my academic career. I would like to thank Professor Wolfgang Marwan and Dr Allan Tucker for agreeing to serve on my defense committee and Dr Simon Taylor for agreeing to serve as the independent Chair for my viva voce.

For their friendship and support, I would like thank everyone in the Computational Systems and Synthetic Biology Research Group, both past and present. I would also like to thank Teresa Czachowska for her various support in organising academic meetings and trips for everyone in David's group.

I would like to thank everyone who has supported me in various ways for my academic career and life at Brunel University.

I am grateful for my parents' invaluable and incredible support for my personal and academic life. I have also been fortunate to receive funding support from a bursary from Brunel University.

Qian Gao

London, UK

February, 2013

Chapter 1

Introduction

1.1 Motivation

With the overwhelmingly growth of experimental data available, Biomodel Engineering (BME) has been considered as a promising approach to meet the need of studying and understanding biological systems, where at the heart of BME is mathematical (computational) modelling. Mathematical modelling in engineering, physics or chemistry has traditionally sought to extrapolate from existing information and underlying principles to create complete descriptions of various systems, which could be analysed and simulated. Often only incomplete abstracted hypotheses exist to explain observed complex patterning and functions. Thus the question is what level of information is required about a biological network in order to mathematically model the network hence to derive explanation about the behaviour of the (biological) system. Thus it should simultaneously characterise the complex and nonintuitive behaviour of a network, while reveal deficiencies in the model and suggest new experimental directions.

A large variety of modelling approaches have already been applied to model a wide array of biological systems (see [Heat 09] for a review). Until now many models have

been at the intracellular level, and indeed have largely ignored locality within the cell; however there is a need to increase the spatial scope of models of biological systems to enable descriptions at the intercellular (cell-cell), tissue, organ and even whole organism scales - moving from single cell to multi-cellular organisms. The motivation has come both from the increasing need for life scientists to use computational models to facilitate the investigation and understanding of multi-cellular systems, and the greater variety of data available at different scales. All these lead to multi-scale modelling, which is a field of study to solve physical problems which have important features at multiple scales, particularly multiple spatial and/or temporal scales. It has a significant impact on a diverse range of scientific and engineering disciplines, including material science, fluid dynamics, meteorology, chemistry and biology. Multi-scale modelling in Systems Biology needs to go beyond the traditional approach of modelling at just one spatial scale. Thus the challenges for modellers includes the development of suitable paradigms and associated tools to create coherent descriptions of biological systems by integrating several spatial scales, and methods for the simulation, analysis, and checking of the models in order ultimately to use them to predict the behaviour of the biological system when disturbed by, for example, mutations, drugs, or stresses. Modelling biological systems beyond one spatial scale introduces a series of challenges which should be addressed, such as repetition of components, variation of components, organisation of components, hierarchical organisation, communication of components, mobility/motility, replication of components, deletion of components, irregular/semi-regular organisation of components, dynamic grid size, differentiation of components, and pattern formation of components [Gao 12, Hein 12a] (see details in Chapter 2).

There are several issues that should be noted in modelling a multi-scale system, which

include: (i) the availability of data - due to technical limits time series data may not obtainable, e.g. Planar Cell Polarity (PCP) signalling occurs in the pupal stage when it is difficult to obtain such data; (ii) choice of a suitable modelling formalism by balancing between simplicity and expressiveness; (iii) computational challenges due to simulating a very large model, for example, simulating a model comprising a large number of biological entities can take hours; (iv) fitting a multi-scale model to data, where problems are mainly due to the lack of data and long simulation runtime; (v) searching for appropriate techniques to analyse and check a multi-scale model at different levels, for example, a biological system being modelled is at a lower level (e.g. biochemical level) but needs to be interpreted at a higher level (e.g. physiological level). All these issues are briefly discussed in Chapter 6.

This thesis has chosen Planar Cell Polarity (PCP) signalling in the *Drosophila* wing as a biological example to present how to develop, construct and analyse a multi-scale model as well as a multi-cellular one, while some of the issues addressed above have been covered.

Planar cell polarity (PCP) refers to the orientation of cells within the plane of the epithelium, orthogonal to the apical-basal polarity of the cells. This polarisation is required for many developmental events in both vertebrates and non-vertebrates. Defects in PCP in vertebrates are responsible for developmental abnormalities in multiple tissues including the neural tube, the kidney and the inner ear (reviewed in [Simo 08]). The signalling mechanisms underlying PCP have been studied most extensively in the epithelia of the fruit fly *Drosophila melanogaster* including the wing, the abdomen, the eye, and the bristles of the thorax. Furthermore, genesis of human cancer has implicated that PCP shares signalling components with the Wnt pathway [Cadi 97], which also motivates to get a better understanding of this pathway.

Genetic studies in the wing and eye in the 1990s led to the proposal of a PCP signalling pathway involving the PCP proteins Frizzled (Fz), Dishevelled (Dsh) and Prickle (Pk) (reviewed in [Wong 93]). In the late 1990s and 2000s further genetic analysis, including the discovery of more PCP proteins, e.g. Flamingo (Fmi) and Van-Gogh (Vang), and data on the sub-cellular localisation of these proteins in normal and mutant situations, led to the formulation of more complex biological models of PCP signalling. With the identification of key components involved in the signalling, the challenge in understanding PCP has moved to discover the connection of the known players in this cascade that will provide adequate explanation to the complex and various phenotypic behaviours of the system. One phenomenon in particular, referred to as domineering nonautonomy, whereby cell clones mutant for some PCP signalling components cause polarity disruptions of surrounding wild-type cells, has been the source of much confusion and debate in the literature [Adle 02, Stru 02d]. A secondary signalling molecule has been proposed to explain domineering nonautonomy (described in Chapter 3 and [Adle 02, Adle 00, Fant 03, Lawr 99, Lawr 02, Stru 97a, Stru 02a, Wehr 98, Zeid 99]), but despite an extensive search, no such molecule has yet been identified [Axel 11].

An alternative model proposes that the most important components that explain domineering nonautonomy have already been identified, but it is only through the complex system dynamics resulting from feedback effects of the signalling network that domineering nonautonomy arises [Axel 01, Cadi 97]. It is not immediately evident, however, that this feedback loop model can adequately explain the variety of behaviours exhibited by the system, and others have argued that it cannot account for some experimental results involving clones of cells mutant for some of the polarity signalling between interactions among the signalling components and the range of behaviours

exhibited by the system [Adle 02, Stru 02d]. Mathematical modelling is ideally suited to make explicit the link between interactions among the signalling components and the range of behaviours exhibited by the system.

Various studies postulated the existence of an unknown and as yet unidentified secreted morphogen signal to which the PCP proteins respond. Despite the lack of positive data identifying such a morphogen signal [Pove 05] such models remain popular in the literature. A second postulated upstream signal biasing PCP is associated with the proto-caderhin molecules Fat and Dachsous (Ds) whose activities are modulated by the golgi kinase Four-jointed (reviewed in [Barr 06, Yang 02]). It has thus been proposed that the function of the Fat and Ds proteins is to facilitate the transport of the distal PCP proteins, mainly Fmi, Fz and Dsh to the distal cell cortex where they are stabilized by interactions with proximal complexes, which is so-called polarised vesicle transport hypothesis (described in Chapter 3).

In the *Drosophila* wing, about 30,000 epithelial cells are hexagonally packed in a 2-dimensional honeycomb lattice. Signal transduction within each cell is coupled with inter-cellular communication through the formation of protein complexes, so that local signalling produces a global effect over the entire organ. Computational model of PCP thus should include the repetition of components in a two-level hierarchy of different geometries, permitting abstraction at the level of cells, with a static two-dimensional organisation which is different at each level. The higher inter-cellular level is that of cells in a rectangular honeycomb grid, representing the epithelium tissue, and the lower level being intra-cellular organisation represented by virtual compartments within one cell in a circular grid. Moreover the computational model should allow variations among cells in the form of clones of mutant cells which either lack a specific signalling protein, or include

an overexpression of a signalling protein.

Moreover, the computational model could be used to investigate properties of proposed biological models for PCP signalling, explicitly demonstrating that it can explain the often counter-intuitive, complex behaviours of the system, obviating the need for additional factors to explain these behaviours, and providing insight into the system dynamics that govern them. This led to an iterative process in which the inability of the model to reproduce certain phenotypes of the system would allow researchers to ask why the model was not able to reproduce the phenotype, and motivated additional areas for further investigation. As additional biological evidence emerged, the model could be expanded, permitting the biologists to then explore the implications of the new evidence on prior beliefs about the mechanisms underlying PCP signalling.

Furthermore, another question that has to be addressed concerns the interaction between the functional modules of polarity signalling. Not enough is known about how the feedback loop interacts with upstream signalling components, which serve to impart a directional bias on the system to align cell polarity along the global axis of the tissue [Axel 01, Ma 03, Stru 02b]. It has not yet been settled whether PCP signalling is established via one pathway (feedback loop interacting with the morphogen signal, or the polarised vesicle transport) or by the two acting in parallel. Therefore, this work not only studies the possible role that each pathway plays but also further investigates the possibility of these systems collaborating in parallel.

1.2 Thesis contributions

Contributions of the work presented in this thesis include:

- Application of a generic modelling approach, hierarchically coloured Petri nets (HCPN), to design and construct compact, parameterisable and scalable models of multi-scale and multi-cellular biological systems, illustrated in depth for Planar Cell Polarity (PCP) in the *Drosophila* wing.

- A family of HCPN models developed to describe the core machinery of PCP signalling, where HCPN have not yet been applied to model this biological system.

- An approach to analyse, visualise and model checking such multi-scale models which allows us to compare in-silico results against semi-quantitative experimental data, e.g. image obtained from biological laboratories.

- Exploration of the mechanisms underlying PCP, which includes testing various hypotheses proposed in literature using the family of HCPN models, and predicting additional phenotypes which could be testable in biological laboratories.

1.3 Publications

Publications raising from this work are listed as follows:

- Q. Gao, D. Gilbert, M. Heiner, F. Liu, D. Maccagnola and D. Tree (2012). Multi-scale Modelling and Analysis of Planar Cell Polarity in the *Drosophila* Wing. IEEE/ACM Transactions on Computational Biology and Bioinformatics, Vol 99. ISSN: 1545-5963, DOI: 10.1109/TCBB.2012.101.

- D. Maccagnola, E. Messina, Q. Gao and D. Gilbert (2012). A Machine Learning Approach for Generating Temporal Logic Classifications of Complex Model

Behaviours. The Proceedings of Winter Simulation Conference (WSC 2012), Berlin. http://dl.acm.org/citation.cfm?id=2429759.2430156.

- Q. Gao, F. Liu, D. Gilbert, M. Heiner and D. Tree (2011). A multi-scale approach to modeling planar cell polarity in *Drosophila* wing using hierarchically coloured Petri nets. The proceeding of the 9th International Conference on Computational Methods in Systems Biology (CMSB2011). Published in the ACM Digital Library, pp 209-218. ISBN: 978-1-4503-0817-5, DOI: 10.1145/2037509.2037538.

- Q. Gao, F. Liu, D. Tree and D. Gilbert (2011). Multi-cell Modelling Using Coloured Petri Nets Applied to Planar Cell Polarity. The proceeding of the International Workshop on Biological Processes & Petri Nets (BioPPN2011). ISSN: 1613-0073. http://ceur-ws.org/Vol-724/.

1.4 Outline of the thesis

This thesis is structured as follows: this chapter has provided the motivation and an overview of the work presented in this thesis. Chapter 2 provides a brief review of the literature regarding biomodel engineering (focused on mathematical modelling) and multi-scale modelling of biological systems, followed by Chapter 3 on the biological background of planar cell polarity and related work. Chapter 4 describes the application of modelling approach, hierarchically coloured Petri nets, by means of the PCP case study, followed by Chapter 5 on the analysis and model checking of the HCPN models of PCP, and finally the conclusion, future work and open problems in Chapter 6.

Chapter 2

Multi-scale modelling of biological systems

This chapter gives a brief review of Biomodel Engineering (BME), and a further review of mathematical modelling of biological systems. It then focuses on multi-scale modelling, including challenges, modelling formalisms and simulators.

2.1 Background

With the rapid growth of data being generated in biology, it has become an informational science. Thus it is necessary to organise the data into coherent models that describe system behaviours, which are subsequently used for simulation, analysis and prediction.

From a systems biology perspective, a biological system is a set of complex interactions (network structure) rather than many individual molecular components. A biological system comprises large numbers of functionally diverse, and frequently multi-functional sets of elements that interact selectively and nonlinearly to produce coherent behaviours. This can be anything from a simple biological process, such as a biochemical reaction cycle, a gene regulatory network or a signalling pathway in a cell, a tissue, an organ, an

entire organism, or even an ecological web.

The core of systems biology is mathematical modelling, which can involve, for example, the process of simulating an abstract model of a biological system to test hypotheses with in-silico experiments or to provide predictions to be tested by in-vitro and in-vivo studies. In order to achieve the goal of answering biological questions, computational models have to reliably depict a biological system and be able to predict its behaviour. Based on the schematic representation of its components and their links, the model is a description of the dynamic behaviour of the biological system, which equips the model with predictive power. However, a model is not a real or exact portrait of the biological system; it is rather a simplified description to assist in the analysis and enhance the understanding of the system. Thus it is often needed to identify key components and processes and to predict biological behaviour: such as which processes and proteins are most important for signalling; why certain genes are oncogenes or tumour suppressor genes; or what effects a particular experimental technology (e.g. RNA interference) or drug will have on a biological system.

2.2 Biomodel engineering

Biomodel Engineering (BME) is the science of designing, constructing and analysing computational models of biological systems [Gilb 09]. It takes place at the interface of computing science, mathematics, engineering and biology, which provides a systematic approach for designing, constructing and analysing computational models of biological systems. Some of its central concepts are inspired by efficient software engineering strategies. BME does not aim at engineering biological systems per se, but rather aims at describing their structure and behaviour, using computational tools and techniques in a

principled way. The two major application areas of BME are systems biology [Hood 03] and synthetic biology [Endy 05, Hein 06] (Figure 2.1). In the former, the aim is to design and construct models of existing biological systems, which explain observed properties and predict the response to experimental interventions; in the latter, BME is used as part of a general approach for designing and constructing synthetic biological systems with novel functionalities [Brei 10, Hein 12a].

Systems biology is the main focus of this thesis. At the heart of systems biology lies mathematical (computational) modelling, thus, the following further explores mathematical modelling.

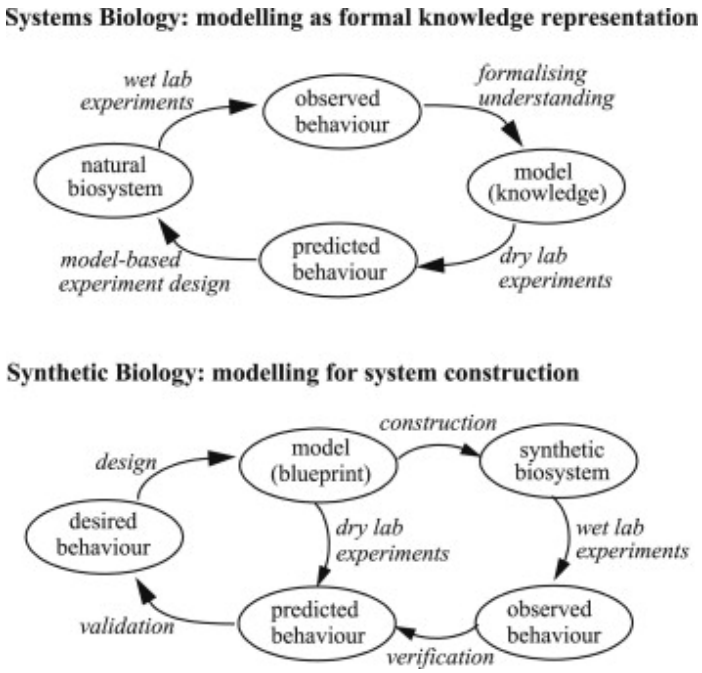

Figure 2.1: The role of formal models in systems and synthetic biology, adapted from [Hein 12a].

2.3 Mathematical modelling

Mathematical modelling in engineering, physics or chemistry has traditionally sought to extrapolate from existing information and underlying principles to create complete description of various systems, which could be analysed and simulated, and from which further abstractions could be made. However, in studying biological systems, it is often only incomplete abstracted hypotheses exist to explain observed complex patterning and functions. The challenge has become to show that enough of a network is understood to explain the complex and nonintuitive behaviour of a network, while revealing deficiencies in the model and suggesting new experimental directions.

While advances in molecular biology, genome sequencing and high-throughput data measurement have been driving factors to build system level models of biological systems [Idek 01, Kita 02a, Kita 02b], modelling methodology must account for the lack of complete data and understanding that will continue to characterise biological problems. The success of bioinformatics has contributed to the characterisation of biology as an informational science [Idek 01]. The next step is to successfully adapt tools from numerical analysis, control theory and optimisation to greater challenges of treating biological systems as complex dynamical systems. While traditional biology provides the logical connections describing the interactions between genes and molecules, it is the dynamics of these interactions that give rise to the relevant behaviour and function of the system. The idea of mathematical modelling of the dynamics of biological systems is not new [Turi 52], but complexity in biology has prevented the application of these methods which have largely been developed for engineered systems or for studying problems in physics and chemistry. Biological systems, however, are often characterised by a large number of interconnected components, many of which participate in several different functions

[Weng 99]. Advances in methods to deal with complexity [Weng 99] and a shift in biology toward quantitative data permits a transition where biology begins to be characterised also as a predictive science [Idek 01].

This view of biology forms the basis for the emergence of systems biology as an approach to study biological systems [Idek 01, Kita 02a, Kita 02b]. Systems biology aims to examine the structure of a biological system or function, and to simultaneously consider all of the relevant component parts and the dynamics of their interactions. The question remains, how to begin modelling a complex biological system. Traditionally, approaches have involved attempting to work from the *bottom-up*, describing and assembling the detailed components and interactions of a system until the complete system has been reconstructed. Alternatively, some approaches begin from the *top-down*, in which abstract models of the complete system are replaced by progressively more detailed models of components' parts. Each of these approaches suffers from various drawbacks. With many biological systems, working from the bottom-up may quickly result in computational intractability far before enough of a system has been sufficiently modelled to address relevant questions regarding that system [Nobl 03]. Furthermore, higher level processes often regulate low-level systems, such as controlling gene expression levels [Nobl 03]. In top-down approaches, high level models are often compatible with many lower level models, without a clear indication as to which choice of decomposition would lead to new insights and which would render the model intractable [Nobl 03]. There has been some efforts to address these issues, such as some efforts to construct a whole cell simulation of single cell organisms, in which the entire system was treated as a whole to understand system-wide functions [Scha 99, Tomi 99]. Another approach attempted to identify and characterise a large number of interactions that might possibly interact in a particular

pathway using high-throughput data collection in order to develop a complete description of the system. Motif finding efforts attempted to identify and characterise recurring patterns in these interactions, providing building blocks upon which abstract models of more general biological systems might be build [Milo 02]. Until these efforts succeed in establishing a basis for studying arbitrary biological system, however, it is left as a challenge for modeller to choose the appropriate scale of a model to address the specific questions that the model is meant to address. Here, the goal is to begin from the available experimental data and exploit existing knowledge of the biological system to work from the *middle-out* [Nobl 03] to construct a model capable of providing insight into the system and propose experiments that will allow parts of the model to be refined and to be extended to provide a more complete model of the system (Figure 2.1).

This systems biology approach promotes interactions between simulation and experiment that allow us to ask better experimental questions, make counter-intuitive predictions, and be able to explore far more complex system dynamics, for example where individual proteins may play multiple interacting roles in the system. This is particular true when trying to understand patterning effects in developing tissues, where often small changes affecting individual cellular processes can have widespread effects. Traditional biological approaches of identifying isolated components and their functions cannot always fully characterise the complete system behaviour necessary for cell and tissue level patterning. It is at this interface between biochemical interactions of the cell and tissue level patterning that the need for systems biology is evident. The critical challenge for such problems is identifying the proper level of detail in the model to adequately establish the nonintuitive link between events occurring at different levels of the system, while remaining computational tractable.

2.3.1 Modelling formalisms

Modelling efforts in systems biology typically lead to the simulation and analysis of models containing an enormous number of components and associated interactions. Non-linear mathematical models of biological processes and systems requires powerful numerical analysis methods to investigate their behaviours. A large variety of modelling approaches has been applied to model a wide array of biological systems [Heat 09] in order to meet the overwhelming need to organise biological data into models that describe system behaviour, which are mathematically tractable. This section describes briefly several modelling approaches which are widely used in systems biology to represent computational models of biological systems.

2.3.1.1 Ordinary differential equations

Ordinary differential equations (ODEs) is one of the most widely used modelling formalisms in systems biology [Chas 02, Tyso 03, Chen 99, Rizz 97], as well as one of the main modelling formalism used in this work. Therefore, a comprehensive review has been given here.

All living organisms and biochemical processes are governed by the laws of biochemistry. In order to support and maintain the complexity of life, biochemical processes control the information and energy flow through a variety of pathways (e.g. signalling and metabolic pathways), which require structures, functions and interactions of cellular components.

Biochemical systems theory (BST) [Sava 69a, Sava 69b, Sava 70] was developed in the late 60s in order to mathematically model biochemical systems using ODEs, in which biochemical processes were represented as power-law expansions in the variables of the

system. One main advantage of BST is that a set of equations can be set up without knowledge of the exact mechanism of each reaction, and identity of reactants and their reactional and regulatory interactions are feasible to design these biochemical models.

Biochemical systems consist of n molecular species (reactants), whose dynamic evolution is determined by the occurrence of m biochemical reactions. A biochemical system is fully characterised by the initial amount for each molecular species X_i ($1 \leqslant i \leqslant n$) and and the description of the biochemical reactions R_j ($1 \leqslant i \leqslant m$) with their kinetic rate laws [Ball 10]. In biochemical systems, the production or consumption of reactants are described by the biochemical reactions, presenting the regulations among these reactants. Biochemical reactions involve zero or more molecular species which can be either reactants or products. Stoichiometric coefficients associated with these reactions specify the number of molecules which are consumer or produced for each molecular species involved in the reactions.

An ordinary differential equation is defined as an equation involving one or more unknown functions and their derivations [Goss 11]. It is a description of how a property of interest, such as [A], the concentration of A, changes over time. This is usually expressed by describing how the rate of change of the concentration is related to the concentration at that moment. Taking a simple biochemical reaction as an example, the decay of substrate A to product B at rate k can be depicted by a simple mass action equation: $A \xrightarrow{k} B$, which can be translated into differential equations as:

$$\frac{d[A]}{dt} = -\frac{d[B]}{dt} = -k \times [A] \tag{2.3.1}$$

Furthermore, many biochemical reactions are reversible, for example $A \underset{k2}{\overset{k1}{\rightleftarrows}} B$, which can be described as the following differential equations:

$$\frac{d[A]}{dt} = -k1 \times [A] + k2 \times [B] \tag{2.3.2}$$

$$\frac{d[B]}{dt} = +k1 \times [A] - k2 \times [B] \tag{2.3.3}$$

Since biochemical reactions are often catalysed by enzymes which are not consumed during the process, they can extend the examples above as the following mass action description and differential equations: $A + E \underset{k2}{\overset{k1}{\rightleftarrows}} A|E \overset{k3}{\rightarrow} B + E$

$$\frac{d[A]}{dt} = -k1 \times [A] \times [E] + k2 \times [A|E] \tag{2.3.4}$$

$$\frac{d[A|E]}{dt} = +k1 \times [A] \times [E] - k2 \times [A|E] - k3 \times [A|E] \tag{2.3.5}$$

$$\frac{d[B]}{dt} = +k3 \times [A|E] \tag{2.3.6}$$

$$\frac{d[E]}{dt} = -k1 \times [A] \times [E] + k2 \times [A|E] + k3 \times [A|E] \tag{2.3.7}$$

Where E is the enzyme and A|E is the substrate-enzyme complex.

Enzymes serve a wide variety of functions inside living organisms. They are indispensable for signal transduction and cell regulation, often via kinases and phosphatases [Hunt 95]. Mass action kinetics [Horn 72] are often used for modelling reactions within signalling pathways whereas Michaelis-Menten kinetics [Ment 13] are often used in modelling the metabolic pathways. This latter kinetic model is relevant to situations where very simple kinetics can be assumed, holding at the initial stage of a reaction before the concentration of the product is appreciable, and makes the assumptions

that the concentration of product is close to zero, no product reverts to the initial substrate and the concentration of the enzyme is much less than the concentration of the substrate [Gilb 09]. In the following equations $[E_T]$ is the total enzyme concentration, V is the reaction velocity, V_{max} is the maximum reaction velocity, and k_M is the Michaelis constant, which is the substrate concentration required for an enzyme to reach one-half its maximum velocity.

$$V = V_{max} \times \frac{[A]}{k_M + [A]} \tag{2.3.8}$$

$$k_{cat} = k3 = \frac{V_{max}}{[E_T]} \tag{2.3.9}$$

$$k_M = \frac{k2 + k3}{k1} \tag{2.3.10}$$

The advantage of using the Michaelis-Menten kinetics is that it enables a single differential equation to describe the enzymatic reaction above:

$$\frac{d[A]}{dt} = -\frac{d[B]}{dt} = -k_{cat} \times [E_T] \times \frac{[A]}{k_M + [A]} \tag{2.3.11}$$

Although ODE-based modelling is quite popular in the field, it has one major drawback, which is that it is reliant on high-frequency sampling and parameter data being available, such as kinetic rates and absolute initial concentrations. However, much data generated by biologists are not directly amenable to modelling. Furthermore, as there is only very little standardisation of measurements, data from different laboratories usually can merely be compared in a semi-quantitative or qualitative fashion [Kolc 00]. Thus, several alternative methods to estimate missing parameter data in an ODE-based model were developed,

such as Flux Balance Analysis (FBA) [Bona 97] and Metabolic Control Analysis (MCA) [Kacs 73].

Other types of differential equations, e.g. stochastic differential equations (SDEs) and partial differential equations (PDEs) can be used respectively to account for stochastic effects and spatial distribution [Turn 04]. In addition, piecewise-linear differential equations (PLDEs) have been used to integrate discrete and continuous features in gene regulatory networks [Batt 05, Jong 04].

2.3.1.2 P Systems

P systems introduced by Gheorghe Păun [Puau 98] are computational models that performs calculations using a biologically inspired process based on the structure of biological cells in which chemicals interact and cross cell membranes. Variations on P systems have led to the formation of a research branch, membrane computing [Puau 98, Puau 99, Puau 00]. P systems were first employed as a computational model, but later they have been applied to modelling of biological systems [Arde 03, Blak 11, Gheo 08, Paun 05, Rome 09b]. To model a biological system, P systems are defined by a set of membranes, each of which contains biochemical entities and rules. Entities in a P system model determine the process in which one entity may react with another to form a product, while rules may cause biochemical entities to pass through membranes or even cause membranes to dissolve [Puau 98, Puau 99]. Moreover, in a cellular environment, a biochemical reaction may only take place when required molecules collide and interact in a random manner. Therefore rules in a P systems model are implemented randomly, which results in a stochastic computation in the model and multiple simulation results being obtained in a repeated computing process [Puau 98, Puau 99, Puau 00]. Computation in a P systems model stops

at a state which no more reactions are enable. Thus simulating a P systems model is a process of all biochemical entities being passed to outside of the outermost membrane or into a specific membrane [Puau 98, Puau 99, Puau 00, Puau 01]. Although most research in P systems is mainly in the study of the computational power of different proposed variants and its application as a modelling formalism at cellular level, their recent applications in multi-scale modelling have emerged [Rome 09a, Twyc 10].

2.3.1.3 Agent-based modelling

Agent-based modelling (ABM), also known as Individual Based Modelling (IBM), is an approach to model systems comprised interacting autonomous agents. It is a rule-based computational modelling methodology that focuses on rules and interactions among the individual components or the agents of the system [An 09]. It is similar in concept to cellular automata, except that instead of using a grid and synchronised time steps, the agents move freely within the containing space. This approach has been used to study complex phenomena and emergent dynamics using populations of agents with simple rules. At the molecular level, it has been mainly used to build models of signalling pathways that account for spatial distribution and the structural properties of the cell [An 09, Gonz 03, Gonz 06, Pogs 08], but also applied to metabolic reactions [Klan 11].

2.3.1.4 Cellular automata

Cellular automata (CA) were created by von Neumann and Ulam in the 40's [Neum 66]. They are discrete dynamic models that consist of a grid of cells with a finite number of states. A cellular automaton has an initial configuration that changes at each time step through a predefined rule that calculates the state of each cell as a function of the state

of its neighbours at the previous step. They are specially suited for modelling complex phenomena in a scale-free manner and have been used in biological studies for a long time [Erme 93]. Due to their spatial features their main applications are related to molecular dynamics and cellular population dynamics.

Application examples at the molecular level include enzyme reaction networks that account for spatial diffusion [Weim 02] and signalling pathways [Kier 09, Wurt 00]. At the cellular level they were used for models such as those of bacterial aggregation [Sozi 05] and HIV infection [Corn 06, Sant 01]. Dynamic cellular automata are a variation of cellular automata that allows for movement of the cell contents inside the grid, mimicking brownian motion. They were used to model enzyme kinetics, molecular diffusion and genetic circuits [Wish 05].

2.3.1.5 Petri nets

Petri Nets (PN) originating from the dissertation of Carl Adam Petri [Petr 62], are an excellent modelling formalism for describing and studying systems that are characterised as being concurrent, asynchronous, distributed, parallel or nondeterministic. The formalism combines an intuitive graphical notation with a number of advanced analysis techniques with a firm mathematical foundation. PN are applied in practice by industry, academia, and other fields to the analysis of systems arising in asynchronous circuit design, communication protocols, distributed computing, production systems, flexible manufacturing, transportation, systems biology etc. Particularly in systems biology different classes of Petri nets have been used to model a large variety of biological systems, see [Bald 10] for a detailed review.

Petri nets are weighted, directed, bipartite graphs, consisting of places, transitions and

arcs that connect them. Places usually represent passive system components like conditions or resources, while transitions represent active system components like events. In systems biology, places may represent species or any kind of chemical compounds, e.g. genes, proteins or proteins complexes, while transitions represent any kind of chemical reactions, e.g. association, disassociation, translation or transcription. For a chemical reaction, its precursors correspond to the preplaces of a transition while its products to the postplaces of the transition. The arcs lead from places to transitions, or from transitions to places, whose weights indicate the multiplicity of each arc, reflecting e.g. stoichiometries for chemical reactions. The arc weight 1 is usually not labeled explicitly. A place may contain an arbitrary (natural) number of tokens, represented as black dots or a natural number. A distribution of tokens over all places of a Petri net represents a state of it, which is called a marking.

A transition is called enabled if each of its preplaces contains at least the number of tokens specified by the weight of the corresponding arc. An enabled transition may fire and the firing of a transition transfers tokens from its preplaces to its postplaces according to their weights. The firing of a transition updates the current marking to a new reachable one. The repeated firing of transitions establishes the behaviour of a net. The set of markings reachable from the initial marking constitutes the state space of the net. These reachable markings and transitions between them constitute the reachability graph of the net. A simple Petri nets model adapted from a stochastic π-machine model in [Blos 08] modelling the well-known repressilator (a synthetic genetic regulatory network designed to exhibit a stable oscillation) [Elow 00] is illustrated in Figure 2.2.

A unifying Petri nets framework, illustrated in Figure 2.3, has been proposed to relate the three major ways of modelling and analysing biochemical networks: qualitative,

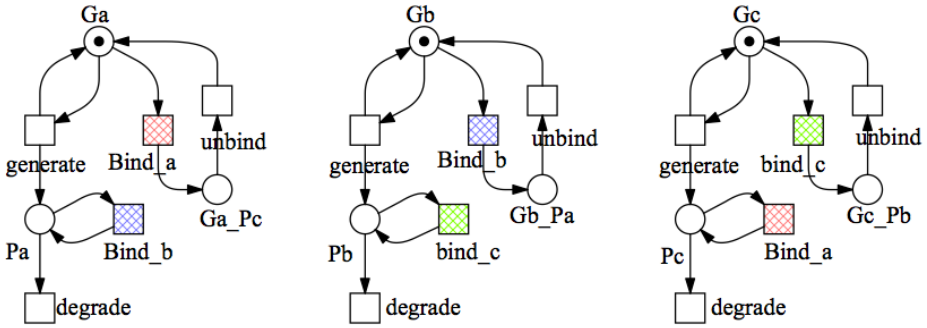

Figure 2.2: A Petri net model for the repressilator. In the model, the highlighted transitions are logic transitions to facilitate the readability of the net, e.g. the two transitions named $bind_a$ are in fact the same transition with two graphic representations, adapted from [Liu 12]

stochastic and continuous [Gilb 07b]. The Petri nets formalism in the framework comprises three classes, sharing nature structure but specialised by their kinetic information, namely qualitative (time-free) place/transition Petri nets (QPN) as well as quantitative (time-dependent) Petri nets: stochastic Petri nets (SPN) and continuous Petri nets (CPN) [Gilb 07b, Rohr 10]. As the qualitative analysis considers all possible behaviour of the system under any timing, the behaviour of QPN forms a discrete state space, which can be analysed in the bounded case, for example, by a branching time temporal logic, one instance of which is Computational Tree Logic (CTL) [Clar 01]. SPN description preserves the discrete state description, but in addition associates a probabilistically distributed firing rate (waiting time) with each reaction, while special behavioural properties can be expressed using e.g. Continuous Stochastic Logic (CSL)[Desh 02]. CPN are represented as a set of ordinary differential equations (ODEs), the state space of which is continuous and linear. It can be analysed by, for example, Linear Temporal Logic with constraints (LTLc) in the manner of [Calz 06]. For more details about Petri nets see [Hein 11] and [Mura 89].

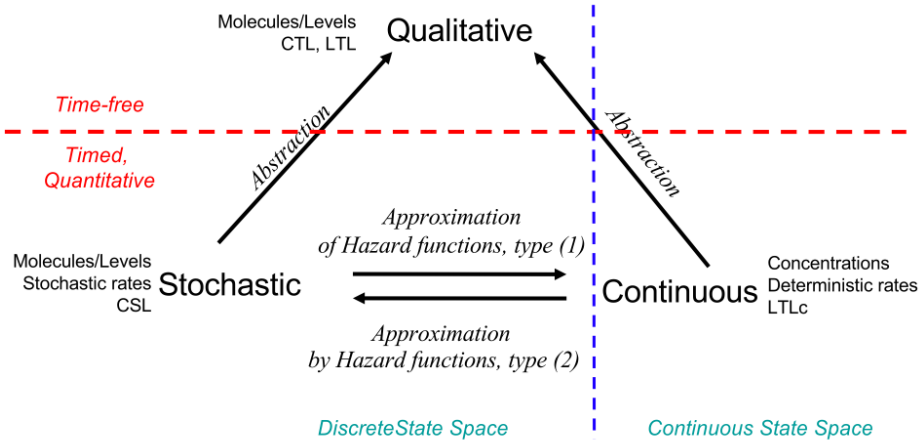

Figure 2.3: A unifying framework for modelling biochemical networks using Petri nets, adapted from [Gilb 07b]

Remarks Several modelling formalisms that are widely used in systems biology have been introduced here, however, other modelling formalisms, such as boolean networks [Kauf 69], bayesian networks [Pear 88], process algebras [Miln 80], cybernetic modelling [Dhur 85, Komp 84], hybrid automata [Ghos 01, Ghos 04], and artificial neural networks [Vohr 01], are also worth mentioning. There is no single formalism which covers the whole spectrum of features desired for modelling all kinds of biological components while Petri nets and rule-based models are among the formalisms that cover most features. Petri nets have several extensions available (e.g. coloured, timed, stochastic, continuous, hybrid, hierarchial and functional), although none of the extensions alone fulfills all requisites, all together they form a very versatile modelling framework. The work in this thesis inclues both QPN and CPN models, while the CPN models are analysed using Probabilistic Linear Temporal Logic with constraints (PLTLc) [Dona 08], which is an extension of LTLc.

2.3.2 Standardisation of models

In the systems biology community, it is widely agreed that models of biological systems should be represented in a common format for exchange and sharing. Currently, Systems Biology Markup Language (SBML) and Cell markup language (CellML) are of the most popular file formats which are widely supported by the community.

2.3.2.1 SBML

SBML is a free and open interchange format for computer models of biological processes [SBML 12]. As a computer-readable file format which can be used to represent models of metabolism, cell signalling, and many other biochemical networks and biological processes. SBML is ultimately proposed to serve as a declarative representation language for computational models in biology. More precisely, the goal of SBML is to serve as a software lingua franca supporting the encoding of models such that those models can be exchanged and interpreted unambiguously by different software systems. It is not intended to encode the details of algorithms used to instantiate the models, nor the procedures used to process and analyse the models. Furthermore, SBML is not linked to any specific software system. It is an XML-like annotation language, in which biological processes and entities involved in biological systems can be described by employing SBML frameworks [Huck 03]. Details about normative definitions and features of SBML can be referred to most recent SBML specification document SBML Level 3 Version 1 Core [Huck 12]. SBML models of biological systems can be found in a repository called BioModels Database (`http://www.ebi.ac.uk/biomodels-main/`).

In addition, a number of software tools are available for SBML-based models [Orto 05, Pett 05], which include CellDesigner (`www.celldesigner.org`), Copasi (`www.`

copasi.org), E-Cell (www.e-cell.org/ecell), Jarnac combined with JDesigner (www.sys-bio.org/software/jdesigner.htm). Additionally, Mathematica (http://www.wolfram.com/) and Matlab (www.mathworks.com) are also compatible with SBML-based models.

2.3.2.2 CellML

CellML is an open standard XML-based markup language, which has been developing by Auckland Bioengineering Institute at the University of Auckland and affiliated research groups. The aim of CellML is to store and exchange computer-based mathematical models, which allows model sharing and enables researchers to reuse components [Mill 10]. With its growth in popularity as a portable description format for computational models, CellML has been used for modelling and developing software tools based on CellML (see CellML case studies at http://www.cellml.org/community/case-studies). CellMl is similar to SBML but provides greater scope for model modularity and reuse, and is not specific to description of biochemistry.

Details about core specifications as well as addition specifications for metadata of CellML can be referred to most recent CellML specification document CellML 1.1 (http://www.cellml.org/specifications/cellml_1.1) and the metadata associated specification documents (the metadata 1.0 http://www.cellml.org/specifications/metadata, simulation metadata http://www.cellml.org/specifications/metadata/simulations, graphing metadata http://www.cellml.org/specifications/metadata/graphs)

In addition, a repository of several hundred biological models encoded into CellML can be found at CellML Model Repository (http://models.cellml.org/cellml).

2.3.3 Simulators

There are various kind of modelling and simulation environment in the systems biology community. Although different software tools define models in different formats to represent biological systems, most of them support importation and exportation of models among several different model formats, for example, SBML-based model files can be imported and exported by Snoopy [Hein 12b, Rohr 10].

Several widely used modelling simulators are listed as follows:

- CellDesigner [Funa 03, Funa 08]: it is a structured diagram editor for gene regulatory and biochemical networks. These networks are drawn based on the process diagram with graphical notation system [Kita 05], which are stored as SBML models. Simulation of these models can be linked to other analysis packages through Systems Biology Workbench (SBW). CellDesigner supports simulation and parameter scanning via an integration with SBML ODE solver and COPASI. Users are allowed to design, construct and modify their models, and simulate and view the dynamics through an intuitive graphical interface.

- COPASI [Hoop 06]: it is a stand-alone software for simulation and analysis of biochemical networks. COPASI is compatible with SBML models, and includes ODEs solver and Gillespie stochastic simulation algorithm. Moreover, arbitrary discrete events can be included in its simulations. Several approaches to analyse the models and its dynamics are provided in COPASI, as well as parameter estimation and optimisation. Details of utilisation of COPASI can be found in [Mend 09b, Mend 09a, Sahl 06].

- BioNessie [Liu 08]: it is a free platform-independent simulation and analysis

environment of biochemical networks, which was developed in Java and could be installed to a number of platforms supporting Java Runtime Environment (JRE) 1.5 or higher. Its Graphical User Interface (GUI) allows users to import, create, edit and export models using SBML standard. BioNessie also provides multiple functions, such as multi-thread/core enabled parameter scans, sensitivity analysis and parameter estimation. By incorporating with National e-Science Centre (NeSC) at Glasgow - BioNessieG benifits from a high performance compute resources across the UK, which enables users to implement large-scale simulations.

- Snoopy [Hein 12b, Rohr 10]: it is a software tool to design and animate hierarchical graphs among other Petri nets. It integrates a unifying framework for graphical display, computational modelling, simulation and bioinformatics annotation of biochemical networks, which can be divided into two levels: uncoloured [Hein 08] and coloured [Liu 12]. A range of Petri nets models are consisted in each level: (i) the uncoloured level contains qualitative (time-free) place/transition Petri nets (QPN) as well as quantitative (time-dependent) Petri nets such as stochastic Petri nets (SPN), continuous Petri nets (CPN), and generalised hybrid Petri nets (GHPN); (ii) the coloured level consists of coloured qualitative Petri nets (QPN^c), coloured stochastic Petri nets (SPN^c), coloured continuous Petri nets (CPN^c), and coloured generalised hybrid Petri nets ($GHPN^c$), illustrated in Figure 2.4. It also includes several ODE solvers for continuous simulations and Gillespie algorithm for stochastic simulations. Snoopy can be downloaded from `http://www-dssz.informatik.tu-cottbus.de/DSSZ/Software/Snoopy`.

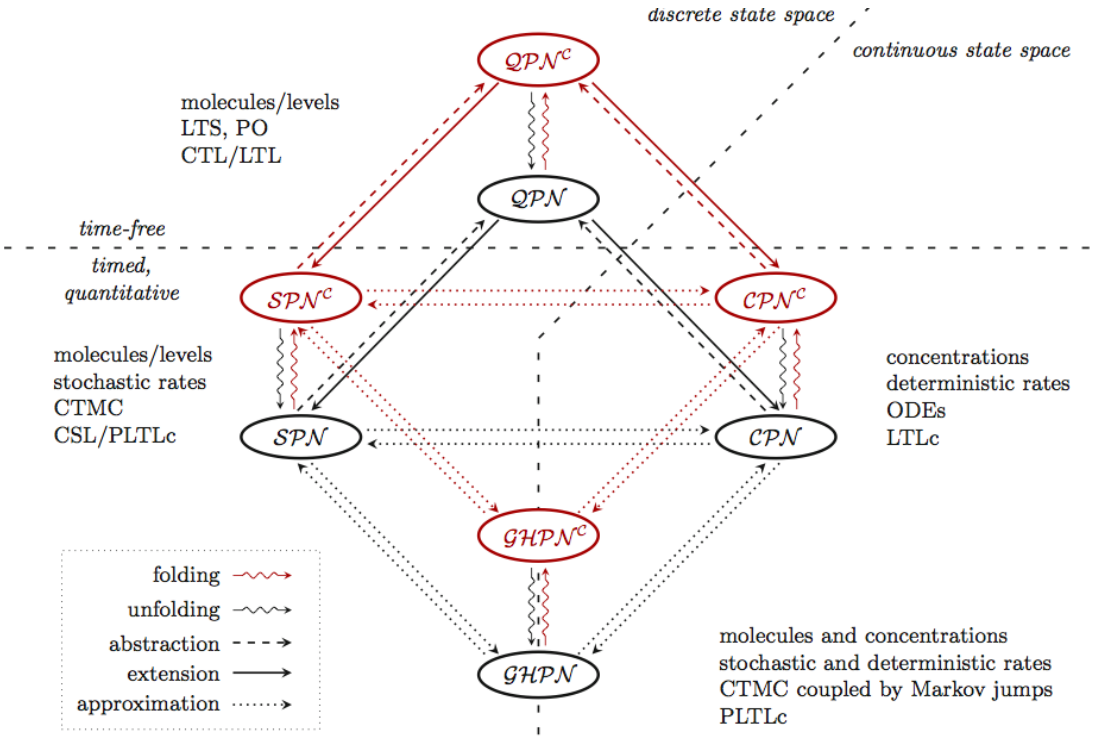

Figure 2.4: Paradigms integrated in Snoopy's unifying framework, adapted from [Hein 12b]

2.4 Multi-scale modelling of biological systems

Until now many models have been at the intracellular level, and indeed have largely ignored locality within the cell, however biological systems are made up of many spatial and temporal scales, each rich and complex. There are different types of scale in biological systems. The organisation of biological systems has been classified into a hierarchy of spatial scales and called this "levels of biological organisation" [Sout 08], which range from gene, to proteins, individual biological cells, tissues, organs, and up to the individual organism that interacts with environment. Associated with this spatially based organisation are the temporal scales of biological processes that range from microsecond (10^{-6} s) for molecular interactions to 80 years (10^{9} s) for the average human life expectancy [Walk 09]. These diverse scales coupled with intra- and inter-scale interactions make a biological system extremely complex, e.g. cell made up of millions of molecules, tissue made up of billions of cells etc.

To understand the behaviour of a biological system requires models that integrate the various interactions that occur on these diverse spatial and temporal scales [Sout 08, Twyc 10] (see Figure 2.5). The complexity of biological systems has made the use of mathematical and computational models to describe and analyse their behaviours and functions an active area of research in recent years. The use of models and experimental data to study how the intra- and inter-scale interactions give rise to their collective behaviours and how they form relationships with their environments is a central theme of systems biology research.

In order to interpret the complexity of biological systems, two major strategies can be employed: 'top-down' or 'bottom-up'. The 'top-down' approach starts with the observation of biological characteristics in the intact biological system and then construct

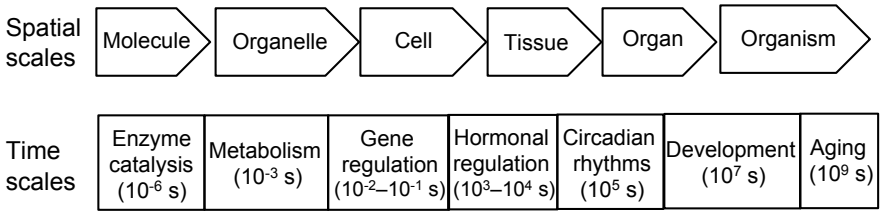

Figure 2.5: Spatial and temporal scales in biological systems [Sout 08, Twyc 10].

theories that would explain the observed behaviours, eventually uncovering the underlying mechanisms. On the other hand, the 'bottom-up' approach begins with the study of the system components in isolation and then integrates the behaviour of each component in order to predict the behaviour of the entire system. Another emerging approach is the 'middle-out' approach, which starts with an intermediate scale (usually the biological cell, the basic unit of life) that is gradually expanded to include both smaller and larger spatial scales [Walk 09].

To fully achieve the objectives of systems biology, models at different scales must be coupled together to produce integrated models across multiple scales. This needs to be carried out with appropriate methods that integrate models representing the different scales involved. The resulting complex integrated models, however, are often computationally expensive and difficult to solve numerically even with the incredible advances in computational power. This is a typical characteristic of multi-scale problems that require the use of suitable multi-scale algorithms, and perhaps with more powerful computers (e.g. distribution computing etc).

In the following, the challenges in multi-scale modelling in systems biology are briefly addressed, and later several multi-scale modelling formalisms, which are playing or will play important roles in such problems in systems biology, are described. More reviews of the existing approaches to multi-scale modelling of physiological and biological systems

can be found in [BASS 05, Dada 11, Schn 07].

Modelling biological systems beyond one spatial scale introduces a series of challenges which should be addressed. These include [Gao 12, Hein 12a]:

1. *Repetition of components* – e.g. the need to describe multiple cells each of which has a similar definition, e.g. epithelial cells in an adult *Drosophila* wing.
2. *Variation of components* – sets of similar components with defined variations. For example, mutants, which widely exist in biological systems, are important to study and understand a biological system, thus in a multi-cellular system, variations among cells are often considered [Kono 71].
3. *Organisation of components* – e.g. how cells are organised into regular or irregular patterns over spatial networks in one, two or three dimensions.
4. *Hierarchical organisation* – enabling the description of (possibly repeated) components which contain repeated sub-components. e.g., cells containing several compartments. This feature enables the use of abstraction regarding the level of detail used to describe components.
5. *Communication between components* – in general communication is constrained to occur between immediate neighbours, but this may be further constrained according to the relationship between neighbours, and the position of a component within a spatial network.
6. *Mobility/Motility* – e.g. transport of components within a system, or actively motile cells.
7. *Replication of components* – e.g. cell division.
8. *Deletion of components* – e.g. cell death.

9. *Irregular/semi-regular organisation of components* – e.g. an approximated honeycomb grid.
10. *Dynamic grid size* – for example alter size and/or topology of grid to model development, may also required for ability to insert/remove components.
11. *Differentiation of components* – e.g. differentiation of embryonic stem cells or immune cells makes a less specialised cell more specialised.
12. *Pattern formation of components* – organising a number of cells in proper one, two or three dimensional structures in space and time.

Consequently these challenges raise crucial demands in the actual process of multi-scale modelling the behaviour of a biological system, which includes seven distinct steps ([Gilb 06, Gilb 10, Hein 12a, Orto 05], see Figure 2.6).

Step 1: System delimitation This step involves selecting the biological system to be modelled, e.g. Planar Cell Polarity (PCP) signalling. More importantly, it is the step to identify the biological question(s) that the model aims to answer, e.g. what is the possible underlying mechanism of PCP signalling?

Step 2: Design and construction This is the key step in the modelling process which involves defining the model to represent the biological system of interest. One of the essential is to draw a detailed topological chart of the system which shows spatial information of the system (e.g. how cells align along the space and if the locality of each cell is static) - only for multi-scale systems, all involved molecular species in each cell (e.g. proteins), what reactions these species are involved and where. It is important to note that most biological systems beyond one scale, if not all, are very complex and not fully understood. Therefore defining such a model often involves making simplifying

assumptions in order to reduce complex and poorly understood systems into simpler ones that can still represent the biological system well enough to explain the observed data.

Step 3: Parameter identification After constructing the structure of the system, the kinetic types for each reaction then need to be defined and the parameter data (e.g. initial concentrations and rate constants) assigned to give a set of detailed kinetic reactions. This step can involve exhaustive searches of the scientific literature to discover what parameter data are already available, as well as performing laboratory experiments to provide data if possible, and often the use of computational techniques to estimate and optimise missing parameters (for example [Webe 13]), whereas the complexity of large-size models makes it even more difficult to search for those missing parameters.

Step 4: Simulation After the model has been defined and quantified, the next step is to translate the kinetic reactions, for example, in an intracellular signalling pathway model(at a single scale), those reactions can be translated into a set of differential equations that describe how the concentration of each species in the system changes over time; this set of differential equations is then simulated over a desired period of time. This is a relatively straightforward step however models with multiple scales raise issues such as: (i) models constructed at different scales sometimes need to be simulated respectively, and then the simulation results at each scale are integrated; (ii) the simulation time is very long due to the complexity of such models, or (iii) it is impossible to simulate such models as there are only a few of multi-scale simulation environments are general frameworks that can be applied to any multi-scale problems in systems biology .

Step 5: Visualisation As multi-scale models always involve different spatial levels (e.g. hierarchical organisation), the visualisation allows the modeller to better comprehend the system behaviour and facilitates the communication between wet lab experimentalists

and dry lab (computational) theoreticians.

Step 6: Validation Simulating a model typically returns a table of data or a curve showing how each species concentration varies over time. These data must then be validated against available experimental data. If the model behaves as the experimental data suggest, then the model can be analysed further. If it does not, then the design and construction step must be revisited, where the model is checked for errors, such as incorrect kinetics or parameter data, over-simplifications of processes and perhaps missing components. This is a debugging loop involving model design, simulation and validation, where the model is refined in order to obtain behaviour which conforms with experimental observations. Sometimes the refinement already reveals useful information about the system. This step is even more intricate for a multi-scale model because: (i) data of biological observations are at a higher level (e.g. tissue level) whereas the simulation results are at the lower level (e.g. biochemical level); (ii) the lack of reliable time-series data for concentrations of key species during to the difficulty of obtaining such data; (iii) biological experimental data are at one time point rather than dynamic (i.e. the final state).

Step 7: Analysis After the model has been validated, it can be analysed and the simulation results are interpreted. Analysis can come in various forms, from the simple examination of the species concentration graphs to more complex statistical analyses. Alternatively model checking can be applied to check the behaviour (simulation results) generated by the model to discover if this model exhibits certain properties. Sensitivity analysis is a commonly used approach that studies the response of system variables to changes in parameter values, and can therefore be used to identify the key reactions and species, as well as monitoring how robust a model is. Robustness is a very important factor in biological systems, as it allows a system to absorb fairly large perturbations

and still function reasonably well; this is because the functionally important behaviour of a system has a certain degree of resilience to damage. If a system variable has a low sensitivity with respect to a parameter, it is robust to alterations in that parameter, however caused. The structural robustness of a model can also be analysed by monitoring how it performs when parts of it are removed; for example, how does the system behave if a species, reaction or entire pathway is removed. This can be useful, because there is often redundancy in biological systems where multiple pathways are available for the production or activation of a certain protein. Overall, the analysis step aims to generate new predictions and hypotheses about biological processes that were not known or unproven before, thus increasing our overall understanding of the system itself.

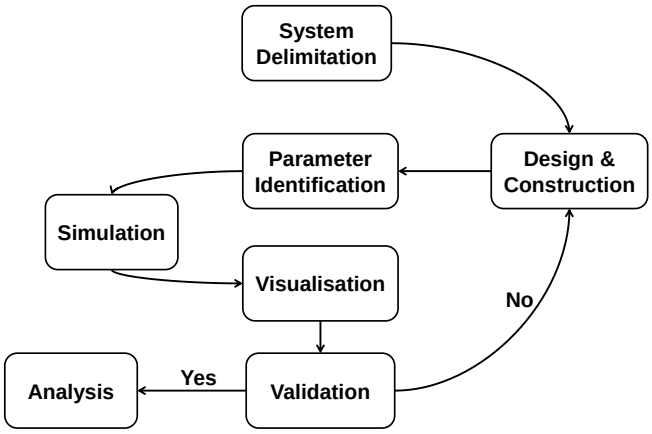

Figure 2.6: The seven steps of multi-scale modelling.

2.4.1 Multi-scale modelling formalisms

A multi-scale model can be defined as a composite model that incorporates more elementary models from several levels of biological organisation or temporal scales

[Sout 08]. With the increasing interest in multi-scale modelling in systems biology, applying proper modelling approaches, which are capable of representing the behaviour of complex systems across a wide range of spatial and temporal scale, becomes critical. Over the past decades, several multi-scale methods have been developed in other fields of science to meet this requirement, which can be adapted to systems biology. As this thesis focuses on spatial scales, thus multi-scale methodologies introduced here are more adaptable to these situations but still could be applicable to temporal scales. As described in Section 2.3.1, P systems can also be applied to multi-scale modelling [Rome 09a, Twyc 10], however, they are not further discussed here.

2.4.1.1 Multi-grid method

Multi-Grid (MG) method developed by Brandt [Bran 77] is a fast and efficient way of solving a wide class of integral and partial differential equations. It requires a series of problems to be solved on a hierarchy of grids with different mesh sizes in order to reduce the problem execution time [Bran 77]. The original MG was later extended by Brandt [Bran 02] taking into consideration different nature of the models at different scales.

2.4.1.2 Complex automata

Complex Automata (CxA) is a general modelling framework developed in [Hoek 08] for multi-scale systems using Cellular Automata (CA) (described in Section 2.3.1), Lattice Boltzmann Models (LBM) and Agent-Based Models (ABM) as building blocks. CxA is based on the ideal that a multi-scale system can be decomposed into N single-scale CA that mutually interact across the scales. A system is first decomposed into its subsystems and a Scale Separation Map (SSM) is built on which each subsystem can be represented as area

according to its spatial and temporal scales as shown in Figure 2.5. The SSM is defined as two dimensional map with the horizontal axis coding for the temporal scales and vertical axis coding for the spatial scales [Hoek 07].

CxA allows the decomposition of the original CA into a number of single-scale CA and let the CAs exchange information in such a way that the overall behaviour of the multi-scale process is reproduced as accurate as possible. Processes having well separated scales are easily identified as components of the multi-scale model [Hoek 08].The interaction across the scales (i.e. exchange of information) is through appropriate coupling mechanisms, which can either be Sub-Domain Coupling (SDC) or Hierarchical-Model Coupling (HMC) method. In the case of SDC, adjacent spatial domains are described by different models on space-time grids of possible different resolution, while some parameters or variables of the main model are first computed locally on the fly by a finer scale model when HMC is used [Hoek 07].

2.4.1.3 Multi-scale Agent-based modelling

Although Agent-Based Modelling (ABM) method has been applied to intra-cellular level (described in Section 2.3.1), its main applications are at the multi-cellular level, where they have been used to study granuloma formation [Sego 04], tumour growth [Enge 08, Zhan 07, Wang 09, Zhan 09], morphogenesis [Gran 06], and several others [Emon 05, Li 08, Loll 06, Mere 07, Thor 8]. These ABM multi-scale approaches were classified into three categories [Walk 09]:

1. cellular-continuum approaches: biological sub-scales are represented in the model as a field of values representing concentrations that are considered to be in steady-state. In this case, single time scale is incorporated in order to represent cellular behaviours.

2. spatially hierarchical approaches which represent sub-cellular components explicitly as a lower hierarchy of agent in the model, but without separation of biological time scales.

3. temporally separated approaches that integrate a second model which represent processes that occur on the a faster time scale.

The major drawback of ABM method is its high computational requirements. The determination of emerging macroscopic system behaviours by simulating the behaviour of all the interacting components can be extremely computational intensive, particularly for higher eukaryotes, where even single organs are composed of large numbers of cells. Efficient techniques for reduction in the computational cost are therefore necessary to help overcome this limitation. Despite this limitation, the important of ABM as multi-scale modelling paradigm in different areas of biological and biomedical problems has been recognised in recent years.

2.4.1.4 Hierarchically coloured Petri nets

Coloured Petri nets Coloured Petri nets (CPN) [Jens 81, Jens 09] are an established discrete event modelling formalism combining the strengths of Petri nets with the expressive power of programming languages. Petri nets provide a sound graphical notation for modelling systems with concurrency, communication and synchronisation. Programming languages offer the constructions of data types, which permit the creation of compact and parameterisable Petri net models. This is the most important advantage of CPN which are going to be exploited in this thesis.

Syntax. CPN comprise – as do standard Petri nets – places, transitions, and arcs. In systems biology, places typically represent species (chemical compounds), while

transitions represent any kind of chemical reactions or transport steps [Hein 08]. In this thesis places represent PCP proteins whereas transitions represent physical interaction and/or signalling events between proteins and polarised transport of proteins.

Additionally, a CPN model is characterised by a set of colour sets (i.e. discrete data types). Each place gets assigned one colour set and may contain distinguishable tokens coloured with a colour of this colour set. As there can be several tokens of the same colour at a given place, the tokens at a place define a multiset over the place's colour set. A specific distribution of coloured tokens at all places constitutes a marking of a CPN. Each arc is assigned an expression; the result type of this expression is a multiset over the colour set of the connected place. Each transition has a guard, which is a Boolean expression, typically over variables occurring in the expressions of adjacent arcs. The guard must be evaluated to true for the enabling of the transition. The trivial guard 'true' is usually not explicitly given.

Behaviour. The variables associated with a transition consist of the variables in the guard of the transition and in the expressions of adjacent arcs. Before the expressions are evaluated, the variables must be assigned values with suitable data types, which is called binding [Jens 09]. A binding of a transition corresponds to a transition instance in the underlying unfolded net.

Enabling and firing of a transition instance are based on the evaluation of its guard and arc expressions. If the guard is evaluated to true and the preplaces have sufficient and appropriately coloured tokens, then the transition instance is enabled and may fire. When a transition instance fires, it removes coloured tokens from its preplaces and adds coloured tokens to its postplaces, i.e. it changes the current marking to a new reachable one. The colours of the tokens that are removed from preplaces and added to postplaces are decided

by the arc expressions. The set of markings reachable from the initial marking constitutes the state space of a given net. These reachable markings and transition instances causing the marking change constitute the possibly infinite reachability graph (state transition system) of the coloured net.

Folding and unfolding. Coloured Petri nets with finite colour sets can be automatically unfolded into uncoloured Petri nets, which then allows the application of all the powerful standard Petri net analysis techniques. Vice versa, uncoloured Petri nets can be folded into coloured Petri nets, if partitions of the place and transition sets are given. These partitions of the uncoloured Petri net define the colour sets of the coloured Petri net. However, the algorithmic identification of suitable partitions is an open research issue. The conversion between uncoloured and coloured Petri nets changes the style of representation, but does not change the structure of the underlying reaction network, see Figure 2.7 for some introductory examples and Chapter 4 for more details.

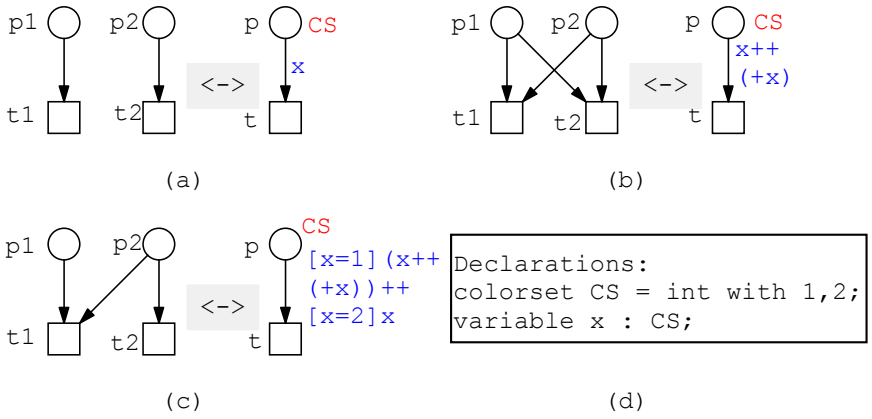

Figure 2.7: Three typical examples for folding-unfolding; notation: ++ multiset addition, (+x) successor, [x=2] guard, adapted from [Gao 12].

The key challenge when unfolding coloured Petri nets is to compute all transition instances. Their computation can be considered as a combinatorial problem, which suffers

from combinatorial explosion. However, when the number of transition instances is only decided by guards, which are in fact logical expressions, Liu [Liu 12] has employed a constraint satisfaction approach [Brai 99, Tsan 93] and an efficient search strategies of Gecode [Geco 11] to greatly improve the unfolding efficiency of coloured Petri nets.

Extensions. Special arcs are specially allowed, e.g. read arcs or inhibitor arcs in the coloured Petri nets. If transitions are additionally associated with stochastic (deterministic) firing rates, as discussed in [Gilb 07a, Hein 08], coloured stochastic (continuous) Petri nets will be obtained. The rate functions defining the usually state-dependent rates can be specified for coloured transitions, or individually for each transition instance; for more details see [Liu 10, Liu 11].

Hierarchically coloured Petri nets [Gao 12, Liu 12]

CPN enable the modelling of (biological) systems comprising repeated components, each of which is associated with a particular colour.

Moreover, colour can be used to encode spatial locality. For example, to model cells in a 2-dimensional lattice, one cell can be represented as a CPN with colour sets denoting the number of copies (cells), and use functions to describe the connectivity between the cells. A colour is a 2-tuple which can then be read as an address identifying locality in space. This can be easily extended to higher dimensions. Moreover, the model is adjustable to different lattice sizes by just changing some constants.

Typically colour sets are not read as unordered sets, but are assumed to be totally ordered, with an implicit successor relation. This idea are generalised, by permitting a partial order relation over a set. One instance of this would result in a hierarchy, i.e. a tree-like structure. Combining the base colour set with the partial order relation, a hierarchical colour set is obtained, and each node in the tree is represented by the path

describing the branch from the root to the node, which enables direct referencing of the position of the node in the hierarchy, and thus permitting the formulation of operations over hierarchy levels.

Operations can be represented over the hierarchically defined colours which correspond to navigation over the hierarchy tree. Operations which move up/down one node are a generalisation of the predecessor/successor function over a totally ordered set. However, operations can also be defined which move over more than one node at a time or jump to nodes in different branches in the tree.

The concept of hierarchy is orthogonal to, and can be freely combined with other colour set constructors such as product or union.

Specifically, the notation for colour sets are described by one, two or three-tuples, corresponding to the number of spatial dimensions being modelled, and note that the underlying colour set is given by the Cartesian product expansion of the colour set tuple. Thus, for example, when modelling cells in, for instance, a rectangular $M \times N$ grid, each cell is associated with a colour (x, y) where $x \in \{1..M\}, y \in \{1..N\}$. A colour set can be associated with a set of constraints which effectively describe the topology of the one, two or three-dimensional grid used to model the layout of the components. Thus a component in a honeycomb (hexagonal) grid may be embedded by imposing the requisite constraints over an underlying rectangular grid. Guards over transitions permit the description of the patterns of connection allowed between cells.

Furthermore in order to describe regular organisation within a cell, this concept can be extent by having a grid at the intra-cellular level and another set of colours to indicate the position inside the inner grid. A separate set of arc functions at the intra-cellular level indicates the inter-component communication at this level. Consequently, a sequence of

tuples are obtained, each tuple referring to the address within a certain level. This concept is called hierarchically structured colours, and the corresponding net class HCPN.

In order to support the description of hierarchically organised systems, the notation for colour sets can be extent as follows. The colour set of the highest level of an L hierarchy is denoted by a tuple T_L, that of the next level by a tuple T_{L-1} and the lowest level by a tuple T_1. When referring to the colour set of a level, its position will be given in the hierarchical tree of colour sets by prefixing the colour sets above it. Thus for example, the colour set for the $L-1$ level is given as $T_L \cdot T_{L-1}$, and the colour set for the lowest level as $T_L \cdot T_{L-1} \cdot \ldots \cdot T_1$. The number of colours in the underlying colour set is given by the product of the number of underlying colours in the colour set tuple from each level.

In order to further facilitate modelling, each level can be denoted by a descriptive label, thus the levels in a three level hierarchy could be denoted by *wing, cell, subcellular-location* and the entire colour set by $wing_{cs} \cdot cell_{cs} \cdot subcellular\text{-}location_{cs}$. The notation can be further expanded for the guards so that they are associated with the level at which they operate. See Chapter 4 for more details on how a HCPN model can be constructed.

In summary, hierarchically structured colours are useful to express repetition and (spatial) locality as what are illustrated in this thesis (see details in Chapter 4). Unlike [Valk 98], HCPN in this thesis does not use a concept of nets-within-nets to describe hierarchy, but colours-within-colours. Two two-tuples (nested two dimensions) are used to address hierarchical systems.

2.4.2 Multi-scale simulators

Most of the available multi-scale simulation environments were developed for specific problems. Only a few of them are general frameworks that can be applied to any multi-scale

problems in systems biology. The most advanced multi-scale simulators come from the heart modelling community. Several packages have been developed for modelling and simulating the heart models by various groups [Nick 05, Usyk 03, Wata 04]. In order to overcome some of problems usually associated with individual groups based simulators, the development of Cancer, Heart and soft Tissue Environment (CHASTE) has been started [Pitt 09]. CHASTE aims at the cardiac electrophsiology and heart modelling and is being developed to code as much modularity and re-use as possible. Among various ABM toolkits, the most popular ones are Repast `http://repast.sourceforge.net/`, Swarm `http://www.swarm.org/index.php/Main_Page`, NetLogo `http://ccl.northwestern.edu/netlogo/`, Mason `http://cs.gmu.edu/~eclab/projects/mason/` and Flame `http://www.flame.ac.uk/`. A comprehensive review of these ABM modelling toolkits can be found in [Grim 06].There are also other multi-scale modelling frameworks, such as Multi-scale Systems Immunology (MSI) [Mith 08], Basic Immune Simulator (BIS)[Folc 07] and Infobiotics Workbench (`http://www.infobiotics.org/`).

Snoopy [Hein 12b, Rohr 10] is a unifying framework and simulator implementing qualitative, continuous and stochastic Petri nets, as well as coloured qualitative, stochastic and continuous Petri nets (see Figure 2.4). Therefore the software tool not only supports traditional modelling and simulation of biochemical networks but also is able to fold and unfold compact and parameterised models, e.g. HCPN models, which gives users the feasibility to construct, edit and simulate multi-scale models using a range of Petri nets methods.

2.5 Summary

With the rapid development of biological technologies and the overwhelming growth of data at various biological levels, there is an urgent need of modelling biological systems with multi-scales (this thesis focuses on spatial scales). The challenges in multi-scale modelling not only come from the complexity of biological systems themselves but also bring issues to modellers to tackle, for example, the choice (perhaps the development as well) of suitable paradigms and associated tools to create coherent descriptions of these biological systems by integrating several spatial scales, and methods for the simulation, analysis and checking of these models (both by human and machine) in order to use them to predict the behaviour of the biological system when disturbed by, for example, mutations, drugs, or stresses. Thus, this thesis uses the application of HCPN modelling formalism to multi-scale modelling of a multi-cellular system, Planar Cell Polarity (PCP) described in Chapter 3, to illustrate how to apply a suitable multi-scale modelling approach to model biological systems with several spatial scales, especially systems organising in a hierarchy, and to search for proper analysis and visualisation techniques to check and verify the models.

Chapter 3

Planar cell polarity

3.1 Introduction

This chapter provides a brief review of planar cell polarity and existing hypotheses regarding its underlying mechanisms. These are used as the biological basis to model PCP signalling using HCPN approach. This is followed by a review of related work in modelling PCP system.

Planar cell polarity (PCP) refers to the coordinated polarisation of cells within the plane of epithelial tissues orthogonal to their apical-basal axes. This critical step is observed in various body regions of the fruit-fly *Drosophila melanogaster*. For instance, the fly eye is comprised of about 800 ommatidia, which are clusters of chirally organised photoreceptors and other cells types [Stru 99, Zhen 95], arranged in a mirror image pattern reflected across the midline (or the equator) toward the dorsal and ventral poles of the eye. Similarly, bristles on the thorax originate from a single sensory organ precursor cell, which polarises in the plane of the epithelium along the anterior-posterior axis to establish the orientation of the bristle [Bell 04, Bell 01a, Bell 01b, Gho 99, Gho 98, Lu 99, Roeg 01a, Roeg 01b].

The adult *Drosophila* wing comprises about 300,000 hexagonal cells [Ma 03] each of

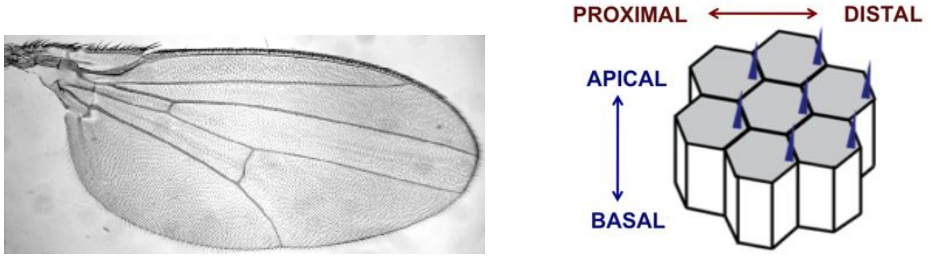

Figure 3.1: *Drosophila*: whole wing (left); scheme of hexagonal cells with hairs (right).

which contains a single hair pointing in an invariant distal direction, see Figure 3.1. This hair comprises actin bundles and is extruded from the membrane at the distal edge of the cell during pupal development, at the conclusion of PCP signalling.

Planar polarity was initially studied using insects such as the milkweed bug *Oncopeltus*, in which gradients of morphogens were hypothesised to direct the local alignment of cell polarity [Lawr 66]. *Drosophila* serves as an ideal experimental platform for studying PCP, because it permits genetic and cell biological analyses, while the hairs and bristles that cover the body surface and the ommatidia of the eye allow for easy assessment of PCP [Tree 02b].

While PCP has been most extensively studied in *Drosophila*, this signalling mechanism is conserved to mammals, and can be seen in the patterning of skin hair, scales in fish, and feathers on birds [Adle 02]. In the organ of Corti in the cochlea of mammalian ears, PCP signalling components have been implicated in the orientation of stereoiliary bundles [Dabd 03]. Defects in the orientation of these stereociliary bundles have been linked to decreased hearing sensitivity or deafness [Comi 89, Dabd 03, Fuji 90, yosh 99]. PCP signalling components also play a role in convergent extension, in which cells elongate, intercalate, and ultimately establish the body axis during gastrulation [Mlod 02, Shih 92, Take 03, Wall 02, Wall 01]. PCP defects can lead to failure of neural tube

closure during development, a condition known as spina bifida, which is a permanently debilitating birth defect among humans. Moreover, PCP signalling components are closely related to the Wnt signalling pathway, which regulates various developmental processes, and the misregulation of which can lead to cancer formation (reviewed in [Cadi 97, Kili 94, Nuss 92]).

The study of PCP is therefore an important step to understand developmental processes and their relationship to intercellular signalling. Indeed, planar polarity serves as a model system for understanding the relationship between cell-level signalling components and their effect on tissue-level patterning and development. These interactions may be quite complex and often may not permit intuitive understanding. Mathematical modelling, especially multi-scale modelling, of such signalling networks can serve to make this link between complex patterning phenotypes and the molecular and genetic interventions that lead to them.

Note that in this thesis, names of genes start with a lower case, whereas names of proteins with a upper case, e.g. *fz* indicates that it is a gene while Fz is used for the corresponding protein.

3.2 Core polarity signalling components

Genetic studies in *Drosophila* have identified a group of core planar polarity genes, so called because they participate in what is called the frizzled pathway, which establishes the polarity of various body regions [Adle 02, Gubb 82, Shul 98, Stru 99]. Key components in this pathway include *frizzled (fz)*, encoding a seven-pass transmembrane protein [Gubb 82, Vins 89], *dishevelled (dsh)*, encoding a cytoplasmic protein [Kili 94, Kras 95, Thes 94], *Van Gogh (Vang)/strabismus (stbm)*, encoding a four-pass transmembrane protein [Bast 03,

Tayl 98, Wolf 98], *starry night (stan)/flamingo (fmi)*, encoding a seven-pass transmembrane protein with cadherin domains [Chae 99, Usui 99], and *prickle-spiny-legs (pk)*, encoding a LIM domain protein [Gubb 82, Gubb 99, Tree 02a]. The *fz* pathway affects ommatidial polarity by regulating the R3/R4 cell fate decision, establishing the chirality and rotation of the ommatidia [Stru 99, Zhen 95]. In the fly wing, the *fz* pathway affects the assembly of prehairs that contain large bundles of actin filaments at the distal vertex of each pupal wing cell by directing the sub-cellular location for prehair initiation [Wong 93].

The core polarity proteins were found to accumulate asymmetrically in cells (see Figure 3.1), forming separate proximal and distal cortical domains in wing cells, or accumulating on the equatorial or polar sides of the R3 and R4 photoreceptor cells in ommatidia, thereby defining their polarity [Adle 02, Adle 01, Axel 01, Bast 03, Eato 03, Fant 04, Feig 01, Jenn 03, Rawl 03, Shim 01, Stru 02c, Stru 01, Stru 02d, Tree 02a, Usui 99, Yang 02]. In the wing, preceding the ultimate manifestation of PCP, signalling occurs such that the proteins adopt an asymmetric localisation within each cell. At the initiation of PCP signalling Fmi, Fz, Dsh, Vang and Pk are all present symmetrically at the cell membrane. At the conclusion of PCP signalling Fmi [Chen 08, Tree 02a] is found at both the proximal and distal cell membrane, Fz [Stru 01] and Dsh [Shim 06, Wong 03] are found exclusively at the distal cell membrane, Vang [Bast 03] and Pk [Tree 02a] are found exclusively at the proximal cell membrane. Through the interpretation of various genetic experiments a consensus view of the signalling events has been formulated that centres on the communication between these proteins at cell boundaries, where the distally localised Fmi, Fz and Dsh recruit Fmi, Vang and Pk to the proximal cell boundary and vice versa. Since the localisation of the distal and proximal proteins appear to be mutually exclusive a completely polarised arrangement of protein localisation results [Axel 01, Feig 01, Shim 01, Stru 01, Usui 99].

The PCP proteins are thus thought to mediate the cell-cell communication that comprises PCP signalling and that they are involved in establishing the molecular asymmetry within and between cells which is subsequently transformed into the polarisation of the wing hairs (reviewed in [Axel 11, Stru 02b]). The result is a polarisation of individual cells and local alignment of polarity between neighbouring cells. This process takes place during the pupal stage of *Drosophila* development, where the core signalling proteins are initially distributed symmetrically within each cell. By 33 hours after puparium formation (APF), shortly before prehairs form, the core polarity proteins have adopted their asymmetric distribution in the cell [Shor 72].

3.3 Overall polarity

Several mechanisms are possible for specifying the overall planar polarity of cells in each body region, cuing the direction in which the core signalling components will accumulate. For example, Lawrence hypothesised a gradient of a morphogen could align the polarity of cells [Lawr 66]. Alternatively, a group of organising cells could initiate a signal that propagates across the tissue. However, a number of experiments argue against the existence of a special group of organising cells in the wing. For instance, *vestigial* mutant wings result in the death of cells on the distal tip of the wing, but the mutation does not disrupt the polarity of rest of the wing [Fris 68, Gubb 82]. Similarly, experiments removing proximal wings cells also do not disrupt polarity [Adle 00, Turn 95]. The temperature-sensitive period for *fz* function was shown to be during the pupal stage, after the loss of the candidate organising cells [Adle 94]. Therefore, any polarity signalling from such organising cells must precede *fz* signalling and act prior to the pupal stage of development.

In the eye, the type II transmembrane protein, Four-jointed (Fj) and the atypical

caherins Dachsous (Ds) and Fat (Ft) provide the *fz* pathway with directional cues [Clar 95, McNe 97, Vill 95, Yang 02, Zeid 99], and also affect polarity in the wing and abdomen [Adle 98, Casa 02, Ma 03, Mata 04, Stru 04, Zeid 00]. In absence of Fj, Ds, or Ft, the proteins in the *fz* pathway remain asymmetrically localized, although not necessarily in the correct direction, suggesting that they act upstream of Fz activity [Axel 01, Ma 03, Stru 01]. Therefore, the long-range expression or activity gradient of Fj, Ds, or Fat may serve to cue the direction of Fz activity in the wing, although it is possible that the direction information may come via Fat from a different source [Ma 03]. Through an unknown mechanism, Ft, which associates with the intracelluar protein Atrophin [Fant 03], biases the direction in which the *fz* feedback loop operates. The PP2A regulatory subunit Widerborst (Wdb), also biases the direction of Fz accumulation, through its relationship to Ft is unknown, and its substrate has not been identified [Hann 02]. However, the function of Ft is essential for directional signals in both the eye and wing, mutations in the genes coding for Ft lead to mutant phenotypes in which cells still coordinate their polarity with neighbouring cells, but fail to align with the tissue axes [Axel 09]. Several mechanisms for how the Fj/Ds/Ft group drive the intracellular asymmetry of the core proteins have been proposed. Some of the first models invoked diffusible morphogens that establish a gradient of ”factor X”, in which the gradient of factor X is translated, by each cell, into a scalar value of Fz, and that each cell can compare this value with those of their neighbours and orient their polarity accordingly [Casa 02, Lawr 04]. Being unable to identify a morphorgen congruous with the properties of the factor X, this model has largely given way to signalling models relying on local interaction [Axel 09]. Following the discovery of the heterotypical complex formed by Ds and Ft, it was suggested that the graded expression of Fj and Ds is converted into intracellular asymmetries of these DsFt heterodimers [Ma 03, Yang 02]. These subcellular

gradients would then transmit directional information to downstream components of PCP pathway, although a mechanism for this interaction has not yet been identified [Wu 09]. Another proposed model for how Ds and Ft could orient cell polarity is the polarised vesicle transport hypothesis. Vesicles containing Fz and Fmi, tagged by green fluorescent protein, and Dsh have been observed to be transported preferentially towards the distal cell cortex on a polarised microtubule cytoskeleton during the build-up of intracellular asymmetry [Shim 06]. However it has not settled whether PCP is established via on pathway or by the two systems acting in parallel. Nonetheless, it is clear that the mechanism for PCP employs partially redundant information and that this partial redundancy increases the robustness of PCP signalling and reduces polarity errors.

3.4 Three-tiered model

PCP signalling can be viewed as three steps, in which

1. a global directional signal provides a cue to the *fz* pathway, which in turn
2. locally propagates the signal and converts it into the asymmetric accumulation of PCP signalling components, and
3. directs the tissue specific readout of the cell polarity, for example, making the site of prehair assembly or directing the R3/R4 cell fate decision in the ommatidia [Tree 02b].

The focus of this work was on the second tier and how the *fz* pathway receives the asymmetry cure, mainly focusing on factor X and polarised vesicle transport, from the first tier.

3.5 Hypotheses of domineering nonautonomy

An interesting property of the establishment of planar polarity was observed in experiments in which clones of cells mutant for *fz* were observed to not only disrupt polarity within the mutant clone, but also to repolarise some of the surrounding genetically wild-type cells [Gubb 82, Vins 87]. This long-range disruption of hair polarity is referred to as domineering nonautonomy (Figure 3.2). Later experiments overexpressing *fz* in a clone revealed that the hair polarity tended to point away from regions of high Fz activity [Alde 97], whereas the hair polarity tended to point toward clones of cells with null *fz* alleles [Gubb 82, Vins 87]. Vang clones show the opposite effect, with hairs pointing away from cell lacking Vang activity and pointing toward cells which overexpress *Vang* [Tayl 98]. However no polarity occurs in both *fmi* clones and neighbouring wild-type cells if lacking Fmi activity, indicating that Fmi homodimers are necessary for communication of polarity information between cells [Usui 99]. Clones of cells mutant for other core polarity genes do not all exhibit such long-range effects, exhibiting cell autonomous behaviours (Figure 3.3). For example, *dsh* clones show only a cell autonomous polarity phenotype [Kili 94, Thes 94], and *pk* clones show only a subtle phenotype even inside of the clone [Amon 05]. In the eye, domineering nonautonomy near mutant clones that disrupt Wingless (Wg) signalling, such as arrow mutant cells, is seen on the equatorial side of the clone [Wehr 98], while clones disrupting *fz* produce domineering nonautonomy on the polar side of clones [Zhen 95]. Interestingly, the overexpression of either *dsh* and *pk* also show nonautonomous phenotypes [Bast 03]. It is not clear how cells mutant for some polarity genes can repolarise neighbouring cells [Adle 02]. Three differing models have emerged to explain this behaviour though their relationship is still controversial.

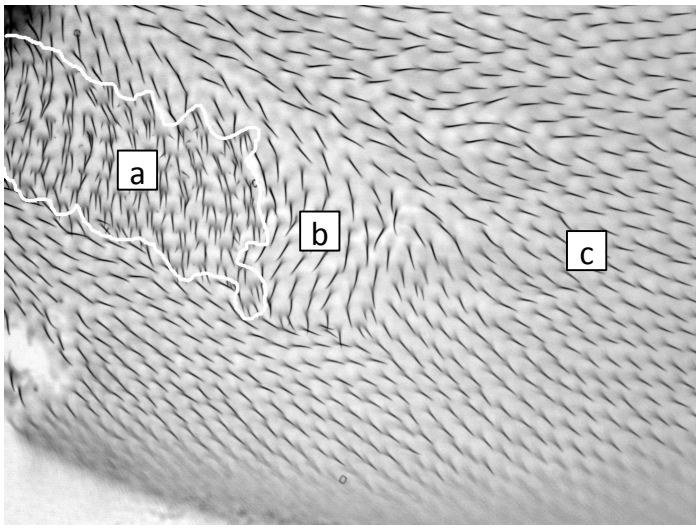

Figure 3.2: Image of domineering nonautonomy obtained from Dr Tree's lab. (a) *fz*- clone, (b) Domineering nonautonomy distal to the clone, (c) wild-type cells far from the clone.

3.5.1 Factor X or Z models

To explain how cells mutant for PCP genes can affect the polarity of neighbouring wild-type cells, a class of models has been proposed in which cells respond to an initial cue by producing and secreting a diffusible second factor whose graded distribution determines polarity [Adle 00, Fant 03, Lawr 99, Lawr 02, Stru 97a, Stru 02a, Wehr 98, Zeid 99]. In the eye, Wingless(Wg) signals in a gradient that is highest at the poles and lowest at the equator. It was hypothesized that Wg signalling induces the dose dependent secretion of "factor X", which diffuses, and in turn regulates Fz signalling [Wehr 98]. Clonal disruption of the response to Wg would perturb the factor X gradient in both the mutant and the neighbouring wild-type tissue, thereby producing nonautonomy. Similarly, in response to a graded upstream signal, Fz and Vang has been hypothesised to regulate production of diffusible "factor Z" that is required for polarity readout [Adle 00, Adle 02]. Interestingly, while for most *fz* alleles, including null alleles, clones of mutant cells produce domineering

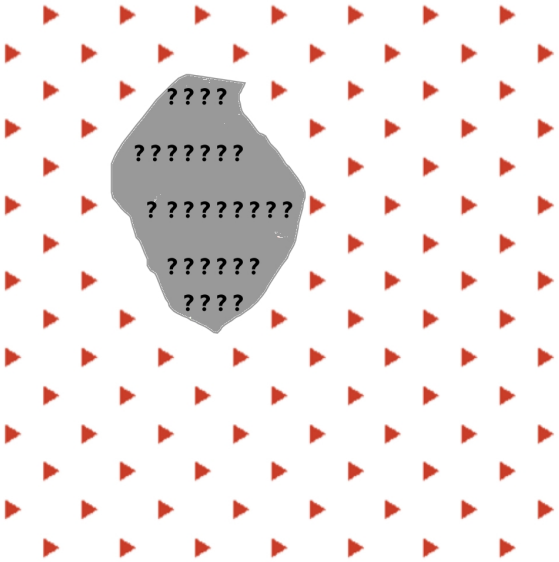

Figure 3.3: Diagram of autonomous clone. Grey region indicates mutant clones.

nonautonomy, clones of cells mutant for some *fz* alleles produce an almost cell autonomous polarity phenotype [Gubb 82, Jone 96, Vins 87]. Clones of cells mutant for *dsh* also produce nearly cell autonomous polarity disruptions [Kili 94, Thes 94]. Diffusible factor models explain these observations with the hypothesis that Fz mediates two separately mutable signalling functions, transducing a cell autonomous signal through Dsh, and a Dsh independent nonautonomous signal [Chae 99, Kras 94, Stru 02d, Tayl 98]. The nonautonomous signal was proposed to be mediated by factor X, or a second, similar, factor Z, in this case produced in response to Fz signalling, and feeding back through Fz, while the autonomous signal was proposed to initiate cell polarisation [Adle 00, Adle 02]. Nonautonomous Fz signalling has been proposed to temporally precede autonomous signalling in both the eye and wing [Stru 02d]. However, despite the ability of these models to explain domineering nonautonomy, diffusible factors X or Z have not been identified [Lawr 02], and no molecular level of understanding of the autonomous and nonautonomous signalling functions of Fz has been obtained [Adle 02, Stru 02d].

3.5.2 Polarised vesicle transport model

It has been observed that Fz, Fmi and Dsh are transported preferentially towards the distal cell cortex via vesicles on a polarised microtubule cytoskeleton during the initiation of intracellular asymmetry [Shim 06]. Moreover, Ds and Ft control alignment of cortical microtubules along the proximal-distal axis and asymmetry of microtubule growth [Haru 10]. The hypothesis that Ds and Ft can orient active transport of core proteins by controlling the polarity of the microtubule cytoskeleton has been supported by data [Olgu 10] as well. The most persuasive piece of evidence comes from studies of the fly abdomen, where excess Ds or Ft in cells within a clone can repolarise surrounding cells receiving this polarity information, even when the cells in the clone and in the surrounding tissue lack Fz or Fmi or both [Casa 06, Lawr 07], which indicates that PCP can be established by the Ft/Ds/Fj system in the absence of the function of core proteins. Further more, studies of the wing have shown that not all cells need to respond correctly to the Ft polarity signal, as subsequent action of the feedback loop, comprising the core proteins, is sufficient to align cells [Ma 03].

3.5.3 Feedback-loop model

Asymmetric cortical recruitment of core PCP complexes also relies on intercellular feedback at the interface of neighbouring cells [Axel 01, Tree 02a]. This is mediated by an asymmetric, homotypic interaction between Fmi bound to Fz in one cell and Fmi associated with Vang in the other [Bast 03, Chen 08, Das 04, Lawr 04, Usui 99, Shim 06]. Mutual inhibition between Fz/Dsh and Vang/Pk complexes is proposed to drive the system to an asymmetric distribution [Axel 01, Tree 02a] (reviewed in [Zall 07]), but this is not yet known in molecular terms. Feedback may depend on molecular

interactions between proximal and distal components, as Vang and Pk both bind Dsh and inhibits its membrane recruitment and function [Bast 03, Tree 02a]. Diego (Dgo) also binds Dsh, which promotes its activity and blocks the antagonistic effect of Vang/Pk [Das 04, Feig 01, Jenn 03, Jenn 05]. Feedback loop may also involve regulated vesicle trafficking [Shim 06, Stru 08]. The assembly of asymmetric intercellular complexes both communicates polarity information between cells and helps define and stabilise the asymmetric cortical domains within cells.

This feedback model was proposed by Amonlirdviman et al. [Amon 05] that experimentally identified local feedback loop and global directional cue comprise the central components of a PCP signalling network, and that examination of their properties may yield testable insights into the mechanisms underlying domineering nonautonomy. However, intuition alone is insufficient to deduce the behaviour of this regulatory network. Indeed, progress in understanding PCP signalling has been severely hampered by an inability to deduce, given a particular signalling network hypothesis, definitive links between molecular genetic interventions and tissue level patterning effects. Thus, it is not readily apparent that this network is sufficient to explain the complex patterns observed in fields of cells containing mutant clones, and others have argued that it cannot account for some of the observed phenotypes [Adle 02, Stru 02d]. For example, it is apparent that removing Dsh and Fz would disrupt the feedback loop, whereas, it is not obvious how the feedback loop in adjacent wild-type cells responds, such that *dsh* mutant clones behave autonomously, while *fz* clones behave nonautonomously. Furthermore, *pk* overexpression promotes the asymmetric accumulation of Dsh and Fz [Tree 02a], while the role of Pk in the feedback loop is to inhibit the membrane recruitment of Dsh. Thus, if the model is correct, then this amplification must be a non-obvious consequence of the feedback loop system

dynamics. Finally, the design and interpretation of additional experiments will require a clear prediction of expected outcomes.

Remarks Although most investigators now favour the model that the core polarity signalling components may function in a feedback loop that acts to both amplify the global directional cue (factor X) and replay the signal to neighbouring cells, several puzzles remain. This thesis investigates different hypotheses derived from experiments in order to shed the light on the mechanisms underlying PCP signalling.

3.6 Mathematical modelling of PCP

Several mathematical and computational approaches have been applied to study PCP signalling. In order to understand how the core proteins interact to produce domineering non-autonomy in *Drosophila* wing, Amonlirdviman et al. [Amon 05] (extended in [Raff 08]) built a model by applying Partial Differential Equations (PDEs) and reaction-diffusion equations which abstract from the spatial dimensions of the PDE model by discretising each cell into a triangular mesh and used periodic boundary conditions to select the grid of cells. Agent-based Modelling (ABM) with stochastic differential equations was used to create a computational model which includes the mechanism in which a Frizzled gradient occurs through feedback-reinforced formation of Flamingo-based asymmetric intercellular complexes in Le Garrec's research [Le G 06] (applied to the *Drosophila* eye in [Le G 08]). Farhadifar et al. [Farh 07] investigated the physical basis of epithelial cell packing in *Drosophila* wing using a vertex model to establish the network of cells in which vertices are subject to mechanical forces. Burak et al. [Bura 09] constructed a semi-phenomenological representation model involving stochastic equations and used statistical mechanics to study how local interactions between

cells impact the dynamics of the process on parameters. Schamberg et al. [Scha 10] built two models by applying reaction-diffusion equations to study the influence of feedback loop, intra-inter-cellular diffusion and whether the re-distribution of proteins depends on the amount of proteins in neighbouring cells during polarisation, which was based on *one-dimensional* line of cells. See [Axel 11] for an overall review of this field.

3.6.1 Amonlirdviman model

Amonlidirdviman et al [Amon 05] developed a mathematical model in order to understand how protein localisation within and between cells, and local feedback loop could drive planar cell polarity in the *Drosophila* wing. Their model centred on the system of core proteins involved in PCP signalling, and a feedback loop between adjacent cells which amplifies an initial asymmetry of protein distributions to later produce huge asymmetric distributions of proteins and complexes as seen in laboratory experiments. This model was constructed by using PDEs based on diffusion and reaction kinetics. It described the concentration of core proteins (Fz, Dsh, Vang and Pk) and their interactions throughout an array of cells, as well as introduced a global directional cue to insert a directional bias in the feedback loop. This PDE-based model was mathematically simulated by discretising to ODEs on a two-dimensional grid of hexagonal cells, which was able to reproduce the characteristic PCP mutant phenotypes, including the effect of domineering nonautonomy in the *Drosophila* wing. Although this model was able to illustrate that a local feedback loop combined with a global directional bias could possibly explain some aspects of the behaviour of PCP mutant clones, the model itself had limitations, for example, it only employed a limited number of proteins of PCP signalling, and parameters were restricted to fit a set of feature constrains such that those simulations would reproduce qualitative

features of the experimentally observed PCP phenotypes (see [Amon 05] supplementary material). Moreover, the effects of stochasticity or geometric variation are not included in the model. Thus a more detailed implementation of this model has been carried out. In contrast to the previous model, the new model was able to correctly reproduce the characteristic swirling pattern created by mutant clones lacking ft [Ma 03, Ma 08, Raff 08].

3.6.2 Le Garrec model

The model developed by Le Garrec et al [Le G 06] focused on the local feedback loop between core proteins and implemented diffusion and reaction kinetics for proteins and complexes that they form, which was similar to the Amonlidirdviman model [Amon 05]. In this model, polarisation was conditional upon three molecular processes: (i) a transient long-range signal over the whole wing tissue in order to establish a small asymmetry in Fz activity; (ii) the ability of adjacent cells to communicate and compare their concentrations of the intermediate complexes Fz:Fmi and Vang:Fmi, and (iii) the presence of an intracellular feedback loop for the amplification of the initial symmetry [Le G 06]. This model is computationally different to the Amonlidirdviman model [Amon 05], because it was built by using Agent-based modelling (ABM), in which PCP in-silico cells were represented by a connected collection of pixels, membranes by closed loop of pixels and cell motion was implemented by exchange of pixels between cells according to a statistical tendency of cells to maintain an optimal size. The cell boundaries were therefore fluid, nonetheless, remain approximately hexagonal, and the dynamics of all processes were noisy and described by stochastic differential equations. The model was also able to reproduce characteristic PCP mutant phenotypes through parameter fitting.

3.6.3 Farhadifar model

Farhadifar et al. [Farh 07] constructed a vertex model of epithelial cells as to investigate physical basis of cell packing in the *Drosophila* wing. The vertices in this model were subject to mechanical forces, such as forces due to the mechanical properties of cell cytoskeleton (contractility) and cell membrane (elasticity), the adhesive properties of cell-cell communications, and the tendency of cells to maintain an approximately constant volume and area. As stationary and stable cell network configurations satisfy a mechanical force balance, cell packing geometries were described by local minima of an energy function in the model, which included terms representing cell area elasticity, line tensions at junctions between cells (the effect of cell-cell adhesion and actin-myosin contractility), and the cortical actin contractility which acts along the cell perimeter. This energy function may have various local minima as well as global minimum depending on parameter values. In addition, several dynamic processes were allowed to affect the cell packing geometry, for instance, cell division, neighbour exchange, cell death (removing from the network). Farhadifar et al. probed the mechanical properties of junctions using laser ablation techniques and in analogy with [Clas 05], used the distribution of polygons in the tissue to characterise the packing geometry. They identified a region in parameter space where simulations matched observed tissue properties by analysing the ground state and equilibrium topologies of the model and comparing their numerical results to laboratorial experiments. They concluded that hexagonal lattice is the global minimum for epithelial cells with physical properties corresponding to this parameter region, and disordered packing geometry caused by stochastic cell division represented a local minimum. Furthermore, line tension on individual cell boundaries was positive, indicating that contractility predominates over adhesive cell-cell interactions. Later,

Mehonic [Meho 11] investigated the model further by introducing noise to the system.

3.6.4 Burak and Shraiman model

In the Burak and Shraiman model [Bura 09], they investigated the role of stochastic dynamics in PCP signalling by developing a semi-phenomenological model based on three biological assumptions essential for reproducing the pattern of protein localisation and the nonautonomy of mutants lacking Fz and Vang. These assumptions are: (i) transmembrane proteins Fz and Vang form complexes across intercellular cell gap which is the source of intercellular interaction; (ii) complex formation on a single intercellular cell gap is bistable such that the formation of polarised complexes inhibits the formation of opposite polarised complexes which leads to segregation of the complexes; (iii) this inhibition acts non-locally within each cell which leads to special segregation of complexes on opposite sides. This model does not place emphasised on the molecular detail and the number of parameters, thus it is significantly smaller than any other models mentioned above. It was used to study the response of the tissue to various types of orienting signals and discovered that global alignment of polarity could be produced by a weak graded signal across the tissue, in contrast, a boundary signal cannot effectively propagate polarity information over a large number of cells. Furthermore, the model was used to illustrate that in the absence of an external orienting signal, swirling patterns would emerge due to random noise in qualitative agreement with the phenotype observed from mutants lacking of *ft* [Ma 03, Ma 08].

In addition, Schamberg et al. [Scha 10] built two reaction-diffusion equations models based on *one-dimensional* line of cells to study the influence of feedback loop, intra-inter-cellular diffusion, and if the re-distribution of proteins depends on the amount of proteins in neighbouring cells during polarisation.

Remarks As discussed above, these computational models have their own advantages and deficiencies. They were designed and constructed to study particular aspects of PCP signalling. However, their common drawback is that these models lack of an approachable way to generate the cell geometry or a grid of hexagonal cells, and are hard to reproduce by other researchers. Thus, it will be a significant contribution to provide an approach which allows the systematical construction of large scale mathematically tractable models in which cell geometry is clearly formalised. Compared with ODEs and PDEs, HCPN are more intuitive for users who do not have much knowledge about modelling. In addition, new aspects of signalling will be incorporated into computational models, such as the inclusion of polarised vesicle transport.

3.7 Summary

This chapter has summarised the current understanding of PCP signalling, which provides the basis of modelling, and then described as well as discussed the computational models developed to study this biological system, while ended with a brief discussion regarding the weakness of existing computational models of PCP signalling.

As PCP system is assumed to be treated as a continuous multi-cellular system with multiple spatial scales (see further discussion in Chapter 4), the choice of a suitable multi-scale modelling formalism does not only depend on the validity and feasibility of the approach but also partly up to the computational tractability of an appropriate simulator to model and simulate such a system [1]. Thus HCPN were chosen because: (i) it can

[1]The work presented in this thesis started from 2010. By then CPN integrated in Snoopy was quite promising to be used as multi-scale modelling and simulation environment. In addition, this work has also driven the development and improvement of Snoopy as the urgent need to construct and simulate large-scale models.

be applied to construct compact, parameterisable and scalable models, (ii) the modelling and simulation tool, Snoopy [Hein 12b, Rohr 10], can cope qualitative, stochastic and continuous PN and CPN models.

In the next chapter, a family of HCPN models are built based on the core machinery of PCP signalling and hypotheses, e.g. feedback loop, factor X and polarised vesicle transport described in this chapter. In addition, new aspects of signalling, such as polarised transport of molecules which has not been considered in previous models, are to be tested using the HCPN models presented in Chapter 4.

Chapter 4

Multi-scale modelling of PCP

4.1 Introduction

This chapter describes the initial attempt to mathematically model PCP, followed by a full development of Hierarchically Coloured Petri Nets (HCPN) template, the ODE models, and the family of HCPN models and their variants. Then, the primary exploration of parameter optimisation using simulated annealing with a description of the implementation details is given. It concludes with a discussion of the design and construction of the HCPN models, and the computational efforts made to accelerate the simulation. Some models presented here were published in [Gao 11, Gao 12].

The primary goal was to develop a mathematical model of PCP signalling that would recapitulate the phenotypes of all known mutant conditions, both loss and gain of function. It is hoped that this model would be able to address questions about the mechanisms underlying the polarising signal, the dynamics of signalling by the individual components and the interactions between the PCP proteins which orchestrate the ultimate morphological manifestation of planar polarity. Ultimately, this model would make predictions about the mechanisms underlying PCP signalling process which would be

testable in a biological laboratory.

The core elements proposed in [Axel 01, Le G 06, Tree 02a] are included in the model in order to illustrate that the model could explain both autonomous and nonautonomous clonal phenotypes. An initial ODE model of PCP signalling was built using a set of ODEs. This model included the production of the PCP proteins in the nucleus, the uncharacterised transport of these proteins to the proximal and the distal membranes, and the core machinery regulating the PCP. Recruitment of Fmi to the membrane and binding of Fmi from two adjacent cells activate the signalling, which was explicitly modelled as a binding (complexation) reaction. Recruitment of Fz, Dsh, Vang, Pk from the interior to and from either the proximal or distal membrane was controlled by the formation of the dimer of the transmembrane protein Fmi from adjacent cells. This initial ODE models of the core machinery of PCP as described above and its variants (by incorporating different hypotheses to the initial model) were used as the basis of creating a family of HCPN models of PCP signalling in order to study and understand the underlying mechanism of this signalling pathway.

The work started off by the construction of an asymmetric model representing the PCP system using Coloured Petri Nets (CPN) [Gao 11] based on a limited inter-cellular communication schema and a crucial way of introducing biasing mechanisms. In this model, each cell was divided into virtual compartments, three each at the proximal and distal sides, and one large central block. This configuration introduced an inherent proximal-distal topological bias across the in-silico wing tissue. In addition, the asymmetric model included three specific intracellular biasing mechanisms in terms of feedback loop, morphogen factor X and polarised vesicle transport (see biological details in Chapter 3). Moreover, only two neighbourhood functions were defined to enable

inter-cellular communication between adjacent cells, forbidding a cell from interacting with its north and south adjacent cells.

The aim was then to construct models that would eliminate the tissue-level bias and permit the exploration of different biasing mechanisms controlling the intra-cellular biasing mechanisms proposed in the field. Such models could present the feasibility of the Hierarchically Coloured Petri Nets (HCPN) model to explain the characteristics PCP phenotypes subject to reasonable assumptions of the properties of the physical mechanisms for the interactions.

4.2 Establishing the hierarchically coloured Petri nets template

Initially each cell was modelled as a (standard) Petri net at a highly abstract level, neglecting details of the intra-cellular pathways, and divided several virtual compartments to facilitate the description of PCP. Because *Drosophila* wing cells form a regular honeycomb lattice, this organisation at the top level of the model has been thus imposed as a hexagonal grid of cells, see Figure 4.2. The Hierarchially Coloured Petri Nets (HCPN) model template was developed by step-wise colouring this spatial information (the construction of the abstract model). The next step was to create a family of HCPN models by adding details of the intra- and inter- signalling mechanisms, and reusing the colouring template of this abstract model.

4.2.1 Coloured Petri nets template

The first model of the wing epithelial cell, illustrated in Figure 4.1, was a high-level representation of the PCP system in order to establish the colour sets required.

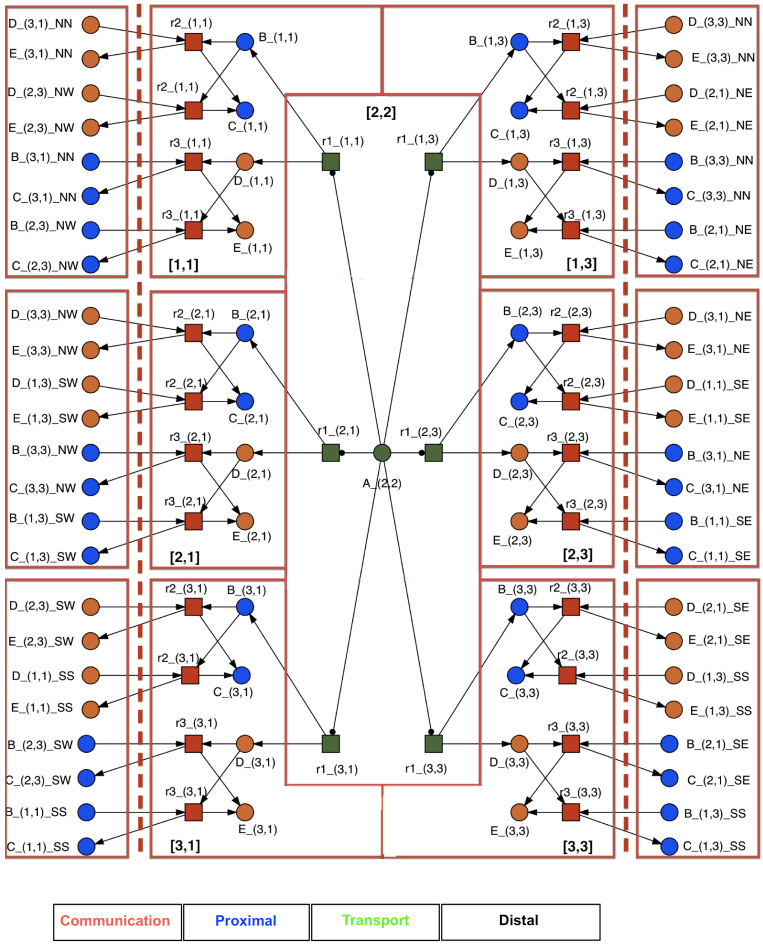

Figure 4.1: Abstract Petri net model for a single cell: (i) one-to-two neighbourhood relationship; (ii) four spatial regions: communication, proximal, transport and distal; and (iii) seven virtual compartments. Dashed red lines indicate cell boundaries [Gao 12].

Each cell was sub-divided into four spatial regions (Figure 4.1, from left to right): (i) the extracellular space (labelled as communication), where the intercellular complexes form, (ii) the proximal cell margin (left-hand side of each cell) in order to process intercellular

signals between two neighbouring cells, (iii) production (read arcs cause an infinite supply of proteins) and intracellular transport of core proteins, and (iv) the distal cell margin (right-hand side of each cell).

In order to facilitate the detection of PCP asymmetry, each cell was then partitioned into seven *virtual compartments*, three each for the proximal and the distal membrane arranged in a circular manner, and one middle compartment for the cytosol, see Figure 4.2, right. This circular structure was imposed on an underlying 3×3 rectangular grid, which explained the numbering of the compartments. Different neighbourhood relationships between adjacent cells were created in order to investigate their influence on the overall behaviour. The two neighbourhood relationships are (i) *one-to-two*, where one individual compartment in a cell communicates with two other compartments in the neighbouring cell(s), see Figure 4.2 (b); (ii) *one-to-one*, where each compartment in a cell communicates with only one compartment in the neighbouring cell, see Figure 4.2 (c).

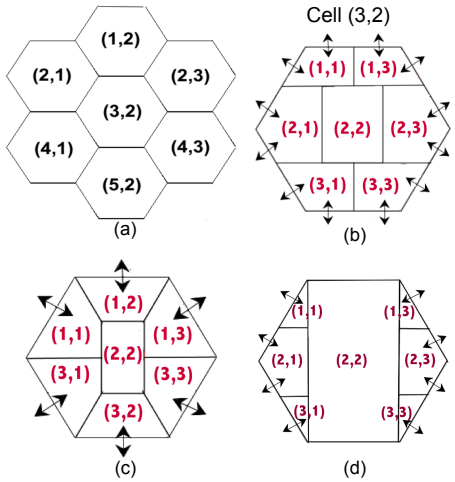

Figure 4.2: *Drosophila* wing epithelial cells. (a) Fragment of wing tissue; coordinates represent honeycomb grid position. (b) Cell with seven virtual compartments in a circular manner in **one-to-two** relationship, arrows denote inter-cellular communication with adjacent neighbouring cells [Gao 12]. (c) Likewise, for **one-to-one** relationship in a circular manner. (d) Cell asymmetrically divided into seven virtual compartments [Gao 11].

Both of these relationships require the definition of six neighbourhood functions (north, north-east, south-east, south, south-west and north-west, abbreviated as NN, NE, SE, SS, SW, NW respectively), see Figure 4.1. Because a honeycomb is a tessellated structure, the pairwise complementary symmetries between neighbouring cells are ignored, thus reducing the need for the 6 rotational axes of symmetry to 3 axes (NN, SW, NW), see Table 4.1.

4.2.2 Hierarchical coloured Petri nets template

This section presents the construction of a HCPN model for a one-to-two neighbourhood relationship and later illustrates how simple it is to adapt this approach to generate a HCPN model for a one-to-one relationship. For the declarations see Table 4.1.

Two constants M, N and a two-dimensional colour set ($CS1$) representing a rectangular $M \times N$ grid were defined, and the subset denoting the coordinates of the hexagonally packed cells (CS_Cell) was selected (Figure 4.2). At this level of hierarchy (wing tissue comprising folded cells) a HCPN model was obtained, which had a similar structure to that of Figure 4.1, but each place had been assigned the colour set CS_Cell.

Next a colour was assigned to each of the seven virtual compartments of a cell. This was done by using a 3×3 grid (CS_ComP) and ignoring colours (1,2) and (3,2) so that the proximal compartments were (1,1), (2,1) and (3,1), the middle compartment was (2,2), and the distal compartments were (1,3), (2,3) and (3,3). The information about cell and compartment locality were combined by defining $CS2$ as product of the colour set to address cells (CS_Cell) and the colour set to address the virtual compartments (CS_ComP). Two subsets $CSmembrane$, and $CSmiddle$ of $CS2$ were introduced to facilitate addressing the components of a cell in a specific region, (i.e. distal or proximal, and middle). The colour defined were hierarchical, thus, each place could be located in

terms of the coordinates $((x, y), (a, b))$, where (x, y) denoted the position of a cell in the honeycomb grid, and (a, b) denoted the position of a virtual compartment within that cell.

The folding was continued by using these colours to obtain a more compact HCPN model – a tissue of cells comprising virtual compartments. This was achieved by folding the six membrane compartments into one, by assigning the colour set $CSmembrane$ to its places. The central compartment $(2, 2)$ was denoted by the colour set $CSmiddle$.

Finally, the two similar communication components (the transitions given in red in Figure 4.1) were folded in each membrane compartment into one. For this a colour set CSI of two colors and a product colour set $CS3$ based on $CSmembrane$ and CSI were defined.

The following describes the necessary steps to construct the compact HCPN model. Having defined the colour sets, variables that were used in transition guards and arc expressions were created. All the transitions in the six membrane compartments were assigned guard $Gd1$ that selected the coordinates of the hexagonally packed cells, while the communication transitions were assigned guard $Gd2$ which mediated the communication between two neighbouring cells. Each arc in the six membrane compartments for the proximal and distal cell edges was assigned an expression $(((x, y), (a, b)), r)$, where the coordinate tuples $((x, y), (a, b))$ described the arcs linking the associated place to a particular transition in compartment (a, b) in cell (x, y), and r took a value of the colorset CSI. In the middle compartment, the arc expression was $((x, y), (2, 2))$. In the communication region, each arc was assigned an expression $NN(x, y, a, b, r)$++$NW(x, y, a, b, r)$++$SW(x, y, a, b, r)$ which defined how neighbouring cells communicated with each other. Finally the HCPN of the abstract PCP model, illustrated in Figure 4.3, were generated. Unfolding this model gives the plain Petri net

model in Figure 4.1.

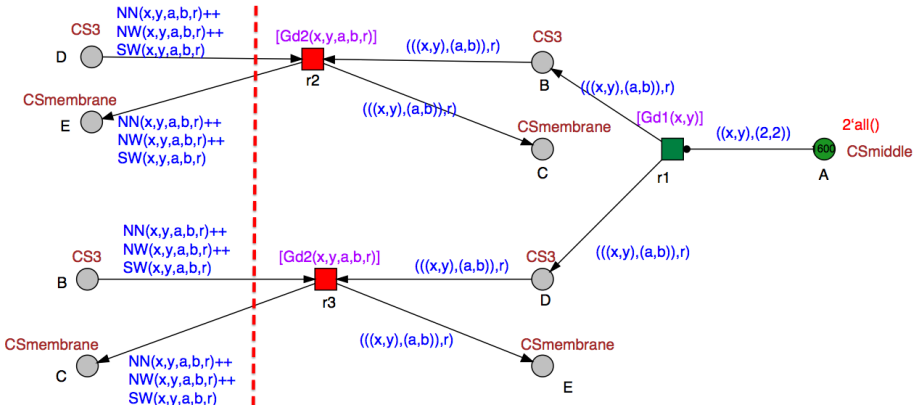

Figure 4.3: Abstract HCPN model, folded version of Figure 4.1. Places B, C, D, and E are logical nodes which are in the distal compartments of each cells [Gao 12].

This is a generic model able to generate honeycomb tissues of arbitrary size by adjusting the two constants M, N. The colour sets define a pattern which can easily be reused to model similar scenarios of spatial locality. For example, the colour sets for a one-to-one HCPN model retain most of what has been established for the above HCPN model. Since this neighbourhood relationship only enabled each compartment of the cell to communicate with one compartment in an adjacent cell, it is achieved by simply removing CSI, $CS3$, and the parameter CSI r in the neighbourhood functions.

In summary, the procedure to construct a HCPN model for a multi-cellular tissue in a hierarchy, e.g. regularly structured compartments inside each cell, can be divided into two steps.

- Fold cells: build up the structure of multi-cells representing the locality of each cell.

- Fold compartments: create the required localisation of compartments within a cell.

This can be trivially extended to model systems with more than two levels of hierarchy.

4.2.3 Model validation

Animation is an important technique for obtaining an intuitive understanding of a Petri net model as it demonstrates the dynamic behavior of the model in a visual way. Nearly all visual tools for modeling Petri nets provide the animation functionality [Nets 12]. Animation of colored Petri nets [Liu 10] enables users to experience the model behavior by following the token flow, which has been used to establish the confidence of the colouring template described above. The animation shows that target species representing report molecular species (FFD) in the abstract model accumulated equally in all six membrane compartments.

Table 4.1: Declarations for the abstract and refined HCPN models. Abbrev used: "const" – "constant", "cs" – "color set", "var" – "variable", and "fun" – "function".

Category	Declaration
const	$int : M = 40, N = 40, R = 3, C = 3$
cs	Row = int with $1 - M$
cs	$Column$ = int with $1 - N$
cs	$CS1$ = product with $Row \times Column$
cs	$CS\_Cell = CS1$ with $x\%2 = 1\&y\%2 = 0 \mid x\%2 = 0\&y\%2 = 1$
cs	$ComR$ = int with $1 - R$
cs	$ComC$ = int with $1 - C$
cs	CS_ComP = product with $ComR \times ComC$
cs	$CS2$ = product with $CS\_Cell \times CS\_ComP$
cs	$CSmembrane = CS2$ with $b = 1 \mid b = 3$
cs	$CSmiddle = CS2$ with $a = 2\&b = 2$
cs	CSI = int with $1 - 2$
cs	$CS3$ = product with $CSMembrane \times CSI$
var	$x : Row, y : Column,$ $a : ComR, b : ComC, r : CSI$
fun	NN (Row x,$Column$ y,$ComR$ a,$ComC$ b,CSI r) $CSMembrane$: $\{[(!(x = 1 \mid x = 2))\&(r = 2\&a = 1\&b = 1 \mid r = 1\&a = 1\&b = 3)], (((x - 2, y), (a + 2, b)), +r)\}$
fun	NW (Row x,$Column$ y,$ComR$ a,$ComC$ b,CSI r) $CMembrane$: $\{[(!(x = 1 \mid y = 1))\&(r = 1\&a = 1\&b = 1 \mid r = 2\& a = 2\&b = 1)], (((x - 1, y - 1), (a + 1, b + 2)), +r)\}$
fun	SW (Row x,$Column$ y,$ComR$ a,$ComC$ b,CSI r) $CMembrane$: $\{[(!(x = M \mid y = 1))\&(r = 1\&a = 2\&b = 1 \mid r = 2\& a = 3\&b = 1)], (((x + 1, y - 1), (a - 1, b + 2)), +r)\}$
fun	$Gd1$ (Row x,$Column$ y) $bool$: $\{x\%2 = 1\&y\%2 = 0 \mid x\%2 = 0\&y\%2 = 1\}$
fun	$Gd2$ (Row x,$Column$ y,$ComR$ a,$ComC$ b) $bool$: $\{(!(x = 1 \mid x = 2))\&(a = 1)\&(r = 1\&b = 3 \mid r = 2\& b = 1) \mid (!(x = 1 \mid y = 1))\&(b = 1)\&(r = 1\&a = 1 \mid r = 2\&a = 2) \mid (!(x = M \mid y = 1))\&(b = 1)\& (r = 2\&a = 3 \mid r = 1\&a = 2)\}$
fun	$RectangleReg$ (Row x,$Column$ y) $bool$: $\{x >= 15\&x <= 25\&y >= 20\&y <= 25\}$
fun	$DiskReg$ (Row x,$Column$ y,Row $x0$,$Column$ $y0$,int radius) $bool$: $\{(abs(y - y0) <= radius)\& (abs(x - x0) + abs(y - y0) <= 2 * radius)\}$
fun	$OvalReg$ (Row x,$Column$ y,Row $x0$,$Column$ $y0$,int rx,int ry) $bool$: $\{(rx < ry)\&(abs(x - x0) <= rx)\& (abs(x - x0) + abs(y - y0) <= ry) \mid (rx >= ry)\&(abs(y - y0) <= ry)\& (abs(x - x0) + abs(y - y0) <= rx)\}$

4.3 Ordinary differential equation models

The development of the ODE models proceeded in several steps, beginning with the choice to represent the continuous movement of the PCP proteins as transport and to represent the interactions between the proteins as the formation of protein complexes. The accumulation of a protein at a location could then be achieved by the reduced ability for these complexes to dissociate. Such complexes could also form across cell membranes with neighbouring cells, restricting some complexes to the shared edge between adjacent cells.

The reactions between individual proteins to form complexes occur as discrete events, resulting in a stochastic process. However, when the number of reacting molecules is large, their quantities may be expressed as concentrations and their reactions can be treated as deterministic, continuous equations by the law of mass action. The total absolute quantity of protein in the system was not known, it was chosen to begin with a deterministic model, since the feedback loop model [Tree 02a] did not reply on random fluctuations in order to function. Furthermore, the fluorescence microscopy techniques used to visualise the sub-cellular localisation of these proteins usually require a large number of proteins for detection [Roge 98].

The feedback loop model is completely symmetric and, as discussed in the previous chapter, functions as an amplification and local error-correction mechanism according to this hypothesis. The interaction of the signalling network with the upstream polarity genes (such as *fat*, *fj* and *ds*) is not well understood, but somehow this interaction must bias the signalling network to uniformly orient the cell polarity toward the distal vertex. The two general possibilities for introducing this asymmetry are through the gradient of some signal throughout the tissue, and through boundary conditions in the cell network that would be propagated by the signalling mechanism, relaying the polarity signal from cell to

cell to correctly polarise the entire tissue. As the biological evidence suggests that global directionality does not come only from the boundaries of the tissue (see Chapter 3), the global directional cue has been chosen to represent a signal influencing the activation of Frizzled (Fz) in every cell. In addition, the interaction of the polarised vesicle transport (see Chapter 3) with PCP signalling is still controversial. Vesicles containing Flamingo (Fmi), Frizzled (Fz), and Dshievelled (Dsh) tagged by green fluorescent proteins, have been observed to be transported preferentially toward the distal cell cortex on a polarised microtubule cytoskeleton during the build-up of intracellular asymmetry [Shim 06]. This biased transport was modelled by using different constant rates for those involved proteins (Dsh, Fmi and Fz) transporting toward the proximal and distal cortexes, whereas keeping the same transport rates for the rest of the core proteins. Furthermore, it has not yet been settled whether PCP signalling is established via one of the mentioned systems or by two systems acting in parallel. Uncertainty about the form and nature of the global directional cue and biased vesicle transport has led to a number of implementations and interpretations concerning their roles in the models (see Chapter 5).

4.3.1 Modelling the core machinery of PCP

An initial ODE model of PCP was created, involving the core proteins: Flamingo (Fmi), Frizzled (Fz), and Dishevelled (Dsh), Prickle (Pk) and Van Gogh (Vang). These proteins participate in two sets of reactions to form complexes, FmiDshFz and FmiVangPk, which can be depicted by the following mass action equations:

Note: ^{a} indicates that the protein is from a neighbouring cell; ^{b} a homodimer formed by two transmembrane proteins, Fmi; † indicates that it is an active form of Fz.

$$Fmi + Fmi^a \underset{kd1}{\overset{k1}{\rightleftharpoons}} FmiFmi^b \tag{4.3.1}$$

$$FmiFmi + Fz^\dagger + Vang^a \xrightarrow{k2} FmiFz + FmiVang^a \tag{4.3.2}$$

$$FmiFmi + Vang + Fz^a \xrightarrow{k3} FmiVang + FmiFz^a \tag{4.3.3}$$

$$FmiFz + Dsh \underset{kd4}{\overset{k4}{\rightleftharpoons}} FFD \tag{4.3.4}$$

$$FmiVang + Pk \underset{kd5}{\overset{k5}{\rightleftharpoons}} FVP \tag{4.3.5}$$

Where k1, k2...k5 are forward reaction rate constants. kd1, kd4 and kd5 are backward reaction rate constants. These sets of reaction represent the initial state of PCP signalling where the formation of FFD are equal along the cell membrane.

This initial ODE model involves only the basic proteins, complexes and reactions. The first attempt was to build a qualitative Petri net model of two neighbouring cells this initial ODE model to confess ourselves, then to quantify the model with parameter values, resulting a quantitative Petri net model with an underlying ODE system (each place in the Petri net model is represented by one ODE). Parameter estimation was initially performed by trying reasonable guesses for parameter values and manually refining these values to reproduce the wild-type phenotype (the accumulation of Flamingo-Frizzled-Dishevelled (FFD) complexes).

Results from this initial quantitative Petri net model were promising, exhibiting the desired behaviour at the initiation of PCP signalling where FFD are equally distributed

along the cell membrane. The model correctly shows the symmetric accumulation of FFD along the cell membrane of simulated wild-type cells.

Although this early model of PCP is successful in demonstrating what has been expected, the model lacks the necessary degrees of freedom to capture additional complicated behaviours that are characteristic of polarity signalling in the wing. For example, this model does not include the feedback loop and the transport of the core proteins from the cytoplasm to the membrane. Additionally, the influence of these two possible systems (morphogen factor X and polarised vesicle transport) which introduce the asymmetry to PCP signalling. Hence, this model is incapable of reproducing both the autonomous and nonautonomous phenotypes.

4.3.2 Modelling the hypotheses

Based on the above assumptions, the model was modified to include the feedback loop, where Vang/Pk inhibits the recruitment of Fz/Dsh to the membrane and Fz/Dsh inhibiting the recruitment of Vang/Pk to the membrane [Axel 01, Tree 02a, Zall 07] which are modelled as the following mass action equations:

Note: ^{pv} and ^{fd} indicates that the complexes form through different processes of the feedback loop.

$$FFD + FVP \xrightarrow{kf1} complex^{pv} \xrightarrow{kf2} FVP + Fmi + Fz + Dsh \tag{4.3.6}$$

$$FFD + FVP \xrightarrow{kf3} complex^{fd} \xrightarrow{kf4} FFD + Fmi + Vang + Pk \tag{4.3.7}$$

Where kf1, kf2, kf3 and kf4 are feedback-loop reaction rate constants. These two feedback reactions are symmetrically introduced into the model, which function as an

amplifier in the asymmetry of PCP signalling. This resulting system is one of the ODE models which later is used to construct a family of HCPN models to explore the underlying mechanisms of PCP.

The production and transport of proteins are modelled as a set of mass action equations as follows:

Note: $Prot$ represents the production of protein; * indicates that the protein is in the nucleus, waiting to transport to the membrane via vesicles.

$$Prot \xrightarrow{kp1} Fmi^\star \xrightarrow{kt1} Fmi \tag{4.3.8}$$

$$Prot \xrightarrow{kp2} Fz^\star \xrightarrow{kt2} Fz \tag{4.3.9}$$

$$Prot \xrightarrow{kp3} Dsh^\star \xrightarrow{kt3} Dsh \tag{4.3.10}$$

$$Prot \xrightarrow{kp4} Vang^\star \xrightarrow{kt4} Vang \tag{4.3.11}$$

$$Prot \xrightarrow{kp5} Pk^\star \xrightarrow{kt5} Pk \tag{4.3.12}$$

Where kp1, kp2...and kp5 are production rate constants. Note that, for convenience, the production have been modelled as a constant supply of proteins. kt1, kt2...and kt5 are transport rate constants, while kt1, kt2 and kt3 can be different due to which cortex those proteins are transported to in the scenarios that involves polarised vesicle transport.

Furthermore, the global directional cue (factor X) which introduces the asymmetry to PCP is modelled as the following mass action equation:

Note: † *indicates it is an active form of Fz which is the only form able to form FFD complexes.*

$$Fz + Ld \xrightarrow{kx} Fz^{\dagger} \tag{4.3.13}$$

Where kx is the reaction constant rate. In order to model the role of the global directional cue, an intermediate form of Fz is introduced after it transports from the cytoplasm to the membrane and assumed that only active form of Fz can participate PCP signalling (Equations 4.3.2 and 4.3.3). Note that this reaction is sequential to Equations 4.3.9, and its product $Fz^{\dagger}$ is the sole form of Fz to be involved in the formation of FFD complexes.

In order to enable the model to reproduce PCP signalling in *Drosophila* wing, an in-silico wing tissue (a multicellular system) is essential to discover the underlying mechanisms. The sets of equations described above are later used to build multi-scale models using Hierarchically Coloured Petri Nets (HCPN).

4.4 Hierarchically coloured Petri nets models of PCP signalling

A set of more detailed PCP models were developed step-wise by refining the model in Figure 4.3. Theses HCPN models incorporate equations described in Section 4.3.

An unbiased model (see Figure 4.4) was first constructed by using the equations described in Section 4.3.1, which did not include any proposed biased systems, to recapitulate the symmetrical localisation of proteins before the polarisation of PCP signalling. This model creates an equal distribution of FFD in all membrane compartments

in those cells having six neighbours (see Figure 4.4). The boundary cells or cells communicating with less than six neighbours do not have a symmetric localisation of FFD due to the lack of intercellular communication.

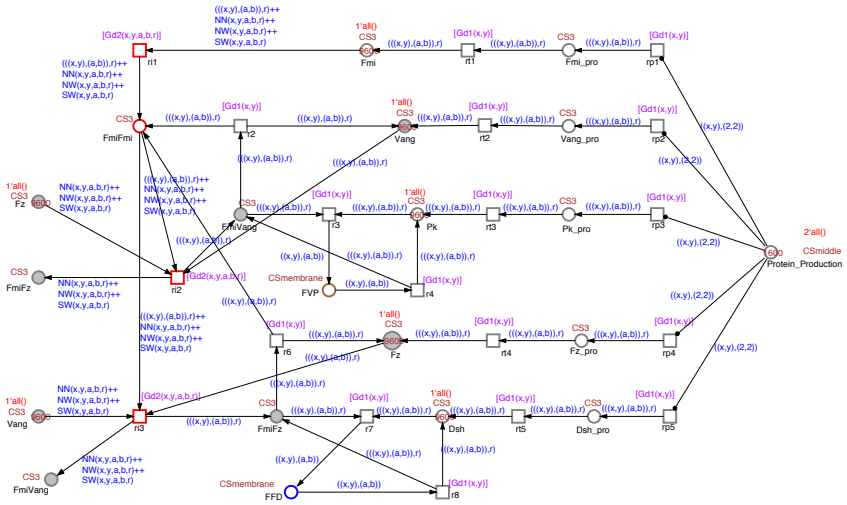

Figure 4.4: HCPN model of unbiased PCP [Gao 12]. It refines the abstract HCPN model given in Figure 4.3.

To construct this unbiased model the same declarations as those used for Figure 4.3 (see Table 4.1) were used, thus, there was no need to start from scratch. Transitions and places were grouped into different spatial regions and virtual compartments, and then the same colour sets were assigned to each region or compartment as done for Figure 4.3, likewise for arc expressions. Additionally, three functions, $RectangleReg$, $DiskReg$ and $OvalReg$ were defined to generate a rectangle, circle-like (hexagonal) or oval region within the whole tissue, respectively (see Table 4.1, Figure 4.5) which would be used to place mutant cells on the tissue to model biological mutant clones. Therefore, the models can not only perform simulations of PCP signalling in normal, wild-type cells but also on clones of mutant cells in a wild-type background.

A family of HCPN models were developed based on this unbiased model by introducing

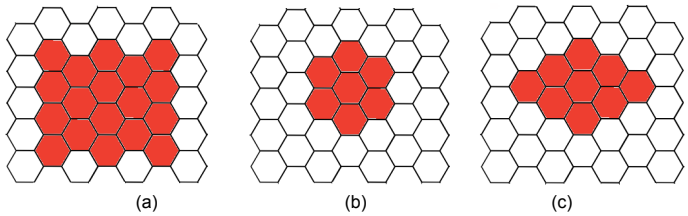

Figure 4.5: Diagram of possible shapes of a mutant clone that can be introduced to the in-silico tissue. From left to right: rectangular region, hexagonal region and oval region (labelled in red).

feedback loop, factor X, polarised transport in different combinations in order to test different hypotheses:

- By introducing the inhibitory loop, which mediate a competition between proximal and distal proteins displayed on adjacent surfaces of neighbouring cells [Axel 01, Tree 02a, Zall 07] (see equations of the feedback loop in Section 4.3), to the unbiased model, a biased feedback model was obtained. Subsequently, a set of HCPN models based on this basic biased model by introducing either factor X or biased transport alone, or both combined, were constructed. These variants can be used to test various hypotheses underlying PCP signalling, e.g. how the global directional cue interacts with the feedback loop, how the inhibitory loop responses to the polarised transport, and if these two factors act in parallel or only one of them plays a critical role.

- Two additional HPCN models were created by introducing only factor X or polarised transport alone to the unbiased model.

In order to simplify the identification of these models, the biased and unbiased versions are distinguished. If the model is biased (B) then it optionally includes feedback loop (F),

biased transport (T) or factor X ligand (X), or both or all of them. The neighbourhood relation can be one-to-one or one-to-two. Finally a biased model can be wild-type (Wt) or contain one of six mutant clones, *fmi*-, *fz*-, *dsh*-, *vang*, *pk*- or *pk*+ (see details in Appendix A). This naming convention can be described by the following simple BNF:

Note: '-' indicates the loss of function of the protein; '+' indicates the protein overexpression.

B [F] [X] [T] [1|2] [Wt|Fmi|Fz|Vang|Dsh|Pk|Pk+].

Details of the unbiased model and variants of the biased model, and the list of mutant clones can be found in Appendix A.

4.5 Quantifying HCPN models

The models were quantified by assigning to each transition a rate function following mass-action kinetics. The kinetic parameters were optimised by using simulated annealing to fit against constraints obtained from qualitative features of PCP phenotypes observed in biological laboratories [Axel 01, Gubb 82] (see details in the following section).

These quantitative models can be equally read as stochastic or continuous models, with appropriate scaling of the kinetic constants. Particularly, a continuous HCPN model uniquely defined a set of ODEs. To simulate wild-type and different mutant conditions, individual marking sets, function sets, and parameter sets (which are maintained within one model file) were considered.

The unspecified model parameters were the initial concentrations for each of the proteins and complexes involved (e.g. denoted $[Dsh]_0$), and the constant rates for each of the reactions. The unit of measure for concentration in the problem was unspecified, thus, the number of parameters were reduced by arbitrarily fixing the initial concentrations

of the core proteins (denoted $[Dsh]_0$, $[Fmi]_0$, $[Fz]_0$, $[Pk]_0$, and $[Vang]_0$) to 1.

4.5.1 Hair growth direction

The hair growth direction for each simulated cell was determined based on the protein concentration distribution at the end of the simulation. The distribution of FFD was chosen to predict the hair growth direction because polarisation of hair cells recruits tissue specific effectors, in this case – FFD, to proximal and distal sides of the cell that control actin polymerisation and bundling to produce a distal hair [Adle 04, Stru 08]. Therefore, the hair growth direction predicted by the HCPN models are represented as the localisation of FFD in each cell.

4.5.2 Parameter sensitivity analysis

Parameter sensitivity analysis was implemented by using an extension of the ODE15s Matlab ODE solver [MATL 11, Sham 97, Sham 99]. It solves ODE systems, at the same time, calculates derivatives (sensitivities) of the solution with respect to parameters [Bock 81, Maly 96].

There are several issues needed to be concerned to use this extension:

- The main question is about the accuracy of the computed derivatives. It is difficult to give a precise answer, but it has been observed that it gives accuracies of the order of the integrator accuracies. These accuracies can be controlled through the original ODE15s tolerances; the tighter the ODE15s tolerances, the better the derivatives accuracy. Therefore, the accuracy attainable is problem-dependent. It has been found that the default relative tolerance in ODE15s (1e-3) is too coarse to obtain good accuracies to the derivatives, so that it has been set to 1e-6.

- As The vectorization of the m-files can make the subroutine fail when: (i) extra parameters are passed; (ii) the parameters, whose derivative needs to be found, have no influence on the system.

The sensitivity of all parameters involved in the signalling was computed with the respect to the concentration of FFD complexes, and found that not all parameter values had a high impact on the behaviour of the signalling. Thus, the number of parameters that would be taken into account in performing parameter optimisation (see Section 4.5) were reduced. Results of sensitivity analysis of BFXFz model are shown in Table 4.2 as BFX model is the most interesting and promising (see discussion in Chapter 5).

Parameter	Sensitivity value	Parameter	Sensitivity value	Parameter	Sensitivity value	Parameter	Sensitivity value
k_1	1.4077	k_5	-0.0664	k_{p1}	0.0028	k_{t2}	0.0005
k_{d1}	0.0195	k_{d5}	-0.0051	k_{p2}	0.0005	k_{t3}	-0.0003
k_2	-0.1699	k_{f1}	-0.0136	k_{p3}	0.0006	$k_{t4}1$	0.0001
k_3	-0.2670	k_{f2}	-0.0226	k_{p4}	0.0000	k_{t5}	0.0000
k_4	1.1606	k_{f3}	0.0175	k_{p5}	0.0000	k_x	0.5092
k_{d4}	-0.0063	k_{f4}	0.0507	k_{t1}	0.0036	k_{mutant}	1.8360

Table 4.2: Overall Sensitivity of all parameters using BFXFz model comprising 18 cells.

According to the sensitivities, 12 highly influential parameters [1] out of the total set of 24 parameters were selected to perform parameter optimisation in order to quantify the HCPN model. Note that k24 is constant rate used for the knock-out, which must be zero in order to reproduce the scenario of loss of function of a particular mutation (e.g. *fz*-). It is crucial to any models with mutant clones consistent with the sensitivity results (k24 has the highest impact on the signalling output).

However, testing the sensitivity of a model to single parameter variations does not guarantee that the model is not sensitive to other combinations of parameter changes, nor is it necessarily appropriate, since the model may be insensitive to simultaneous changes

[1] Influential parameters: parameters whose sensitivity is higher than 0.01 or lower than -0.01

in several parameters, while remaining sensitive to changes in the relative values of some parameters. Methods to reliably identify model parameters that are robust to parameter variations remains an open problem and will be discussed further, along with the biological basis for this form of robustness in Chapter 6.

4.5.3 Parameter Optimisation

4.5.3.1 Simulated annealing

Simulated annealing (SA) is a randomised local search method based on the simulation of annealing a metal, which was firstly described by Kirkpatrick et al in 1983 [Kirk 83]. The application of SA has been widely used to address optimisation problems with/without constraints. The principle of SA models is to heat and lower system temperature iteratively to reduce system defects and to enable the system to reach a minimum energy status.

In order to search for optimum solution, SA randomly generates a trial solution based on the current solution at each step. The trial solution is evaluated by an objective function to decide if the trial solution can be accepted to replace the current one, which is called an acceptance probability. The system favours to go downhill by lowering the temperature when the temperature is high. Whereas uphill moves enables the system to escape from the local minima, without which, the system would be trapped into a local minimum. This probability is controlled by temperature in such a manner that at the beginning of the procedure the temperature is sufficiently high, in which a high probability is available, and as the calculation proceeds the temperature is gradually decreased, lowering the probability.

The procedure of SA consists of two nested loops: an outer loop and an inner one, where the cooling and the equilibrating processes are performed, respectively. This search procedure is iteratively repeated and won't stop until the terminated conditions reach.

The SA routine is represented in pseudocode format in Algorithm 1, and implementation details can be found in the following.

```
Require:
  choose the initial solution; calculate the cost (fitness);
  repeat
    outer loop (cooling)
    repeat
      inner loop (equilibrating)
      perturb the current solution;
      calculate the cost;
      accept the trial solution as a new solution by acceptance probability;
    until
    iteration limit is reached;
    temperature is lowered according to the cooling schedule;
  until termination condition is satisfied;
  return Optimal solution
```

Algorithm 1: Pseudocode representation of simulated annealing

4.5.3.2 Implementation of parameter optimisation

The selected parameters were optimised by constraining them so that the simulation would reproduce qualitative features of the experimentally observed PCP phenotypes. For example, one case involves simulating clones of cells lacking *fz* function surrounded by wild-type cells, and constraining the parameters to make sure at the end of the simulation, reversed hair polarity is produced in a wild-type cell that is distal to the mutant cells. The degree of asymmetry was scored in this cell based on the asymmetry distribution of the FFD concentration at the end of a simulation for a given set of parameters.

$$S_{fz} = -\frac{[FFD]_{total}^{distal} - [FFD]_{total}^{proximal}}{[FFD]_{total}} \tag{4.5.1}$$

Then the constraint was implemented as a quadratic penalty function with weight, K_{fz}, that was nonzero when the polarity in this cell was not reversed above a threshold target value, T_{fz}.

$$F_{fz} = \begin{cases} 0 & S_{fz} \geqslant T_{fz} \\ K_{fz} \times (T_{fz} - S_{fz})^2 & S_{fz} < Tfz \end{cases}$$

This form of the penalty function has the advantage of having a continuous derivative and does not require that the initial parameter set be feasible. Two feature constraints corresponding to the characteristic PCP phenotypes listed in Table 4.3 were used to train the models. A fitness function evaluation of the SA procedure described in Algorithm 1 was applied to different cases. The parameters were selected by iteratively running the optimisation routine after adjusting the weights of the penalty functions for each of the individual feature constraints, cooling and equilibrating processed until the desired behaviour of the system was achieved. In order to reduce the amount of computational effort required for each fitness function evaluation, cases involving mutant clones were simulated on lines of wild-type cells extending parallel to the proximal-distal axis and interrupted by mutant cells (see Figure 4.6) using models in a size of either 6 or 18 cells. However, it still demands very heavy computational power to run the optimisation routine (see Figure 4.7) due to the search space for the rate parameters and the number of underlying ODEs of the simulated model to solve (e.g. 6-cell model comprising of 1,086 underlying ODEs, while 18-cell consists of 3,258 underlying ODEs) each time manipulating the penalty function.

The same SA optimisation routine was applied to the family of models generated in Section 4.4 .

Once parameters that satisfied the constraints on the simple lines of cells had been

Model	Fitness	Constraint description
Biased model with a *fz* clone	$F_{fz} > 0$ in cells distally adjacent to the *fz*-clone; $F_{fz} < 0$ in other wild-type cells far away from the clone	Distal domineering nonautonomy [Gubb 82]

Table 4.3: Characteristic PCP phenotypes used as constraints to fit the HCPN models.

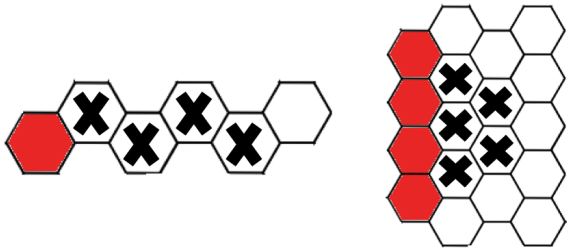

Figure 4.6: Cell configuration for parameter optimisation: (a) simulated cell array comprises 6 cells: four cells distally to the mutant cell are chosen to be fit (labelled by x), (b) simulated cell array comprises 18 cells: five cells are selected to fit (labelled by x). Mutant cells are labelled in red.

identified, they were verified over a larger two-dimensional honeycomb lattice to see if the model could still exhibit the desired behaviour which was used to produce the feature constraints. The selected parameters are shown in Table A.2.

The parameter optimisation routine (SA) was implemented in Matlab ®7.11.0 [MATL 11] because Matlab facilitates to plot and examine the fitness function values makes it easier to follow the progress of the parameter optimisation. A procedure was followed in order to implement the optimisation routine: (i) exporting ODEs of an unfolded Petri nets at a required size from Snoopy [Hein 12b, Rohr 10]; (ii) converting these ODEs to a Matlab readable format; (iii) applying the SA routine to optimise the parameters in the ODE model. The computational effort required by the optimisation algorithm is small compared to the effort required in evaluating the fitness functions due to the time-costing of solving

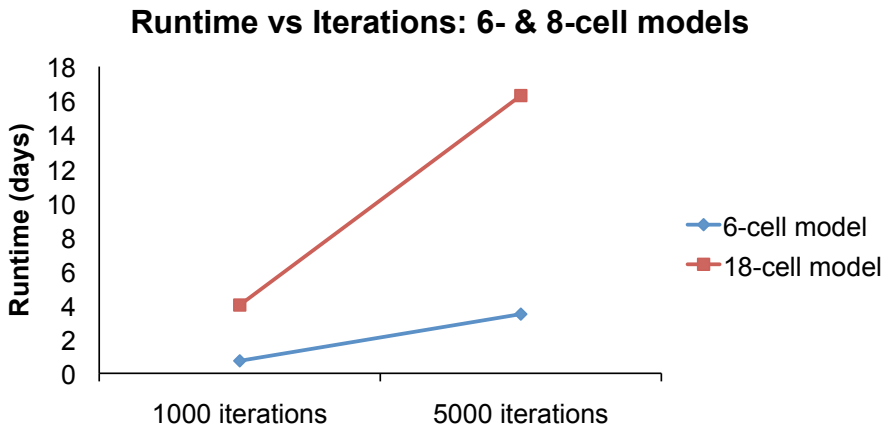

Figure 4.7: Runtime for parameter optimisation (time unit is in days). It shows the relationship between the runtime of SA routine and the number of iterations as well as the number of ODEs to be solved (representing by the size of the model): (i) the more iterations the SA routine implements, the longer runtime is; and (ii) the larger the model is, the more computational effort is required.

the large underlying system of ODEs.

4.6 Discussion

A number of assumptions were made to construct the HCPN models of PCP. This section summarises the most important of these, discussing their basis and validity, and noting the consequent limitations on the models and conclusion drawn from their results.

In addition, in this work, the simulation and fitting of a multi-scale model consisting of a large number of biological entities, e.g. a model comprises over 150,000 ODEs (see Table 4.5), requires high computing power. Therefore, efforts made in this work in order to speed up computation in simulation and parameter optimisation have been discussed in this section.

4.6.1 Kinetics - deterministic versus stochastic

It is assumed that conditions are met such that protein concentrations may be treated as continuous, deterministic variables. The discrepancy between deterministic models and stochastic models of these reactions increases when the number of reacting molecules is small, the reactions proceed very slowly, or if the behaviour of the signalling network depends on the random fluctuations in protein concentrations, rather than on their average values [Blou 91, Gill 77, Kurt 72, McAd 97]. The biological model does not rely on fluctuations of protein concentrations, and the timescale for the asymmetric accumulation of Fz, Dsh, Pk, and Vang is much greater than the time necessary for protein-protein interactions. Furthermore, the minimal level of detection of fluorescent tagged molecules, localised to the plasma membrane, used to visualise these proteins suggests that the number of molecules present in each cell is likely to be on the order of thousands or more [Tsie 98]. Therefore, it will not be expected that the increased computational effort of stochastic models would be necessary, and indeed, it has shown that a deterministic model is sufficient

to reproduce the behaviour of the system.

4.6.2 Cell geometry

The planar nature of the signalling network permits the mathematical modelling of the biological system in two-dimensions. The computational grid geometry describes a regular hexagonal lattice of cells (see Figure 4.2) that remains fixed throughout the simulation. *Drosophila* wing epithelial cells are generally hexagonally packed, but the precise size and shape of the cells vary over the wing and during PCP signalling. These geometric variations likely influence precise hair angles, the degree of asymmetric localisation of proteins, and the extent of nonautonomous polarity disruptions in regions near some mutant clones. This work only attempted to reproduce general qualitative features of PCP phenotypes, and thus these variations were not considered in the HCPN models.

4.6.3 Feedback loop

The underlying mechanisms for the interactions in the local feedback loop are not yet fully understood, the essential logic of this feedback loop was preserved by representing these interactions as binding to form protein complexes. Direct biochemical evidence exists for the FmiVangPk complex [Bast 03, Jenn 03]. It has been observed that Fz induces the recruitment of Dsh to the cell membrane in vivo and in heterologous systems, and evidence of direct, albeit weak, binding exists for Fz and the PDZ domain of Dsh [Axel 01, Axel 98]. A direct FzFmi-FmiVang interaction through a homodimer Fmi-Fmi was modelled. Fz and Vang localise to opposite sides of the cell, and Vang accumulation depends on the Fz allele (autonomous or nonautonomous) in the adjacent cell (see Section 2 and [Amon 05]). Fz and Vang are linked through a mutual affinity for a bridging cadherin

such as Fmi, which is presented in both membranes, and is required for assembly of the other components [Usui 99]. Diego (Dgo) [Das 04, Feig 01], Wdb [Hann 02], and other proteins known to be involved in PCP signalling, were not included in the model, as their roles were in- sufficiently understood. Fmi is required for the membrane localisation of the other proteins [Bast 03], Dgo localises to the proximal cell surface [Das 04], and both are required to generate asymmetry. Fmi bound to Fz in one cell and Fmi associated with Vang in the other [Bast 03, Chen 08, Das 04, Lawr 04, Usui 99, Shim 06]. It is unknown how Dgo exactly influence other proteins. Thus, this work only includes Fmi in the models but not Dgo. When the role of Dgo is more fully understood, it may be included in the model without perturbing the function of the feedback loop, and could provide additional degrees of freedom that may enhance the concordance with experimental data. Therefore, the feasibility of the model should not be sensitive to these assumptions.

Both genetic and cell culture data previously led us to propose that Pk inhibits Fz dependent membrane recruitment of Dsh [Tree 02a]. In a cell culture assay, Pk was shown to cell-autonomously antagonises Dsh recruitment [Tree 02a]. This result has been replicated by some, but not others [Bast 03, Carr 03, Take 03, Veem 03]. New experimental results demonstrated a dose dependent inhibition of Dsh membrane recruitment by both Pk and Vang, and a combinatorial effect when both are expressed. This is consistent with the hypothesis that Pk and Vang work together on the proximal side of the cell, and with the observation that, like Pk, Vang binds Dsh [Bast 03, Park 02, Tree 02a]. The differences between these results and those of others might therefore depend on differences in the uncontrolled levels of host cell Vang, or other uncontrolled aspects of the experiment. It has observed that colocalisation of Dsh with Pk and Vang at the end of PCP signalling, suggesting that Dsh does not remain bound to Pk and Vang after this inhibition has taken

place [Axel 11]. Therefore, it is expected that the mechanism for this inhibition is more complicated than competitive binding, such as some modification that then destabilises the FmiFzDsh complex. The implementation of Pk and Vang inhibition used in the mathematical model only required that these proteins inhibit the FmiDshFz interaction, without relying on a specific molecular mechanism for this activity.

4.6.4 Modelling implementations

Several possible mechanisms exist for providing the asymmetric signal that the feedback loop interprets, and then use to produce, from an initially symmetric distribution, the asymmetric localisation of the proteins included in the model (Fz, Dsh, Vang, and Pk).

The gradient of global directional cue was modelled as a tissue-level gradient throughout the whole wing. This was implemented by assigning different initial concentrations of the factor X (denoted by $[Ld]_0$) in different regions, e.g. $[Ld]_0$ forms a gradient from the proximal end of the in-silico tissue to the distal end.

In addition, the modelling of the polarised transport of Fz, Fmi and Dsh towards the distal cell cortex via vesicles was achieved by assigning different parameter values to reactions which control the transport of corresponding proteins to the distal & proximal, respectively.

4.6.5 Computational speed-up

As the HCPN models requires computational experiments over very large underlying models which are represented by systems of ordinary differential equations (ODEs), a crucial point is how many cells that can be simulated in terms of current computational capabilities, i.e. what tissue size can be actually analysed.

4.6.5.1 Choice of parameters

The lack of a detailed biological understanding limits the ability of the mathematical model to make quantitative predictions about the detailed behaviour of PCP signalling. It has only enforced qualitative feature constraints during parameter selection, and so quantitative results from the model, such as the degree of asymmetry observed in each of the proteins or the extent of nonautonomy observed around mutant clones, could vary with the specific choice of parameters without violating the feature constraints. Further, the feature constraints were enforced only based on features of the results at the end of each simulation. Therefore, the dynamic behaviour of the model, which also depends on the choice of parameters, does not necessarily reflect the behaviour of the biological system. Finally, the relative values and stabilities of the model parameters are themselves dependent on the choice of the other parameters in the model. A different choice of parameters might exhibit different relative parameter values or stabilities, and so it has refrained from making any inferences about the relative importance of various parameters based on the current parameter set.

As mentioned in Section 4.5.3.2, evaluating the fitness functions is the most computational expensive procedure in the SA routine using Matlab. This is due to the large number of ODEs to be solved each time. Thus this step was accelebrated by converting the system of ODEs to jacobian matrix, binary matrix and sparse matrix, which could be critical to make stiff ODE solvers (e.g. Matlab ODE15S solver) execute faster. A massive speed-up has been achieved (see Table 4.4 for details), which as a result dramatically reduce the runtime for the SA routine.

Properties	Simulation time (in seconds)
ODEs	115.3
Jacobian matrix	84.5
Jacobian matrix & pre-allocation	76.4
Binary matrix	16.2
Sparse matrix	3.7

Table 4.4: Simulation time for 6-cell model using different properties for Matlab ODE15s solver. A massive speed-up of simulation has been achieved by using Sparse matrix.[a]

[a] performed on a MacBook Pro Intel Duo-core, CPU 2× 2.93GHz, memory (DDR 3) 4 GB;

4.6.5.2 Simulation of HCPN models

In order to simulate HCPN models, there are three technical key problems: unfolding, ODEs construction, and simulation. The coloured continuous Petri nets are automatically unfolded which can be considered as a kind of compilation. Afterwards, the corresponding ODE needs to be constructed for each place; the runtime for this step is negligible. Finally, the model has to be simulated using an appropriate continuous simulation algorithm. Considering the huge size of the models (see the number of places and transitions in Table 4.5).

Size				Runtime	
Grid($M \times N$)	Cells	Places	Transitions	Unfolding runtime (seconds)	Simulation runtime (seconds)
5×5	12	2,028	2,802	1.154	3.145
10×10	50	8,450	11,826	2.613	14.618
15×15	112	18,928	26,622	4.495	42.586
20×20	200	33,800	47,646	9.231	88.886
40×40	800	135,200	191,286	83.162	371.647
40×40^b	800	164,000	229,686	120.186	7,399.544

Table 4.5: Unbiased PCP model size and runtime[a] for unfolding and continuous simulation over 1000 time units, illustrating the efficiency of constraint solver.

[a] performed on a Mac Quad-core Intel Xeon, CPU 2× 2.26GHz, memory (DDR 3) 8 GB; [b] for the biased model BFXWt.

These improvements of unfolding, ODEs construction, and simulation in Snoopy [Hein 12b, Rohr 10] were conducted by Professor Heiner's group at the Brandenburg University of Technology, in Cottbus.

4.7 Summary

This chapter has illustrated the development of the HCPN colouring template which could be easily adapted by other systems with a similar hierarchical organisation, followed by the construction of a family of HCPN models which would enable us to test different hypotheses of PCP signalling and possibly make predictions of the underlying mechanisms of the signalling. In addition, the challenges of simulating such large models and optimising parameters in these models have been discussed. The next chapter describes how to check and verify the HCPN models constructed in this chapter in order to have some understanding of the biological system of PCP signalling.

Chapter 5

Validation and analysis

5.1 Introduction

This chapter illustrates an approach to analyse and check the multi-scale models of PCP signalling constructed in the previous chapter (Chapter 4). This is followed by a discussion, including insights into the domineering nonautonomy, as well as additional phenotype predications.

Hierarchically coloured Petri nets (HCPN) enjoy a large variety of analysis techniques, ranging from informal animation to formal static or dynamic analysis techniques. In the following, the simulative methods to analyse dynamic properties of the models in the continuous setting have been used, because the infinite state space caused by the infinite supply of proteins precludes the application of dynamic exhaustive analysis techniques. Specifically, automated clustering and visualisation techniques to identify sets of cells with similar behaviors are used, complemented by PLTL model checking and visual analysis to both primary and secondary time-series data from representative cells of automatically identified clusters.

The analysis begins with the experiments by using the unbiased and biased HCPN

models to generate an in-silico wild-type tissue based on a honeycomb grid comprising 800 cells imposed on an underlying 40×40 rectangular grid. Next, the effect of a clone of mutated cells lacking the key signalling molecules Frizzled (Fz) or Van Gogh (Vang), denoted by Fz- and Vang-, respectively, in an otherwise wild-type field of cells by zeroing the concentration and switching off transport in the corresponding Petri net places, is modelled. A mutant clone of *fz-* or *vang-* in a circle-like shape with 80 mutant cells,i.e. 10% of the in-silico tissue, using the $DiskReg$ function (see Table 4.1) are produced.

Later the influence of lacking of Dishevelled (Dsh), Flamingo (Fmi), Prickle (Pk), and overexpression of Pk, denoted by Dsh-, Fmi-, Pk- and Pk+, respectively, are modelled using the same approach as described above. Details for naming convention can be found in Section 4.4.

The analysis has been conducted as a proof of principle that the models indeed have the ability to capture the given biological phenomena and make sensible predictions. Some results presented in this chapter has been published in [Gao 12].

5.2 Simulations

All simulations were run over 1,000 time units reported at 200 time points. The time span represented PCP signalling beginning at 33 hours APF (after puparium formation) in-vivo [Shor 72]. Snoopy's (version: 02-7-2012) [Hein 12b, Rohr 10] built-in continuous simulators were used to simulate the models. Simulation traces had been written to csv files, which were then been further processed by Matlab®7.11.0 [MATL 11] for clustering and visualisation, and MC2 [Dona 08] for model checking.

5.3 Clustering analysis

The approach to analyse the models is based on the behaviour of each cell, represented by the Flamingo-Frizzled-Dishevelled (FFD) complex concentrations for each of the six membrane compartments because these are associated with the PCP response [Stru 01] (explained in Chapter 4 Section 4.5.1). In the first step, the time points of the FFD time series from the six membrane compartments were taken and then used as 'features' for each cell; they comprised the characteristics which defined the behaviour. With e.g. 200 time points for each of the six time series 1,200 features were obtained, which contained some less significant variables. This could slow down the clustering algorithm, and the clustering result may be diluted by the background noise.

Thus a feature selection technique, Principal Component Analysis (PCA) [Joll 86], was applied to reduce the number of features. This way, the original features were converted into new features, which are a linear combination of the previous ones, see Figure 5.1. PCA allows the calculation of the variance explained by each new feature in order to find how many new features that are needed to describe the space. The number of features that were needed to explain at least 95% of the variance were then checked.

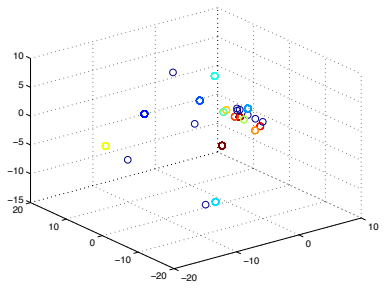

Figure 5.1: Representation of the most important three features derived by PCA for model BFXFz. PCA reduces the large number of features generated by the time series data, in which those original features are converted into a linear combination of them.

Once the new dataset had been obtained through PCA, a clustering algorithm was applied to find groups of cells with similar behaviour. There are many clustering algorithms available in literature, such as the commonly used K-means [MacQ 65] and hierarchical clustering [Hast 09]. However, these two traditional methods can not be relied on because the approach should not require a priori an expected number of clusters, but more importantly the clusters are not necessarily spherical. For this reason a density-based clustering algorithm called DBScan [Este 96], which permits finding clusters of any shape, was applied in this work. DBScan identifies dense areas of space by looking at the neighbourhood of each point. An area of a given radius is considered to be dense if it contains at least a given number of points. The algorithm then merges connected dense areas to form a cluster. All the points left out of the clusters are labelled as outliers. The best number of clusters to fit the scenario can be found by inspecting the three dimensional data space and adjusting the parameters (radius and size of the neighbourhood). The results were then plotted on a hexagonal grid, which represented the layout of the cells in the model, see experiments below. The cells in each cluster were allocated a unique colour and labelled by the cluster number generated by the technique. All outliers were allocated -1 as their cluster number.

To confirm the efficiency of the approach mentioned above [Gao 12], another two widely used clustering techniques, K-means and hierarchical clustering, were applied as comparison. A visual inspection of the data showed that clusters of common behaviours were characterised by non-spherical shapes. Moreover, the time series exhibited rather smooth behaviour, partly due to being the results of deterministic simulation. Therefore the application of clustering procedures belonging to the clustering categories, such as partitioning methods, hierarchical methods and model-based methods, did not lead to

useful results as shown in Figure 5.2, which was also confirmed by the cluster evaluation results. The clustering technique was expected to detect at least: (i) the boundary cells at the edge of the in-silico tissue; (ii)the mutant clone; (iii) the neighbouring cells to the clone. Hierarchical clustering failed to even detect the boundary cells, which K-means was successful to achieve. However, K-means is less efficient to capture features of the time series compared to DBScan combined with PCA (see Figure 5.2), e.g. boundary cells at the four corner of the in-silico tissue behave differently due to their localities, which was assigned as separate clusters from other boundary cells by DBScan with PCA but not distinguished by K-means (cluster -1 in Figure 5.2 (c)). To evaluate the goodness of clusters, a validity index named Composed Density between and within clusters (CDbw), which was introduced by [Halk 08], was calculated. This index measures the quality of the clusters by considering multiple representative points per cluster and therefore accommodates non hyper spherical cluster geometry. The higher the CDbw value is, the better the quality of clusters is (for more details, refer to [Halk 08, Macc 12]).

Clustering algorithms were encoded and implemented in Matlab. DBScan for Matlab was taken from [Dasz 01], which was later combined with PCA by [1]. K-means and hierarchically clustering were done by using Matlab's built-in functions. Applications of this clustering approach can be found in [Gao 12, Macc 12].

[1] D. Maccagnola was responsible for developing and encoding the clustering approach (DBScan combined with PCA), as well as the validity index CDbw into Matlab. Address: Dipartimento di Informatica Sistemistica e Comunicazione, Università degli Studi di Milano-Bicocca, Italy.

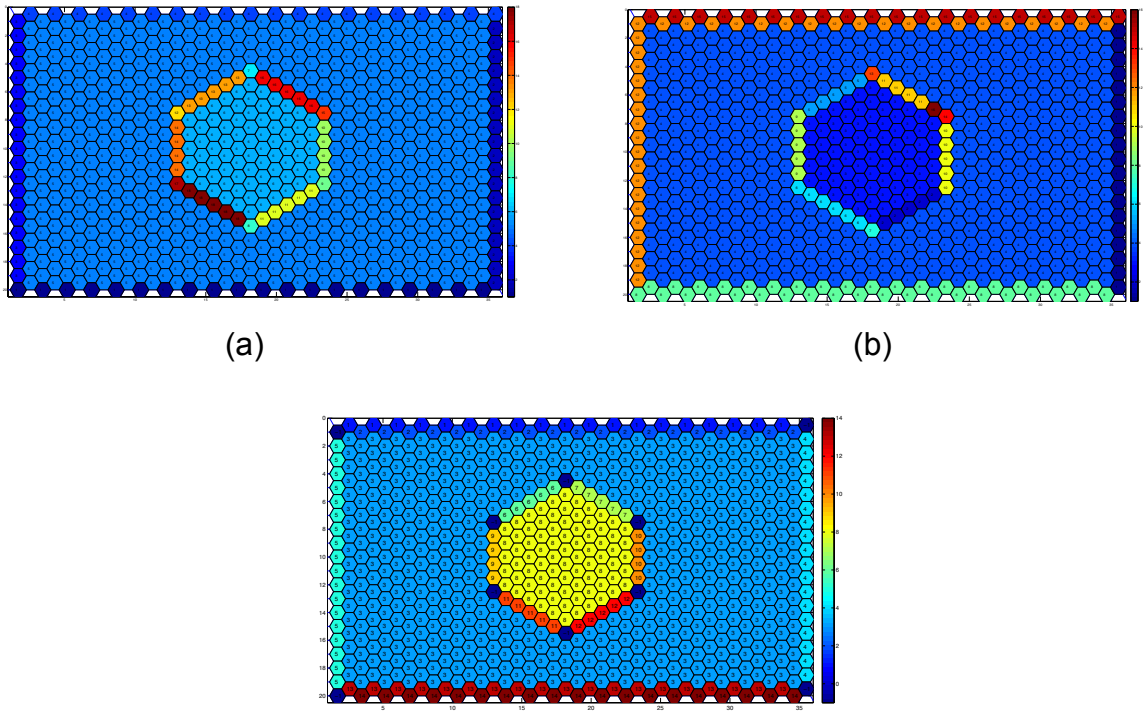

Figure 5.2: Comparison of validity of different clustering techniques over time-series data obtained from *BFXFz* model in **one-to-two** neighbourhood relationship. (a) Hierarchical clustering assigns 18 clusters; (b) K-means allocates 16 clusters; (c) DBScan combined with PCA detects 14 clusters. The approach, DBScan combined with PCA, has successfully captured most features of the time series, if not all, with the least number of clusters. Hierarchical clustering fails to detect boundary cells even by assigning the most number of clusters, while K-means is able to detect some of the features from the dataset, however, it still lacks the capability to acquire the other important features. Thus, compared to the other two clustering techniques, DBScan combined with PCA gives the best clusters.

5.4 Model Checking

In order to explore in more detail the behaviour of a model, model checking was performed on representative cells from the clusters identified by the cluster analysis. Model checking is a technique which is used to check the validity of properties of a dynamic model expressed in a temporal logic; this can be performed either analytically (requiring the entire state space to be explored), or over finite time-series traces produced by the simulation of the model. This work used simulative model checking because of the infinite state space of the models, and specifically employed the MC2 model checker [Dona 08] which operates over Probabilistic Linear Temporal Logic with constraints (PLTLc). Both the primary data output from a model (i.e. time series of concentrations) as well as secondary data – in this case time series of accumulated concentrations (known as cumulative rewards in CSL model checking [Schw 11]) were checked.

This secondary data has been checked because the localisation of PCP signalling components at any given time point is the result of the cumulative effect of the sum over the signalling events until that point. This was done by computing the accumulation of the concentration of FFD at each point in the time series trace, and called as *cumulative time series*. The cumulative time series data "smooths away" small variations which are apparent in the primary data.

5.5 Models

All experiments reported in this chapter refer to the one-to-two neighbourhood relationship models, namely the unbiased model and variants of the biased model, as defined in Section 4.4 (details see Table A.3), as one-to-two neighbourhood is more interesting

because: (i) in topological point of view, one-to-two neighbourhood could provide more variations in cell behaviours and may give more insights into the mechanisms underlying PCP signalling compared to one-to-one neighbourhood (Figure 5.3); (ii) the simulation results for mutant conditions confirm that one-to-one can not reproduce the biological observations, where cells adjacent to the clones behaves similarly to other wild-type cells (not distal domineering nonautonomous), see Figure 5.4.

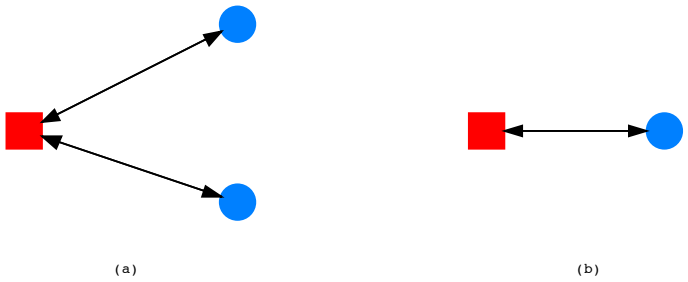

Figure 5.3: Representation of one-to-one and one-to-two neighbourhood relationships. (a) *one-to-two*, where each compartment in a cell communicates with two other compartments in the neighbouring cell(s); (b) *one-to-one*, where on individual compartment in a cell communicates with only one compartment in the neighbouring cell.

In addition, boundary cells (the most outside cells) are not considered when analysing and checking the model.

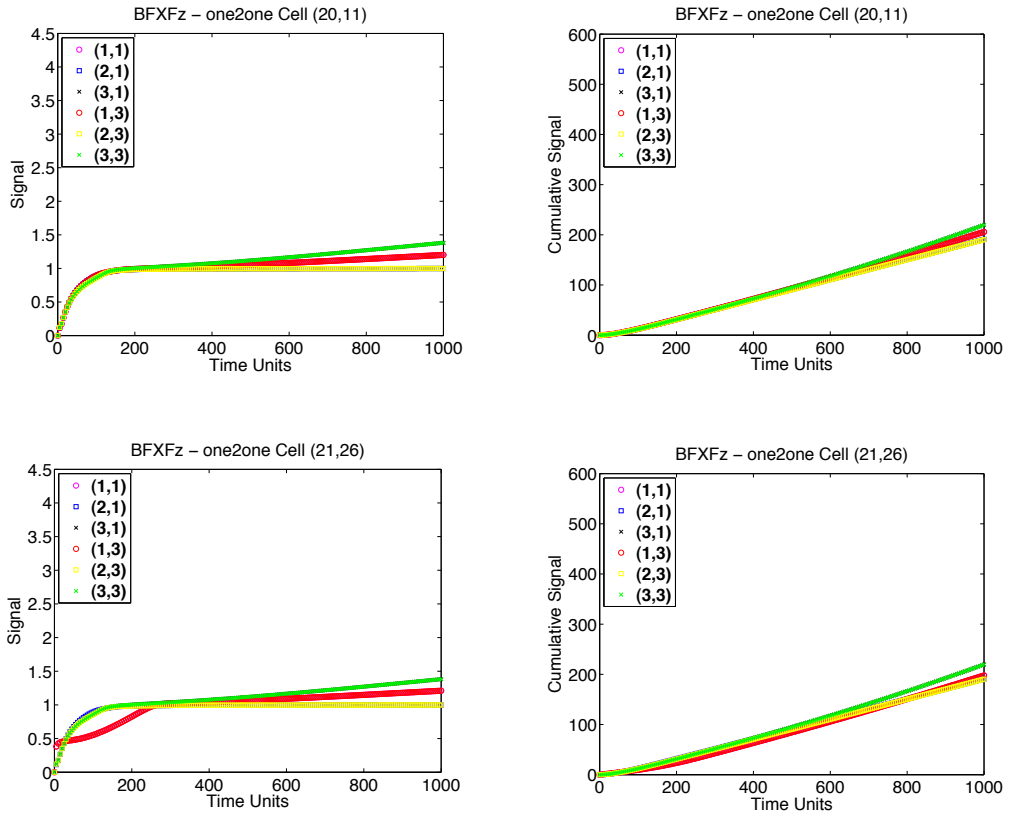

Figure 5.4: BFXFz model in one-to-one neighbourhood: continuous simulation for a representative of wild-type cells away from the clone (top left); cumulative signal for a representative of wild-type cells away from the clone (top right); continuous simulation for a representative of cells distally adjacent to the clone (bottom left); cumulative signal for a representative of cells distally adjacent to the clone (bottom right).

5.6 Experiments and results

In order to validate the ability of the models to recapitulate biological phenomena observed in the wet lab, the experiments start off by simulating and analysing the wild-type models, before considering the effect on neighbouring wild-type cells of a clone of mutant cells lacking Fz or Vang proteins or other core signalling proteins mentioned at the beginning of this chapter. Then mutant clones are introduced to the in-silico tissue in order to reproduce well-documented PCP phenotypes.

An approach including clustering using DBScan with PCA and PLTLc model checking, is proposed in this section to analyse and check the family of HCPN models against qualitative features of the experimentally observed PCP phenotypes. The approach could be applied to analyse other multi-scale models with similar features. Some results were published in [Gao 11, Gao 12].[2]

Experiment 1: Unbiased model

Current biological models of PCP in the literature [Stru 01] show that at initialisation of signalling PCP proteins are symmetrically distributed along the cell membrane. In order to mimic this, the unbiased model was used to reflect this symmetry.

Clustering. It is expected that FFD complexes would be equally distributed over the six membrane compartments in those cells which communicate with six neighbouring cells (i.e. except boundary cells). The continuous simulation result confirms the expected behaviour, as shown in Figure 5.5; all cells with six neighbours belong to cluster 3 (labelled in green in Figure 5.5). Cells at each boundary are separately assigned to other clusters;

[2]Model checking queries: the concentration of FFD in the proximal and distal compartments (1,1), (2,1), (3,1), (1,3), (2,3), and (3,3) are represented by P_1, P_2, P_3, D_1, D_2 and D_3.

due to the different number of communication neighbours resulting from the honeycomb grid. In the following analyses tissue border cells are ignored, since these are an artificial artifact, and there are no real biological data or observations for these.

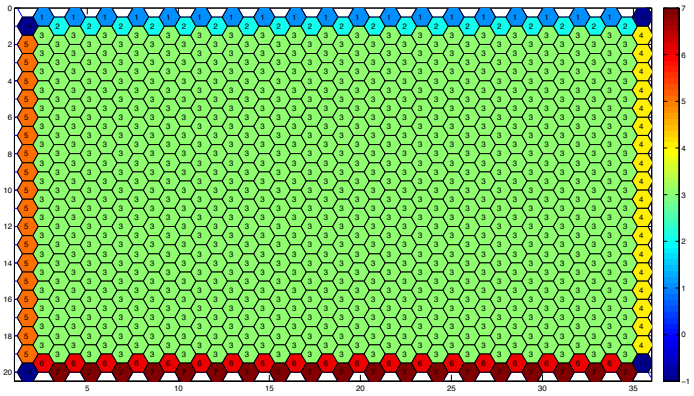

Figure 5.5: Clustering for continuous simulation of unbiased and BFXWt models.

Model checking, primary data. The unbiased model is constructed to reproduce the lack of polarisation, causing proteins to be symmetrically distributed along the cell membrane. This is confirmed by the results, see Figure 5.6 which shows identical time-series plots for all six membrane compartments.

For this, the following query was formed, where P_1, P_2, P_3, D_1, D_2 and D_3 denoted the concentration of FFD in the proximal and distal compartments.

$$P_{=?} \quad [G(P_1 = P_2 \pm \delta \wedge P_1 = P_3 \pm \delta \wedge P_1 = D_1 \pm \delta \\ \wedge P_1 = D_2 \pm \delta \wedge P_1 = D_3 \pm \delta)]$$

Note that $P_1 = P_2 \pm \delta$ is a short-hand for $P_1 < P_2 + \delta \wedge P_1 > P_2 - \delta$. A small value $\delta = 0.003$ was used because the query was dealing with real numbers, and there could be slight variations between traces due to, for example, effects in the solver. This query holds for all non-boundary cells.

Model checking, secondary data. Because the wild type cells (except boundary cells) exhibit the same time-series behaviour in all compartments, the accumulation of FFD is

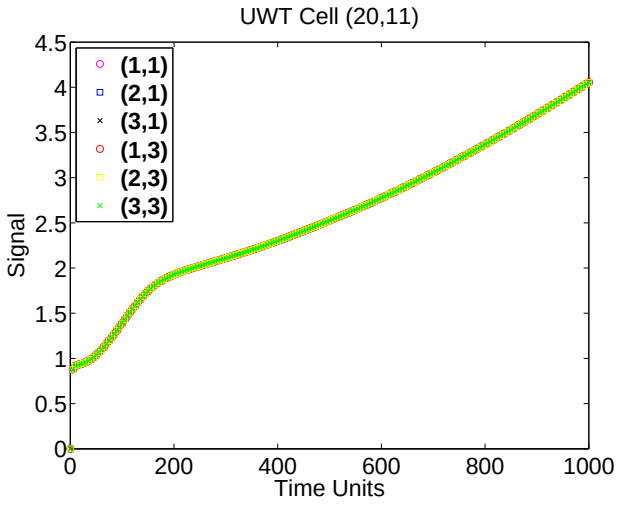

Figure 5.6: Unbiased model, continuous simulation for a representative of cluster 3 (see Figure 5.5).

identical in all membrane compartments (see Figure 5.7).

The same query as above was used, but set $\delta = 0.04$ due to the cumulative nature of variations; again this query holds for all non-boundary cells.

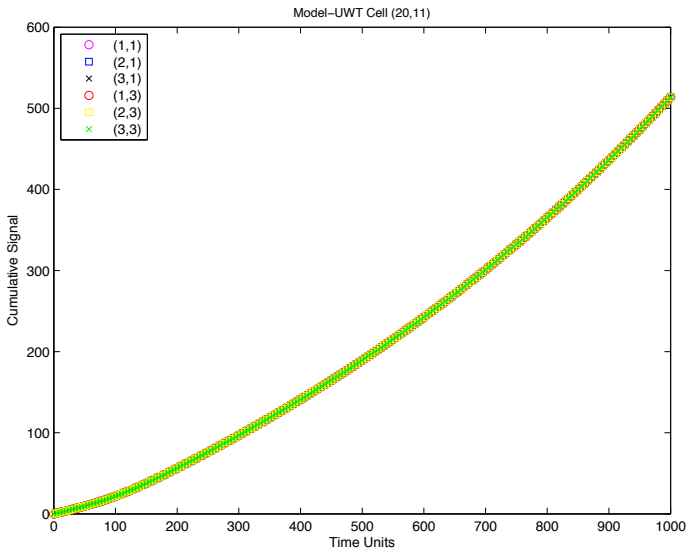

Figure 5.7: Unbiased model, cumulative signal for a representative cell, relating to Figure 5.6.

Experiment 2: Biased model with feedback loop and factor X, wild-type (BFXWt)

Clustering. The feedback loop mediates the competition between proximal and distal proteins between adjacent surfaces of neighbouring cells and amplifies the asymmetric localisation of these proteins. In wild-type cells, Dsh and Fz localise on the distal membrane, while Vang and Pk accumulate to the proximal membrane as seen in vivo [Axel 01, Bast 03, Stru 01, Tree 02a]. The results obtained by clustering analysis for this model exhibit similar characteristics to the wild-type unbiased model (Experiment 1, Figure 5.5).

Model checking, primary data. The time-series behaviour for a representative cell is shown in Figure 5.8. The middle proximal and the distal compartments were chosen for comparison because they exhibited the largest differences in proximal-distal intra-cellular behaviour. At the initial time point all concentrations are zero; after this the concentration

of FFD in the middle distal compartment ((2,3) in Figure 5.8) represented by D_2 is always the highest of all the compartments, and that in the middle proximal compartment ((2,1) in Figure 5.8) represented by P_2 is always the lowest. The upper and lower distal compartments are equal, lower than the middle distal and higher than the corresponding upper and lower proximal compartments.

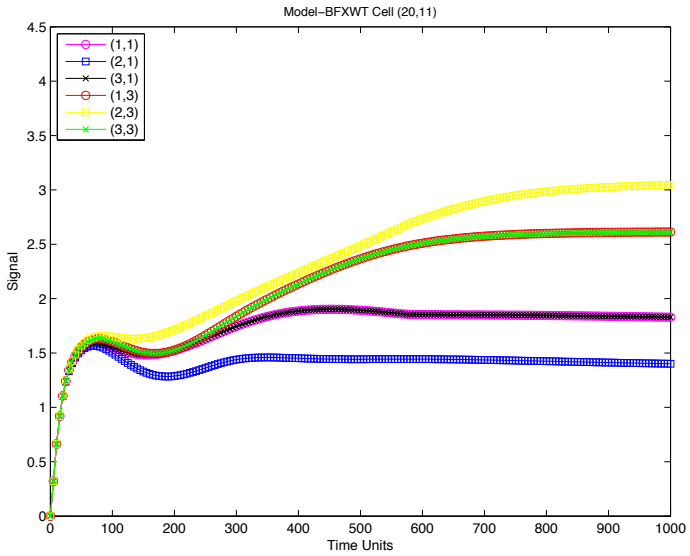

Figure 5.8: BFXWt model, continuous simulation for a representative of cluster 3 (see Figure 5.5).

The following query was constructed to reflect the observation above, where $\delta = 0.002$:

$$P_{=?} \quad [time > 0 \rightarrow G(D_2 > D_1 \wedge D_1 = D_3 \pm \delta \wedge$$
$$D_1 > P_1 \wedge P_1 = P_3 \pm \delta \wedge P_1 > P_2)]$$

It has been observed that all traces exhibit one peak followed by a trough, and formulate a query for compartment D_2 as an example, using the differential function:

$$P_{=?}[F(d(D_2) > 0 \wedge F(d(D_2) < 0 \wedge F(d(D_2) > 0)))]$$

The queries were then extended in the obvious manner to show that the proximal compartments have two peaks whilst the distal compartments have one peak in the traces

over the time period shown here.

$$P_{=?}[F(d[P_2] > 0 \wedge F(d[P_2] < 0 \wedge F(d[P_2] > 0) \wedge F(d[P_2] < 0 \wedge F(d[P_2] > 0)))))]$$

Model checking, secondary data. It has been observed that the middle distal compartment always has a higher cumulative time series than the other compartments, whereas the middle proximal compartment is consistently the lowest. The upper and lower distal compartments are very similar, less than the middle distal and higher than the corresponding upper and lower proximal compartments (see Figure 5.9).

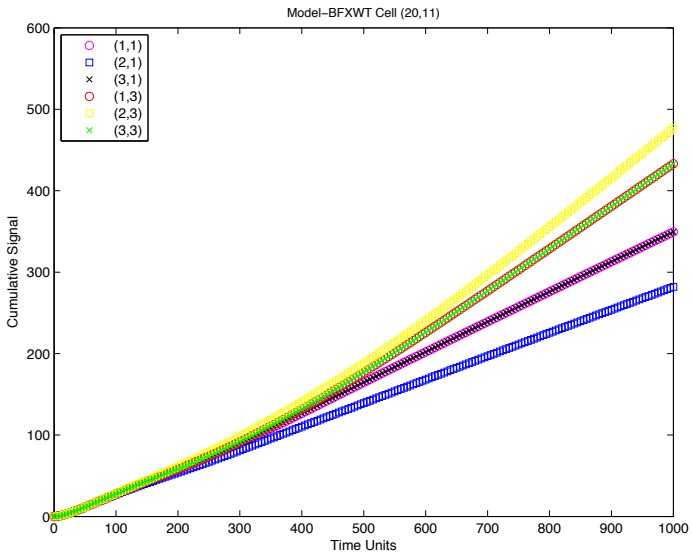

Figure 5.9: BFXWt model, cumulative signal for a representative cell, relating to Figure 5.8.

The cumulative variables for the membrane compartments $P_1, \ldots, D_3$ by $CP_1, \ldots, CD_3$ were denoted. The equivalent query to that which was formulated for the time series traces holds for the cumulative signal, again with $\delta = 0.002$:

$$\begin{aligned} P_{=?} \quad & [time > 0 \rightarrow \\ & G(CD_2 > CD_1 \wedge CD_1 = CD_3 \pm \delta \wedge \\ & CD_1 > CP_1 \wedge CP_1 = CP_3 \wedge CP_1 > CP_2)] \end{aligned}$$

Free variables in MC2 [Dona 08] were used to compute the range over the difference

in the maximum and minimum cumulative signals:

$P_{=?}$ [$\$\gamma = max(CD_2) - max(CP_2)$]

In this case $\gamma = 194$, indicating the clear signal in the Wt for the formation of the hairs in the central part of the distal edge of the cell.

Experiment 3: Biased models with with feedback loop, factor X and mutated clones (BFXFz and BFXVang)

This experiment presents the results of both *fz*- and *vang*- mutant clones. As reported in [Gubb 82, Vins 87], cells in a *fz*- clone have incorrect polarity and occasional multiple hairs. Wild-type cells distal, but not proximal to the clone have incorrect polarity, pointing more proximally towards the clone [Gubb 82, Vins 87], see Figure 3.2. Regarding the capability of the current model, it is expected that: (i) cells in the clone have incorrect polarity (FFD does not form); (ii) wild-type cells distal to the clone have FFD accumulated at the proximal rather than distal edge of cell.

As reported in literature [Tayl 98, Ma 03], cells in a Vang- clone have proximal domineering nonautonomy characteristic, the wild-type neighbouring cells proximally next to the clone point away from the clone rather than towards the clone.

Clustering: *fz*- mutant clone. The result demonstrates the impact of a clone of *fz*- mutant cells on the neighbouring wild-type cells surrounding it, see Figure 5.2 (c), where all of the mutant cells are in one cluster, and disruption was detected in all the wild-type cells directly adjacent to the clone. In more detail, different clusters are detected, and hence different responses in the wild-type cells to each of the six sides of the clone. Note that each of the six cells at the vertices of the clone were detected as individual outliers by the clustering technique. Because Fz is knocked out in all mutant cells, the supply of Fz to

form FFD is completely cut off, leading to zero FFD in the mutant cells. In addition, FFD could be affected by adjacent cells due to intercellular communication. These factors have an influence on the clusters that are detected.

Clustering Vang- mutant clone. The result shows the impact of a clone of Vang- mutant cells on their neighbouring cells around it (see Figure 5.10). In this case the central cells in the clone is in one cluster, while the mutant cells at the edges of the clone are in other clusters – one cluster at each of the six edges. Disruption has been detected in all the wild-type cells directly adjacent to the clone, resulting in several clusters, one at each edge of the clone. Only cells at four of the vertices of the clone are detected as individual outliers by the clustering technique. The lack of Vang in the central mutant cells indirectly affects the formation of FFD in adjacent cells due to intercellular communication. Thus while the central Vang- clone cells has zero FFD due to the lack of Vang in their surrounding cells, mutant cells at the boundary of the clone have some FFD because they maintain some communication with adjacent wild type cells. Proximal domineering nonautonomy was reported for Vang- clone[Tayl 98], cells proximal but distal to the clone point away from the clone, while cells distal to the clone somehow behave as the wild-type cells (pointing distally). Thus cells distal and proximal to the clone are expected to belong to different clusters, which is confirmed by the clustering results. Cells in the distal or proximal edge of the clone is assigned to three clusters, respectively. Because cells at the same edge to the clone have similar overall behaviour pointing away from the clone, however the angle of the hair in each cell may be slightly different from each other due to different communication patterns with its neighbouring cells.

Model checking, primary data, *fz*- clone. In the following, only the middle distal and middle proximal compartments, D_2 and P_2, are considered because they exhibit the highest

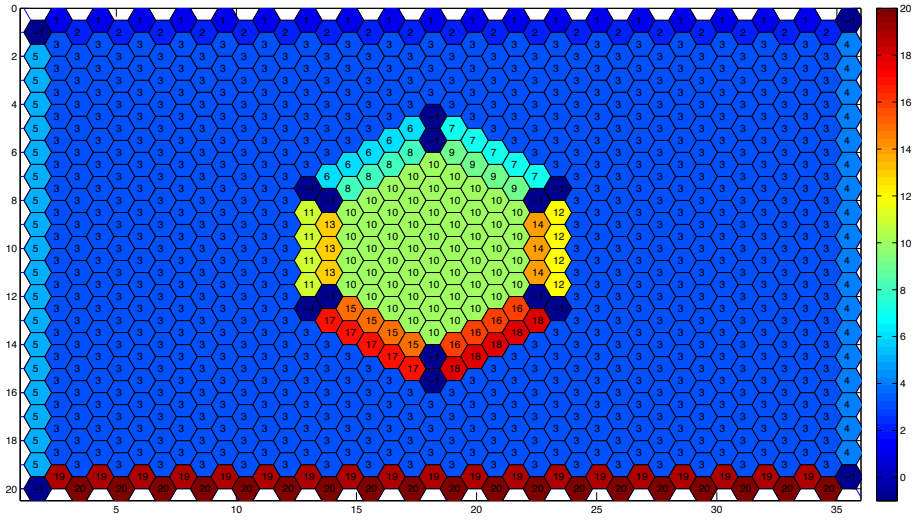

Figure 5.10: Clustering for continuous simulation of model BFXVang.

PCP signal.

Unlike in the wild-type cells, for the cells distally neighbouring to the clone the concentration of FFD in the middle distal compartment is always lower than that of the middle proximal compartment, see Figure 5.11:

$P_{=?}\ [time > 0 \rightarrow G(D_2 < P_2)]$

Moreover, the trace of D_2 exhibit a peak followed by a trough, which is not true for P_2:

$P_{=?}[F(d(D_2) > 0 \wedge F(d(D_2) < 0 \wedge F(d(D_2) > 0)))]$

However the cells immediately proximal to the clone exhibit behaviour similar to wild-type cells, i.e. the middle distal compartment is always higher than the middle proximal compartment (see Figure 5.12). The concentration of FFD in all the clone cells is zero at all time points.

Model checking, primary data, *vang*- clone. All the non-boundary cells within the clone have zero FFD at all time points, however the clone cells directly adjacent to the clone boundaries do exhibit some accumulation of FFD in those compartments directly abutting

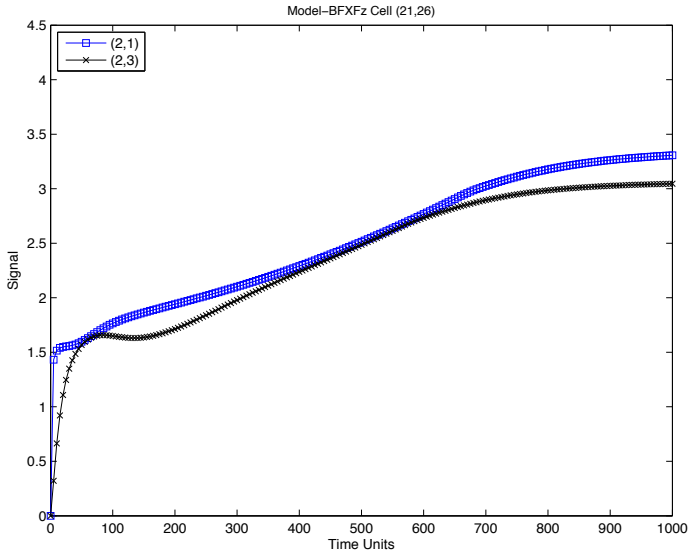

Figure 5.11: BFXFz model, continuous simulation for a representative of cells distally adjacent to the *fz*- clone (cluster 10, coloured in ECE amber see Figure 5.2 (c)).

wild-type cells; this was to be expected because of the formation of FFD could take place with participation of Vang from the adjacent wild-type cells.

As in the wild-type cells, the concentration of FFD in the middle distal compartment of the cells distally neighbouring to the clone is always higher than that of the middle proximal compartment, but this latter is always zero. The trace for the distal compartment exhibit a peak and a trough, see Figure 5.13.

$$P_{=?}[(time > 0 \rightarrow G(D_2 > P_2)) \wedge G(P_2 = 0)]$$

$$P_{=?}\ [F(d(D_2) > 0 \wedge F(d(D_2) < 0 \wedge F(d(D_2) > 0)))]$$

The behaviour of the cells immediately proximal to the clone is similar to wild-type cells, in that middle proximal compartment is always higher than the middle distal compartment, while the latter is always zero (see Figure 5.14).

$$P_{=?}[(time > 0 \rightarrow G(P_2 > D_2)) \wedge G(D_2 = 0)]$$

$$P_{=?}\ [F(d(P_2) > 0 \wedge F(d(P_2) < 0 \wedge F(d(P_2) > 0)))]$$

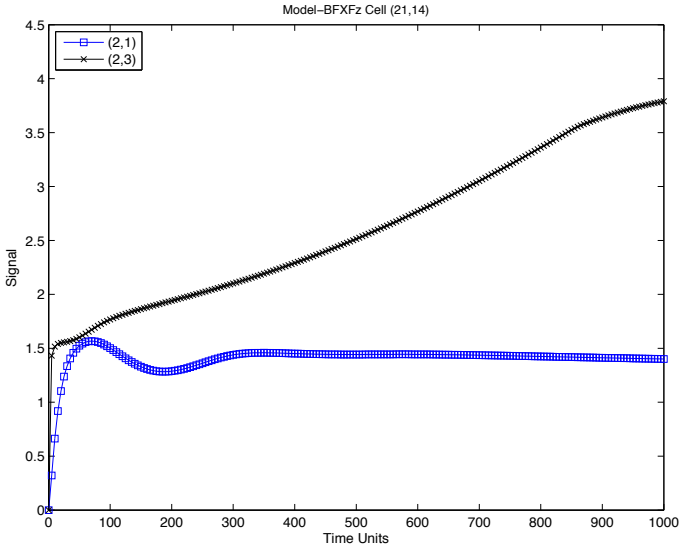

Figure 5.12: BFXFz model, continuous simulation for a representative of cells proximally adjacent to the *fz*- clone (cluster 9, coloured in amber see Figure 5.2 (c)).

Model checking, secondary data, *fz*- clone. All biased models with the *fz*- clone (BFXFz, BFTFz and BFXTFz) always exhibit a relatively higher cumulative signal in the middle proximal compartment compared to the middle distal compartment in those cells distally directly next to the *fz*- clone (see Figure 5.15 for model BFXFz):

$P_{=?}\ [time > 0 \rightarrow G(CD_2 > CP_2)]$

Run the following query to check wild-type cells not adjacent to the clone:

$P_{=?}\ [\$\gamma = max(CD_2) - max(CP_2)]$

In this case, the range over the difference in the maximum and minimum cumulative signals $\gamma = 72$, indicates a severe loss of polarising signal compared to BFXWt model.

The wild type cells in the tissue (i.e. away from the clone area) display behaviour similar to the wild type cells in the BFXWt model (see Figure 5.9).

$P_{=?}[time > \epsilon \rightarrow G(CD_2 > CD_1 \wedge CD_1 = CD_3 \pm \delta \wedge$

$CD_1 > CP_1 \wedge CP_1 = CP_3 \wedge CP_1 > CP_2)]$

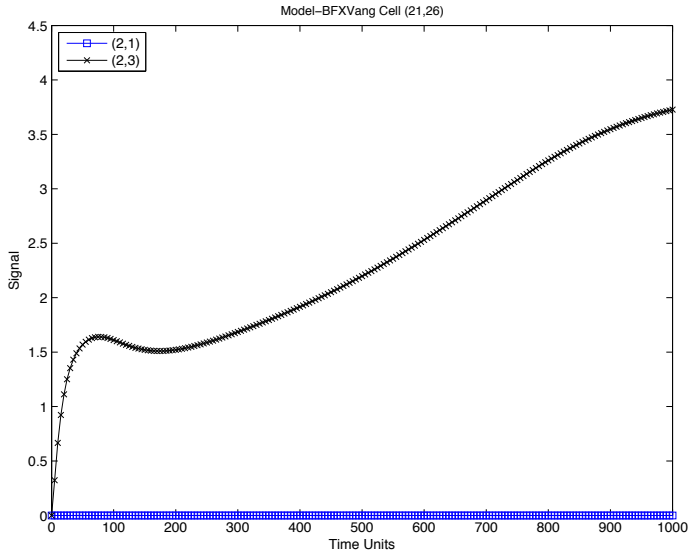

Figure 5.13: BFXVang model, continuous simulation for a representative of cells distally adjacent to the *vang-* clone (cluster 12, see Figure 5.10).

where $\epsilon = 50$ and $\delta = 0.2$

The behaviour of the cells proximally next to the *fz-* clone (see Figure 5.16) is similar to that of the wild-type cells in the model. All clone cells within the *fz-* clone have zero accumulation of FFD in all compartments because there is no supply of Fz to form the FFD complex.

Model checking, secondary data, *vang-* clone. In contrast to the *fz-* clone, for all Vang- clone models (BFXVang, BFTVang and BFXTVang) the cumulative signal in the proximal cells next to the clone is always higher in the middle proximal compartment compared to the middle distal compartment, and the latter is always zero (see Figure 5.18 for the BFXVang model):

$P_{=?}[(time > 0 \rightarrow G(CP_2 > CD_2)) \wedge G(CD_2 = 0)]$

In this case the range over the difference in the maximum and minimum cumulative signals between the compartments is given by $\gamma = 421$ over compartments CP_1 and CP_3 as

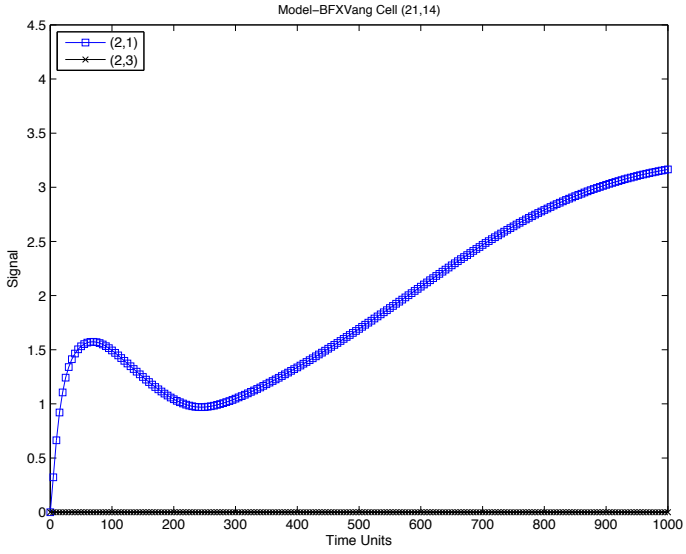

Figure 5.14: BFXVang model, continuous simulation for a representative of cells proximally adjacent to the *vang-* clone (cluster 11, see Figure 5.10).

the maxima and CD_2 as the minima indicating strong polarisation in the opposite direction to the Wt cells.

The behaviour of the distal cells next to the clone (see Figure 5.17) exhibit a similar trend to that of the wild-type cells; however the cumulative value for FFD in the middle proximal compartment is always zero, and the values in the upper and lower proximal compartments are lower than in the wild-type cells. This is because the knockout of Vang in the adjacent cells in the clone causes the lack of FFD in the distally neighbouring cells.

$$P_{=?}[(time > 0 \rightarrow G(CD_2 > CP_2)) \wedge G(CP_2 = 0)]$$

$$P_{=?}\ [\$\gamma = max(CD_2) - max(CP_2)]$$

In this case $\gamma = 472$, indicating very strong polarisation in the normal (Wt) direction.

The findings regarding the analysis of the secondary data are consistent with observations reported in the literature, for example [Gubb 82, Vins 87, Axel 11], namely that PCP disturbance was reported in cells distal to *fz-* clones, whereas the disturbance

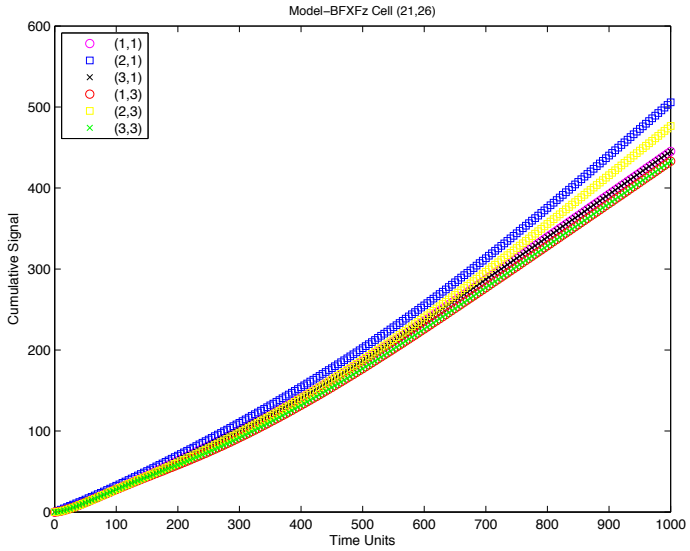

Figure 5.15: BFXFz model, cumulative signal for a representative of cells distally adjacent to the *fz*- clone, relating to Figure 5.11.

is on the proximal side of Vang- clones. The cause of this disturbance is indicated by the cumulative plots for these cells, where there is a lack of an orienting signal in the distal neighbours to the *fz*- clone (see Figure 5.15), and the near-zero values in the distal compartments for the proximal neighbours to the Vang- clone (see Figure 5.18).

Other clones, e.g. *fmi*, *dsh*, *pk*, were also modelled using *BFX* model. Simulated the clone of cells lacking *dsh* function results in the disruption of polarity within the mutant cells, but only shows a mild effect on cells outside of the clone. The nearly, though not fully cell autonomous phenotype is similar to that which was observed experimentally [Adle 04], despite allusions in the literature to a fully cell autonomous phenotype for *dsh*. Clones of cells lacking all *pk* function show only a very subtle phenotype, consistent with the data obtained from this work and published reports ([Adle 01]).

Table 5.1 summarises simulation results of distributions of all core proteins using BFX model, and the results are highly concordant with the available published experimental

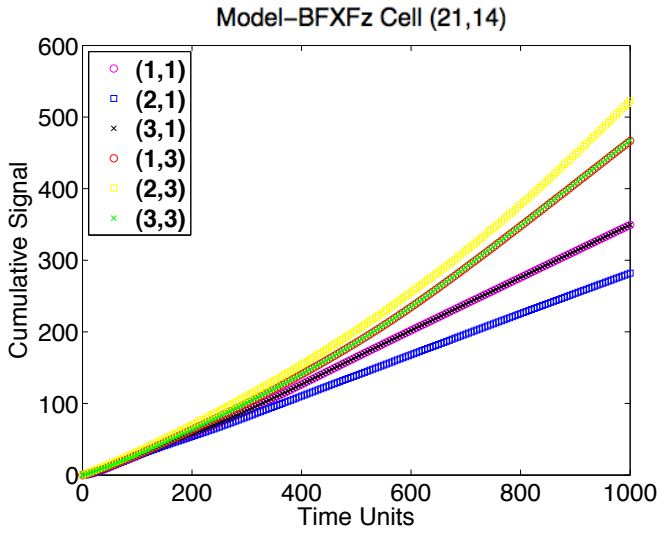

Figure 5.16: BFXFz model, cumulative signal for a representative of cells proximally adjacent to the *fz-* clone, relating to Figure 5.12.

observations.

Clone	Simulation results	Experimental observations
fmi-	No Polarity in both clone and neighbouring wild-type cells	Ref [Amon 05, Usui 99]
fz-	Distal domineering nonautonomy	Ref [Gubb 82]
vang-	Proximal domineering nonautonomy	Ref [Tayl 98]
dsh-	Polarity disruption inside of the mutant clone	Ref [Kili 94, Thes 94]
pk-	No polarity reversal	Ref [Adle 01] [Amon 05]
pk+	Distal domineering nonautonomy	Ref [Bast 03]

Table 5.1: Comparison of simulated and experimental protein localisation. For loss of function and overexpression clones obtained from BFX model. By systematically comparing the simulations to available experimental results, in nearly all cases, the simulations closely approximate the experimental results.

Remarks Other HCPN models were also analysed following the same approach as demonstrated above in this section. The findings suggest that this approach can satisfy the need to analyse and check multi-scale models in both human and machine compatible ways. In addition, by analysing the simulation results, BFX model is able to provide the most

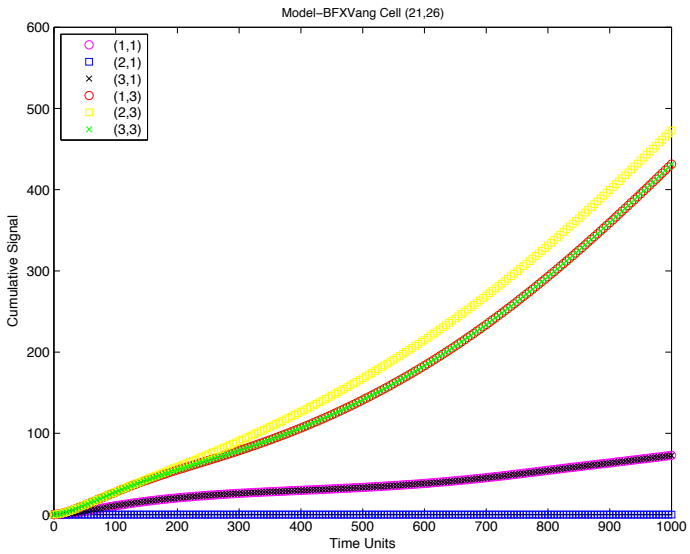

Figure 5.17: BFXVang model, cumulative signal for a representative of cells distally adjacent to the *vang*- clone, relating to Figure 5.13.

promising description of PCP by reproducing nearly all biological observations reported in literature compared to other models. Therefore, it was used as the example to present the analysis and model checking approach above.

Some of the logic statements can be used as classifier over different clusters, for example, the query describes that the concentration of FFD is always zero:

$$P_{=?}[G(P_1 = 0) \wedge G(P_2 = 0) \wedge G(P_3 = 0) \wedge G(D_1 = 0) \wedge G(D_2 = 0) \wedge G(D_3 = 0)]$$

In the case of the *fz*- clone, MC2 returns P=1 (true) only to cells in the clone, while P=0 (false) for other cells. This query thus can be considered as a classifier to characterise a particular kind of behaviour.

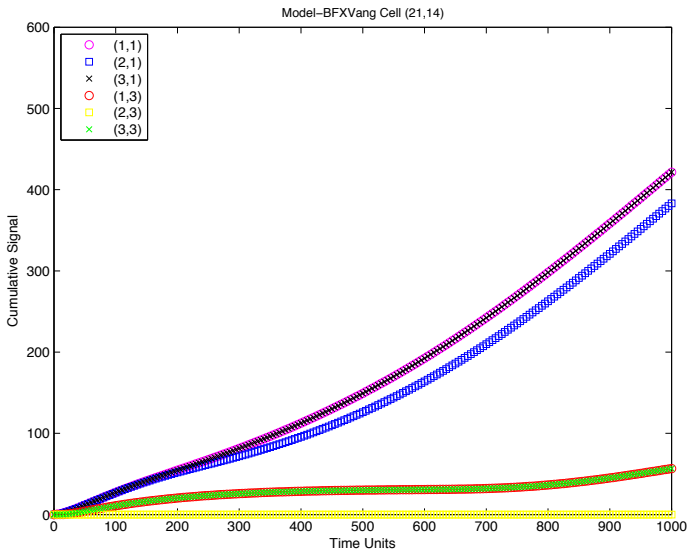

Figure 5.18: BFXVang model, cumulative signal for a representative of cells proximally adjacent to the *vang*- clone, relating to Figure 5.14.

5.7 Discussion

This section discusses insights into domineering nonautonomy by comparing the family of HCPN models constructed based on different hypotheses, and subsequent experiments supporting model predictions.

5.7.1 Insight into domineering nonautonomy

The analysis of each HCPN model provides a clear view of the model itself, whether it can reproduce characteristic PCP phenotype and how much it can recapitulate those biological observations. However, in order to understand the possible mechanisms of PCP signalling, it is necessary to compare different models to have insights into nonautonomy. Therefore, the family of HCPN models were studied by comparing their effects on the strength of the signalling.

In order to discover model(s) which would be able to capture most biological observations, all models were used to create an in-silico tissue of 800 cells imposed in a 40 by 40 rectangular grid, where introducing 80 mutant cells lacking of *fz*, respectively. After simulating each model, a score S_{fz} (the penalty defined in Section 4.5.3.2 in Chapter 4) for each model was calculated. As reported in literature, PCP disturbance occurs in cells distal to *fz*- clones, indicating that FFD in the distal membrane should be higher than that in the proximal membrane in those cells. Thus it is expected that S_{fz} would be greater than zero in cells distally neighbouring to the *fz*- clone, while smaller than zero in wild-type cells far away from the clone. The results show that simulations obtained from BFXFz model is most consistent with biological observations (see Figure 5.19), and show the strongest polarising signal among those models which can correctly reproduce the PCP phenotype. This is the reason that BFX model was used as a representative to illustrate the analysis approach in this chapter. This result somehow support the point that the feedback loop acts to both amplify the global directional cue and relay the signal to neighboring cells [Axel 01, Axel 11, Tree 02a]. More results can be found in Appendix B.

5.7.2 pk-engrailed overexpression pattern

When originally formulating the feedback loop models, it was found that *pk* overexpression in the posterior wing domain enhanced the accumulation of FFD at cell boundaries, and it was proposed that this reflected an enhancement of accumulation at the opposing cell surface through the feedback mechanism [Tree 02a]. However, that the mechanism would show this property was not obvious. To examine this more rigorously, this experiment was modelled in a 15 by 15 array comprising 112 cells with a clone of mutant cells variably overexpressing Pk in a band of 15 cells. Consistent with experiments overexpressing *pk*

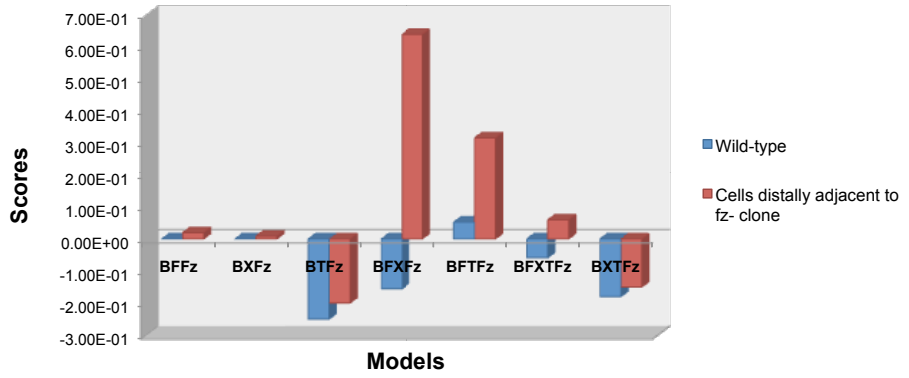

Figure 5.19: Score S_{fz} for representative cells in models with *fz*- clone. As reported in literature [Gubb 82, Vins 87], the *fz*- clone produces a distal domineering nonautonomy. In the model, this can be translated as follows: (i) $S_{fz} > 0$ in cells distally next to the *fz- clone*; (ii) $S_{fz} < 0$ in other wild-type cells in the background.

in the engrailed (en) domain in the posterior of the wing ([Tree 02a]), FFD are seen to accumulate to higher levels in the region overexpressing *pk* than in the wild-type region (use middle distal compartment D_2 as representation, see Figure 5.20). Therefore, despite the cell autonomous inhibition of FFD complex formation, the feedback loop causes the system to respond to excess Pk by maintaining high levels of membrane bound Dsh, Fz, and Pk.

5.7.3 Predicting additional phenotypes

Until now, results haven been shown reproducing known characteristic PCP phenotypes. It is not expected that the models would necessarily reproduce additional mutant conditions without additional training of the model parameters. However, the model could be expected to predict combinations of the conditions that were used in selecting the model parameters. Three such mutant conditions were observed in the *Drosophila* abdomen [Lawr 04, Lawr 10] (see Table 5.2). The models were developed and trained on the

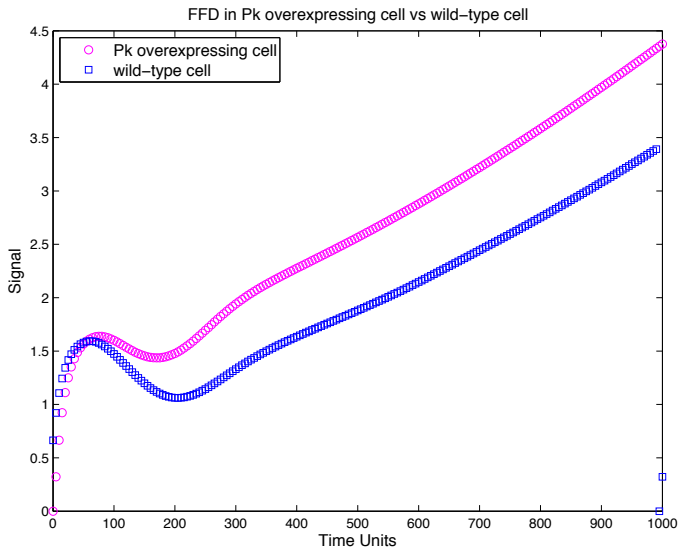

Figure 5.20: BFXPk+ model, continuous simulation for middle distal compartment (D_2) in two representative cells: (i) cells with overexpressing Pk (labelled in pink); (ii) wild-type cells in background (labelled in blue).

- *fz*- clone in a pk^{null} background
- overexpressing *fz* and *vang*- clone
- overexpressing *fz* clone in a *fz*- background

Table 5.2: Three mutant conditions observed in the *Drosophila* abdomen

Drosophila wing, but the polarity of hairs on the fly abdomen behave similarly to those on the wing, and the same core polarity signalling components have been implicated in establishing planar polarity in the abdomen. These conditions were simulated using the same parameter set previously identified, and it was found that the simulation results corresponded well with the published results. Note all simulations were run over an in-silico tissue based on a honeycomb grid imposed on an underlying 30×30 rectangular grid comprising 450 cells, with mutant clone(s) in a circle-like shape with 30 mutant cells.

When *fz*- clones appear in a pk^{null} background (no Pk protein exists in the tissue), the simulation predicts that cells surrounding the clone repolarise to point toward the

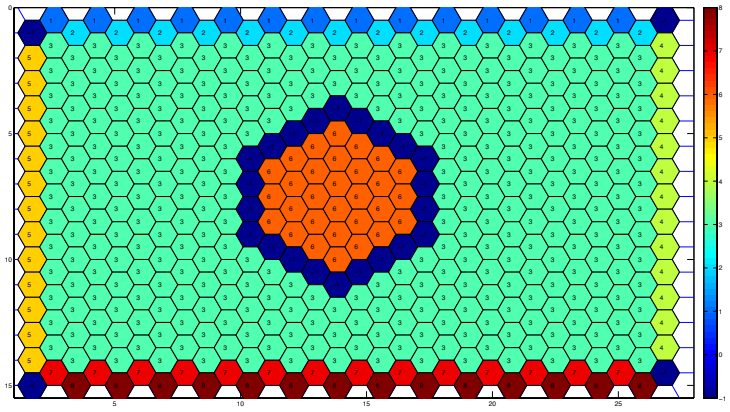

Figure 5.21: Clustering for continuous simulation of BFX model: *fz*- clone in a pk^{null} background. Cells surrounding the clone repolarise to point towards the *fz*- clone, and pk^{null} background cells show very little asymmetry. Compared with Ref. [Lawr 04].

fz- clone (see Figure 5.21). The remaining pk^{null} background cells show very little asymmetry, and variation in the direction of these hairs would be predicted for an actual wing. This is consistent with data from the abdomen, where nonautonomous repolarisation of surrounding pk null background cells was observed [Lawr 04, Lawr 10]. This serves as another example where the feedback loop mechanism can lead to nonautonmous behavior despite the global absence of one of the feedback loop signalling components, pk.

Simulations of *fz* overexpression (Fz+) in a *vang*- clone (see Figure 5.22) look similar to results obtained when simulating Vang- clones, resulting in a nonautonomous polarity reversal in wild-type cells proximal to the clone. In the interior of the clone, polarity is reduced, predicting a randomised hair pattern in this region. These predictions are consistent with results described for the abdomen [Lawr 04, Lawr 10], although they do not observe the ordered polarity of cells within the clone that touch the clone border as predicted by the simulation. A possible explanation for this discrepancy, consistent with

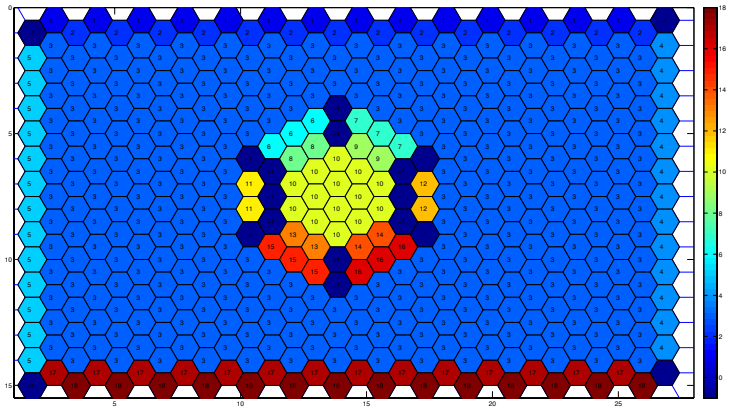

Figure 5.22: Clustering for continuous simulation, BFXFz+Vang- model (a mutant clone comprising of *fz*+ and *vang*-). The simulations look similar to results obtained when simulating *vang*- clones, resulting in a nonautonomous polarity reversal in wild-type cells proximal to the clone. In the interior of the clone, polarity is reduced similar to experiment of Vang- clone (see Figure 5.10). Described but not shown in Ref. [Lawr 04], also described in Ref. [Lawr 10].

observations from the following simulation case, is that final distributions of Vang or Pk may also play a role in directing the site of prehair initiation, and so the loss of Vang function inside the clone may disrupt polarity to a greater extent than is predicted by the simulation.

When fz is overexpressed in a clone of cells within a fz- background, the overexpressing cells are predicted to polarise (see Figure 5.23). The data from the abdomen show a single row of cells outside of the overexpressing cells that have also been repolarised to align with the fz overexpressing cells [Lawr 04]. The simulation results show a repolarisation of some of the proteins within cells adjacent to the fz overexpressing cells that is consistent with the abdomen data for these cells, but in the absence of Fz, there is no mechanism for the accumulation of Dsh to the membrane to in turn direct the predicted hair polarity of

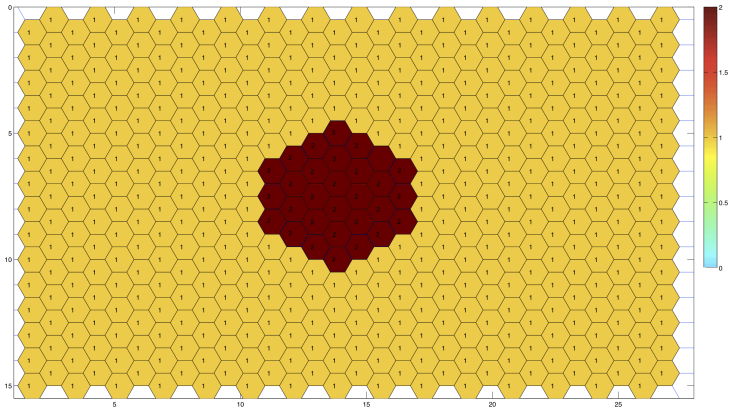

Figure 5.23: Clustering for continuous simulation of BFX model: *fz*+ clone in a *fz*- background. This shows that the overexpressing cells are polarised (labelled in red), while *fz*- background cells are not (labelled in yellow). Compare with Ref. [Lawr 04].

these cells (therefore the influence of neighbouring cells of the clone in the results obtained from clustering FFD can not be seen). This suggests that these fz- background cells retain some ability to interpret the asymmetric accumulation of Pk and Vang to direct the site of prehair initiation.

5.8 Summary

With the growth of various data resources and urgent need to scale up in systems biology, proper analysis and visualisation techniques for multi-scale modelling becomes more and more crucial. In this chapter, the approach comprising clustering with visualisation and model checking provides the feasibility to analyse and check the models by both human and machine. The HCPN models have been proven to reproduce well-studied PCP phenotypes, such as *fz*- or *vang*- clone in a wild-type background. The clustering method provides

a visual speculation of tissue-level behaviour using time-series data obtained from the intracellular level. The dynamic behaviour of each protein of interest is characterised by temporal logic statements, and verified by model checking. Some of the logic statements can be used as classifier over different clusters. Some insights into the mechanisms underlying PCP have been discussed with the results obtained from the HCPN models. Moreover, several predictions of combination of mutant clones reported in the *Drosophila* abdomen have been made using the BFX model, which would be testable in a biological laboratory. In addition, the findings have demonstrated that hierarchically as a modelling formalism permits easy construction of large scale (e.g. multi-scale) models, especially those with a hierarchical organisation, and allows comparison between model variants.

Chapter 6

Conclusions, future work and open problems

6.1 Conclusions

This chapter briefly summarises the efforts of developing and analysing multi-scale models, the key results from the multi-scale models of PCP, and points out some general directions for extending the model that may yield additional insight. Several open problems in multi-scale modelling of PCP signalling and other such systems are later described, along with some thoughts on how some of these problems may be addressed.

The complexity and diversity of biological phenomena, the range of spatial and temporal scales over which they act, and the intricate way in which they are interwoven, make practically unfeasible the understanding of living systems through intuition alone. Therefore, multi-scale modelling approaches are essential tools in the quest for a quantitative understanding of living systems. This thesis has presented the application of hierarchically coloured Petri net techniques to construct a family of multi-scale computational models in order to explore the mechanisms that drive Planar Cell Polarity in *Drosophila* wing tissue. Results have shown that the family of models recapitulate

signalling phenomena known to occur in wild-type and mutant clones. The approach has involved developing sophisticated patterns of communication over hierarchically organised components. The thesis demonstrated the power of the new concept of hierarchically coloured Petri nets (HCPN) introduced in [Gao 12] to model multi-scale systems, especially those with a hierarchy structure, and associated analysis approach (cluster analysis and simulative model checking of primary and secondary data) which has been developed in [Gao 11, Gao 12]. The computational experiments have been over very large underlying models each comprising over 150,000 ODEs. The issues highlighted in the efforts of modelling include

- Illustrating how colours can be used to encode both space and hierarchy
- Computational challenges due to the fact that currently simulations must be performed at the unfolded level rather than at the coloured level. The largest model that has been constructed so far consists of 800 cells, each with 7 virtual compartments, comprising 19 places and 23 transitions as a CPN model which when unfolded yields 164,0000 places (biochemical entities, each of which is described by an ODE) and around 210,000 transitions (reactions). Unfolding takes 2.5mins, while simulation takes up to several hours to run on a dual core computer (see Table 4.5).
- Fitting the multi-scale model to data, where problems are due to (i) the lack of reliable time-series data for concentrations of key biochemical entities because the time period being modelled is during the pupal stage when it is difficult to obtain such data; (ii) data of hair orientation is at the tissue level whereas the kinetic parameters being fitted are at the biochemical level in the virtual compartments; (iii) these data are at one time point (the adult state - i.e. the final state); (iv) the complexity of the

model as it takes hours to run each simulation of a large-size model, e.g. a model of 800 cells, whereas simulation is iteratively repeated in the fitting procedure.

Much work remains in order to have a full understanding of PCP signalling, but with the development of a family of multi-scale computational models of PCP, progress has been made in linking observations about individual signalling components and hypotheses about their interactions to the system behaviours that are ultimately of interest. Further development of these model can also continue to motivate, integrate, and test new experimental findings and hypotheses, making the multi-scale models an integral component in exploring the mechanisms that drive PCP signalling in *Drosophila* wings.

A family of multi-scale computational models of PCP in the *Drosophila* wing have been developed based on hypotheses derived from experimental results. First hypothesis, which relies on local feedback loop acting at the cell membrane, can adequately account for the autonomous and nonautonomous phenotypes of various mutant clones has been the source of some controversy in the field, and some have argued that it cannot account for some observed phenotypes [Adle 02, Stru 02d]. Second hypothesis, cells responses to an initial cue by producing and secreting a diffusible second factor (Factor X) whose graded distribution determines polarity [Adle 00, Adle 00, Fant 03, Lawr 99, Lawr 02, Stru 97a, Stru 02a, Wehr 98, Zeid 99]. The third is that the transport of core proteins, Fmi, Fz and Dsh, are oriented by controlling the polarity of microtubule cytoskeleton [Olgu 10]. By modelling the feedback loops, Factor X and polarised vesicle transport of the three core proteins, the family of multi-scale models were constructed using HCPN. BFX model was able to reproduce in simulation all of the most characteristic PCP phenotypes, demonstrating the feasibility of the underlying biological model as a PCP signalling mechanism. Simulation results from the models suggested it highly possible

that the early aspect of PCP establishment requires a global directional cue to orient cell polarity in which feedback loop directs Fz accumulation as amplifier, which yields results consistent with literature [Axel 11]. In addition, BFX model was also able to predict combinations of the conditions, in which simulation results corresponded well with the published results[Lawr 04, Lawr 10].

6.2 Future work

The multi-scale models presented thus far serve as a starting point, from which may be continuously used to refine the understanding of PCP signalling and incorporate additional knowledge to explain new experimental observations. Many signalling components known to be involved in PCP signalling have not been included in the multi-scale models, because their roles remains largely unclear. One natural extension to the models, however, would be the inclusion of Diego (Dgo) [Feig 01], which promotes the maintenance of apical localisation of Fmi [Das 04]. This function of Dgo is redundant with Pk and Strabismus (Stbm), and only appreciable in double mutant tissue [Das 04, Jenn 05]. The initial membrane association of Dgo depends on Fz, and Dgo physically interacts with Pk and Stbm through its Ankyrin repeats, while Dgo and Pk compete with one another for Dsh binding, thereby modulating Fz/Dsh activity and ensuring tight control over Frizzled PCP signalling [Jenn 05]. A mathematical model that includes Dgo may yield some insight into whether or not Diego is in a direction dependent fashion. This may discover how the initial localisation of PCP signalling proteins maintained, i.e. if Dgo stabilises the initial multiprotein complex, either a direct effect of Dgo and Pk or mediated through their effect on Fmi affects the localisation of Dsh. Another possible extension to the models would be the inclusion of downstream of core components. For example, the

small GTPase RhoA and its associated kinase Drok, known regulators of actin dynamics, contribute to organising distal actin assembly and prehair formation [Stru 97b, Wint 01]. The FH3-domain protein Multiple Wing Hairs and the membrane proteins Inturned (In) and Fuzzy (Fy) act cell autonomously on the proximal side and appear to be involved in restricting the site of wing hair initiation [Park 96, Coll 97, Stru 08]. This may give more insight into the downstream control of PCP signalling to direct wing hair formation.

Parameter optimisation for the ODE model was performed using essentially small regular hexagonal arrays. Better agreement between simulation results from the ODE model and experimental data may be achieved by instead measuring and fitting on larger cell geometries and hair angles from clones of the mutant conditions listed in Table 5.1. This is especially true for mutant conditions that result in domineering nonautonomy, where the extent to which wild-type cells near a clone are repolarised can be used to train the mathematical models to achieve the same degree of nonautonomy. The parameters selected thus far have not used this information, and so they currently only qualitatively reproduce PCP phenotypes. The collection of time series data, such as quantitative measurements of protein localisation at various stages during PCP signalling, may be used to better fit the dynamic behavior of the mathematical models. Currently, only the final protein localization state is used, and so there is no basis for training and verifying the dynamic behavior of the mathematical model. Temperature-sensitive alleles of fz may also permit experiments in which the effect of turning the signalling on and off at various times during the development of the wing can be measured. Simulating the same conditions may also permit to train and validate the dynamic behavior of the mathematical models.

As experimental methods continue to be developed and refined, it may soon be practical to continuously capture images of a developing wing. By extending the model to allow for

simulations on time-evolving cell geometries, the mathematical models may be able to account for the time and geometry dependence of the feedback loop.

6.3 Open problems

The relationship between the feedback loop and the remaining tiers of the signalling hierarchy is not well understood and represents a challenge that multi-scale modeling may help address. For instance, only the upstream asymmetry input signal was included in this thesis. Therefore, an important direction for future work would be to extend the multi-scale models to explicitly include these upstream signalling components and further elucidate their precise roles in PCP signalling.

Little is also known about the signalling components that convert the asymmetric localisation of the core polarity proteins at the end of PCP signalling into the site of prehair initiation, e.g. extending the models to include downstream signalling components to drive additional experiments to better understand how the core polarity genes interact with downstream signalling components to direct hair polarity.

An increasingly important and open problem in systems biology is the identification of mathematical model parameters that describe robust biological systems. The parameter optimisation method that has used are concerned primarily with demonstrating the feasibility of the mathematical model to reproduce some observed behaviour of the biological system. Only after a parameter set has been identified can the resulting model be verified if it is not overly sensitive to parameter variations and reflects a plausible representation of the system. It would be interesting to incorporate the knowledge of the robustness properties of these systems into the parameter optimisation process itself, so that the resulting parameter set yields a more valid representation of the system, and so

that the computational effort of verifying the robustness of the mathematical model can be combined with the computational effort of parameter optimisation. Given the great uncertainty that often characterises models of complex biological systems, understanding and measuring the robustness properties of a system may give some indication of the validity of a model.

In search for a proper modelling formalism perhaps the most important aspect to consider is the balance between simplicity and expressiveness. There is a price to pay for the amount of features provided by a formalism, which may come at the cost of increased model complexity. The complexity of the representation and the number of parameters determines the amount of experimental data required for model construction. This concern is most critical when not only the parameters but also the network structure are unknown. This is even more critical for selecting an appropriate multi-scale modelling formalism as the natural growth of complexity in such a biological system. In addition, availability of suitable modelling frameworks and simulation environment are likely be one of the most important factor in the wider adoption of multi-scale modeling in systems biology. Most of the present modelling and simulation tools are developed for single scale modelling and alone not sufficient for multi-scale modelling. Some may still play major roles in the new emerging frameworks of multi-scale modelling as simulation of multi-scale models will probably involve assembly of different simulators in an integrative manner, each individual being used in its own area of specialisation, e.g. linking different formalisms to create integrative one model.

Similarly searching for suitable approaches to analyse and check multi-scale models is very challenging due to: (i) the output of such biological systems is semi-quantitative experimental data; (ii) the level of behaviour which can be observed in the model is

different from the level which can be observed in experimental data; (iii) the need to consider the physiological geometries (one, two or three dimensions) as well as pattern formation.

HCPN is a generic modelling formalism imposing a hierarchical structure over colours to reflect the hierarchy inherent in a system being modelled. This hierarchically structured colouring allows locality to be encoded in colours, for example, a tuple (x,y) or (x,y,z) representing two or three dimensions geometry. It can also be employed to model multi-scale systems (Chapter 4 gives a detailed example, also see [Gao 12]), where a hierarchy is encoded as (nested) tuples of tuples, e.g. $((x_1, y_1), (x_2, y_2))$ representing a 2-dimensional grid (x_2, y_2) within a larger 2-dimensional grid (x_1, y_1).

Bibliography

[Adle 00] P. N. Adler, J. Taylor, and J. Charlton. "The domineering nonautonomy of *frizzled* and *Van Gogh* clones in the *Drosophila* wing is a consequence of a disruption in local signaling". *Machanisms of Development*, Vol. 96, No. 2, pp. 197–207, September 2000.

[Adle 01] P. N. Adler and H. Lee. "Frizzled signaling and cell-cell interactions in planar polarity". *Current Opinion in Cell Biology*, Vol. 13, No. 5, pp. 635–640, October 2001.

[Adle 02] P. N. Adler. "Planar signaling and morphogenesis in *Drosophila*". *Developmental Cell*, Vol. 2, No. 5, pp. 525–535, May 2002.

[Adle 04] P. N. Adler, C. Zhu, and D. Stone. "Inturned localizes to the proximal side of wing cells under the instruction of upstream planar polarity protein.". *Current Biology*, Vol. 14, pp. 2046–2051, 2004.

[Adle 94] P. N. Adler, J. Charlton, K. H. Jones, and J. Liu. "The cold-sensitive period for *frizzled* in the development of wing hair polarity ends prior to the start of hair morphogenesis". *Mechanisms of Development*, Vol. 46, No. 2, pp. 101–107, May 1994.

[Adle 98] P. N. Adler, J. Charlton, and J. Liu. "Mutations in the cadherin superfamily member gene *dachsous* cause a tissue polarity phenotype by altering *frizzled* signaling". *Development*, Vol. 125, No. 5, pp. 959–968, March 1998.

[Alde 97] P. N. Alder, R. E. Krasnow, and J. Liu. "Tissue polarity points from cells that have higher Frizzled levles towards cell that have lower Frizzled levles". *Current Biology*, Vol. 7, No. 12, pp. 940–949, December 1997.

[Amon 05] K. Amonlirdviman, N. A. Khare, D. R. P. Tree, W. S. Chen, J. D. Axelrod, and C. J. Tomlin. "Mathematical modeling of planar cell polarity to understand domineering nonautonomy". *Science*, Vol. 307, pp. 423–426, 2005.

[An 09] G. An, Q. Mi, J. Dutta-Moscato, and Y. Vodovotz. "Agent-based models in translational systems biology". *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, Vol. 1, No. 2, pp. 159–171, 2009.

[Arde 03] I. I. Ardelean and M. Cavaliere. "Modelling biological processes by using a probabilistic P system software". *Natural Computing*, Vol. 2, No. 2, pp. 173–197, 2003.

[Axel 01] J. D. Axelrod. "Unipolar membrane association of Dishevelled mediates Frizzled planar cell polarity signaling". *Genes & Development*, Vol. 15, pp. 1182–1187, 2001.

[Axel 09] J. D. Axelrod. "Progress and challenges in understanding planar cell plarity signaling". *Seminar Cell Developmental Biology*, Vol. 20, No. 8, pp. 964–971, 2009.

[Axel 11] J. D. Axelrod and C. J. Tomlin. "Modeling the control of planar cell polarity". *WIREs Systems Biology and Medicine*, pp. 865–882, 2011.

[Axel 98] J. D. Axelrod, J. R. Miller, J. M. Shulman, R. T. Moon, and N. Perrimon. "Differential recruitment of Dishevelled provides signaling specificity in

the planar cell polarity and Wingless signaling pathways". *Genes & Development*, Vol. 12, No. 16, pp. 2610–2622, 1998.

[Bald 10] P. Baldan, N. Cocco, A. Marin, and M. Simeoni. "Petri nets for modelling metabolic pathways: a survey". *Natural Computing*, Vol. 9, No. 4, pp. 955–989, 2010.

[Ball 10] P. Ballarini and M. L. Guerriero. "Query-based verification of qualitative trends and oscillations in biochemical systems". *Theoretical Computer Science*, Vol. 411, No. 20, pp. 2019–2036, April 2010.

[Barr 06] J. Barrow. "Wnt/PCP signaling: A veritable polar star in establishing patterns of polarity in embryonic tissues". *Seminars in Cell and Developmental Biology*, Vol. 17, No. 2, pp. 185–193, 2006.

[BASS 05] J. B. BASSINGTHWAIGHTE, H. J. CHIZECK, L. E. ATLAS, and H. QIAN. "Multiscale Modeling of Cardiac Cellular Energetics". *Annals of the New York Academy of Sciences*, Vol. 1047, pp. 395–426, June 2005.

[Bast 03] R. Bastock, H. Strutt, and D. Strutt. "Strabismus is asymmetrically localised and binds to prickle and dishevelled during *Drosophila* planar cell polarity". *Development*, Vol. 130, No. 13, pp. 3007–3014, 2003.

[Batt 05] G. Batt, D. Ropers, H. D. Jong, J. Geiselmann, R. Mateescu, M. Page, and D. Schneider. "Validation of qualitative models of genetic regulatory networks by model checking: Analysis of the nutritional stress response in Escherichia coli". *Bioinformatics*, Vol. 21, No. Suppl 1, 2005.

[Bell 01a] Y. Bellaíche, M. Gho, J. A. Kaltschmidt, A. H. Brand, and F. Schweisguth. "Frizzled regulates localization of cell-fate determinants and mitotic spindle rotation during asymmetric cell division". *Nature Cell Biology*, Vol. 3, No. 1, pp. 50–57, January 2001.

[Bell 01b] Y. Bellaíche, A. Radovic, D. F. Woods, C. D. Hough, M.-L. Parmentier, C. J. O'Kane, P. J. Bryant, and F. Schweisguth. "The partner of Inscuteable/Discs-Large complex is required to establish planar polarity during asymmetric cell division in *Drosophila*". *Cell*, Vol. 106, No. 3, pp. 355–366, August 2001.

[Bell 04] Y. Bellaíche, O. Beaudoin-Massiani, I. St'uttem, and F. Schweisguth. "The planar cell polarity protin Strabismus promotes Pins anterior localization during asymmetric division of sensory organ precursor cells in *Drosophila*". *Development*, Vol. 131, No. 2, pp. 469–478, January 2004.

[Blak 11] J. Blakes, J. Twycross, F. J. Romero-Campero, and N. Krasnogor. "The Infobiotics Workbench: an integrated in silico modelling platform for Systems and Synthetic Biology". *Bioinformatics*, Vol. 27, No. 23, pp. 3323–3324, 2011.

[Blos 08] R. Blossey, L. Cardelli, and A. Phillips. "Compositionality, stochasticity and cooperativity in dynamic models of gene regulation". *HFSP Journal*, Vol. 2, No. 1, pp. 17–28, 2008.

[Blou 91] D. Blount. "Comparison of stochastic and deterministic models of a linear chemical reaction with diffusion". *The Annals of Probability*, Vol. 19, No. 4, pp. 1440–1462, October 1991.

[Bock 81] H. G. Bock. *Numberical treatment of inverse problems in chemical reaction kinetics*, Chap. Modelling of Chemical Reaction Systems, pp. 102–125. Vol. 18 of *Springer Series in Chemical Physics*, Springer, Heidelberg, 1981.

[Bona 97] H. P. J. Bonarius, G. Schmid, and J. Tramper. "Flux analysis of underdetermined metabolic networks: the quest for the missing constraints". *Trends Biotechonolgy*, Vol. 15, pp. 308–314, 1997.

[Brai 99] S. C. Brailsford, C. N. Potts, and B. M. Smith. "Constraint Satisfaction Problems: Algorithms and Applications". *European Journal of Operational Research*, Vol. 119, No. 3, pp. 557–581, 1999.

[Bran 02] A. Brandt. "Multiscale Scientific Computation: Review 2001". In: T. Barth, T. Chan, and R. Haimes, Eds., *Multiscale and Multiresolution Methods*, pp. 3–95, Springer Berlin Heidelberg, 2002.

[Bran 77] A. Brandt. "Multi-Level Adaptive Solutions to Boundary-Value Problems". *Mathematics of Computation*, Vol. 31, pp. 333–390, 1977.

[Brei 10] R. Breitling, R. Donaldson, D. Gilbert, and M. Heiner. "Biomodel Engineering - From Structure to Behavior". *Lecture Notes Bioinformatics*, Vol. 5945, pp. 1–12, 2010.

[Bura 09] Y. Burak and B. I. Shraiman. "Order and Stochastic Dynamic in *Drosophila* Planar Cell Polarity". *PLoS Computational Biology*, Vol. 5, No. 12, p. e1000628, 2009.

[Cadi 97] K. M. Cadigan and R. Nusse. "Wnt signaling: a common theme in animal development". *Genes and Development*, Vol. 11, No. 24, pp. 3286–3305, 1997.

[Calz 06] L. Calzone, N. Chabrier-Rivier, F. Fages, and S. Soliman. *Machine Learning Biochemical Networks from Temporal Logic Properties*. Vol. 4220 of *Lecture Notes in Computer Science*, Springer, 2006.

[Carr 03] F. Carreira-Barbosa, M. L. Concha, M. Takeuchi, N. Ueno, S. W. Wilson, and M. Tada. "Prickle 1 regulates cell movements during gastrulation and neuronal migration in zebrafish". *Development*, Vol. 130, No. 17, pp. 4037–4046, September 2003.

[Casa 02] J. Casal, G. Struhl, and P. A. Lawrence. "Developmental compartments and planar polarity in *Drosophila*". *Current Biology*, Vol. 12, No. 14, pp. 1189–1198, 2002.

[Casa 06] J. Casal, P. A. Lawrence, and G. Struhl. "Two separate molecular systems, dachsous/fat and starry night/frizzled, act independently to confer planar cell polarity". *Development*, Vol. 133, No. 22, pp. 4561–4572, 2006.

[Chae 99] J. Chae, M.-J. Kim, J. H. Goo, S. Collier, D. Gubb, J. Charlton, P. N. Adler, and W. J. Park. "The *Drosophila* tissue polarity gene starry night encodes a member of the protocadherin family". *Development*, Vol. 126, No. 23, pp. 5421–5429, 1999.

[Chas 02] C. Chassagnole, N. Noisommit-Rizzi, J. Schmid, K. Mauch, and M. Reuss. "Dynamic modeling of the central carbon metabolism of Escherichia coli". *Biotechnology and Bioengineering*, Vol. 79, No. 1, pp. 53–73, 2002.

[Chen 08] W. S. Chen, D. Antic, M. Matis, C. Y. Logan, M. Povelones, G. A. Anderson, R. Nusse, and J. D. Axelrod. "Asymmetric homotypic interactions of the atypical cadherin flamingo mediate intercellular polarity signaling". *Cell*, Vol. 133, No. 6, pp. 1093–1105, 2008.

[Chen 99] T. Chen, H. L. He, and G. M. Church. "Modeling Gene Expression with differential equations". *Pacific Symposium Biocomputing*, 1999.

[Clar 01] E. M. Clarke, O. Grumberg, and D. Peled. *Model checking*. MIT Press, 2001.

[Clar 95] H. F. Clark, D. Brentrup, K. Schneitz, A. Bieber, C. Goodman, and M. Noll. "Dachsous encodes a member of the cadherin superfamily that controls imaginal disc morphogenesis in Drosophila". *Genes & Development*, Vol. 9, No. 12, pp. 1530–1542, June 1995.

[Clas 05] A. Classen, K. Anderson, E. Marois, and S. Eaton. "Hexagonal packing of Drosophila wing epithelial cells by the planar cell polarity pathway". *Developmental Cell*, Vol. 9, pp. 805–817, 2005.

[Coll 97] S. Collier and D. Gubb. "Drosophila tissue polarity requires the cell-autonomous activity of the fuzzy gene which encodes a novel transmembrane protein". *Development*, Vol. 124, pp. 4029–4037, 1997.

[Comi 89] S. D. Comis, J. O. Pickles, M. P. Osborne, and C. B. Pepper. "Tip-link organization in anomalously-orientation hair cells of the guinea pig cochlea". *Hearing Research*, Vol. 40, No. 3, pp. 205–211, July 1989.

[Corn 06] D. W. Corne and P. Frisco. "Dynamics of HIV infection studied with cellular automata and conformon-P systems". *BioSystems*, Vol. 91, No. 3, pp. 531–544, 2006.

[Dabd 03] A. Dabdoub, M. J. Donohue, A. Brennan, V. Wolf, M. Montcouquiol, D. A. Sassoon, J.-C. Hseih, J. S. Rubin, P. C. Salinas, and M. W. Kelley. "Wnt signaling mediates reorientation of outer hair cell stereocilliary bundles in the mammalian cochlea". *Development*, Vol. 130, No. 11, pp. 2375–2384, 2003.

[Dada 11] J. O. Dada and P. Mendes. "Multi-scale modelling and simulation in systems biology". *Integr. Biol.*, Vol. 3, pp. 86–96, 2011.

[Das 04] G. Das, A. Jenny, T. J. Klein, S. Eaton, and M. Mlodzik. "Diego inteacts with Prickle and Strabismus/Van Gogh to localize planar cell polarity complexes". *Development*, Vol. 131, No. 18, pp. 4467–4476, September 2004.

[Dasz 01] M. Daszykowski, B. Walczak, and D. Massart. "Looking for natural patterns in data: Part 1. Density-based approach". *Chemometrics and Intelligent Laboratory Systems*, Vol. 56, No. 2, pp. 83 – 92, 2001.

[Desh 02] J. Desharnais and P. Panangaden. “Continuous Stochastic Logic Characterizes Bisimulation of Continuous-time Markov Processes”. *J. of Logic and Alg. Progr*, Vol. 56, pp. 1–2, 2002.

[Dhur 85] P. Dhurjati, D. Ramkrishna, M. Flickinger, and G. Tsao. “A cybernetic view of microbial growth: modeling of cells as optimal strategists”. *Biotechnology and Bioengineering*, Vol. 7, No. 1, pp. 1–9, 1985.

[Dona 08] R. Donaldson and D. Gilbert. “A Model Checking Approach to the Parameter Estimation of Biochemical Pathways”. *LNCS/LNBI*, Vol. 5307, pp. 269–287, 2008.

[Eato 03] S. Eaton. “Cell biology of planar polarity transmission in the Drosophila wing”. *Mechanisms of Development*, Vol. 120, No. 11, pp. 1257–1264, November 2003.

[Elow 00] M. B. Elowitz and S. Leibler. “A synthetic osillatory network of transcriptional regulators”. *Nature*, Vol. 403, pp. 335–338, 2000.

[Emon 05] T. Emonet, C. Macal, M. North, C. Wickersham, and P. Cluzel. “AgentCell: a digital single-cell assay for bacterial chemotaxis”. *Bioinformatics*, Vol. 21, No. 11, p. 2714, 2005.

[Endy 05] D. Endy. “Foundations for engineering biology”. *Nature*, Vol. 438, No. 7067, pp. 449–453, 2005.

[Enge 08] J. Engelberg, G. Ropella, and C. Hunt. “Essential operating principles for tumor spheroid growth”. *BMC Systems Biology*, Vol. 2, No. 1, p. 110, 2008.

[Erme 93] G. B. Ermentrout and L. Edelstein-Keshet. “Cellular Automata Approaches to Biological Modeling”. *Journal of Theoretical Biology*, Vol. 160, pp. 97–133, Jan. 1993.

[Este 96] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise". In: *KDD*, pp. 226–231, 1996.

[Fant 03] M. Fanto, L. Clayton, J. Meredith, K. Hardiman, B. Charroux, S. Kerridge, and H. MacNeil. "The tumor-suppressor and cell adhesion molecule Fat controls planar polarity vai physical interactions with Atrophin, a transcriptional co-repressor". *Development*, Vol. 130, No. 4, pp. 763–774, 2003.

[Fant 04] M. Fanto and H. McNeil. "Planar Polarity from flies to vertebrates". *Journal of Cell Science*, Vol. 117, pp. 527–533, February 2004.

[Farh 07] R. Farhadifar, J.-C. Raper, B. Aigouy, S. Eaton, and F. Jalicher. "The influence of cell mechanisms, cell-cell interactions, and proliferation on epithelial packing". *Current Biology*, Vol. 17, No. 24, pp. 2095–2104, December 2007.

[Feig 01] F. Feiguin, M. Hannus, M. Mlodzik, and S. Eaton. "The ankyrin repeat protein Diego mediates Frizzled-dependent planar polarization". *Developmental Cell*, Vol. 1, No. 1, pp. 93–101, July 2001.

[Folc 07] V. Folcik, G. An, and C. Orosz. "The Basic Immune Simulator: an agent-based model to study the interactions between innate and adaptive immunity". *Theor Biol Med Model*, Vol. 4, 2007.

[Fris 68] D. Fristrom. "Cellular degeneration in wing development of the mutant vestigial of Drosophila melanogaster". *Journal of Cell Biology*, Vol. 39, No. 2, pp. 488–491, November 1968.

[Fuji 90] H. Fujita. "Mutant golden hamsters with an abnormal outer hair cell stereociliary arrangement". *Hearing Research*, Vol. 44, No. 1, pp. 63–69, February 1990.

[Funa 03] A. Funahashi, M. Morohashi, H. Kitano, and N. Tanimura. "CellDesigner: a process diagram editor for gene-regulatory and biochemical networks". *BIOSILICO*, Vol. 1, No. 5, pp. 159 – 162, 2003.

[Funa 08] A. Funahashi, Y. Matsuoka, A. Jouraku, M. Morohashi, N. Kikuchi, and H. Kitano. "CellDesigner 3.5: A Versatile Modeling Tool for Biochemical Networks". *Proceedings of the IEEE*, Vol. 96, No. 8, pp. 1254–1265, Aug. 2008.

[Gao 11] Q. Gao, F. Liu, D. Gilbert, M. Heiner, and D. Tree. "A Multiscale Approach to Modelling Planar Cell Polarity in *Drosophila* Wing using Hierarchically Coloured Petri nets". In: *Proc. 9th International Conference on Computational Methods in Systems Biology (CMSB 2011)*, pp. 209–218, ACM digital library, September 2011.

[Gao 12] Q. Gao, D. Gilbert, M. Heiner, F. Liu, D. Maccagnola, and D. Tree. "Multiscale Modelling and Analysis of Planar Cell Polarity in the *Drosophila* Wing". *IEEE/ACM TCBB*, Vol. 99, 2012.

[Geco 11] Gecode. "Gecode: An Open Constraint Solving Library". http://www.gecode.org, 2011.

[Gheo 08] M. Gheorghe, N. Krasnogor, and M. Cámara. "P systems applications to systems biology". *Biosystems*, Vol. 91, No. 3, pp. 435–437, 2008.

[Gho 98] M. Gho and F. Schweisguth. "Frizzled signalling controls orientation of asymmetric sense organ precursor cell divisions in Drosophila". *Nature*, Vol. 393, No. 6681, pp. 178–181, May 1998.

[Gho 99] M. Gho, Y. Bellaíche, and F. Schweisguth. "Revisiting the Drosophila microchaete lineage: a novel intrinsically asymmetric cell division generates a glial cell". *Development*, Vol. 126, No. 16, pp. 3573–3584, August 1999.

[Ghos 01] R. Ghosh and C. J. Tomlin. "Lateral Inhibition through Delta-Notch Signaling: A Piecewise Affine Hybrid Model". *Lecture Notes in Computer Science*, Vol. 2034, pp. 232–246, 2001.

[Ghos 04] R. Ghosh and C. Tomlin. "Symbolic reachable set computation of piecewise affine hybrid automata and its application to biological modelling: Delta-notch protein signalling". 2004.

[Gilb 06] D. Gilbert, H. Fu, X. Gu, R. Orton, S. Robinson, V. Vyshemirsky, M. J. Kurth, C. S. Downes, and W. Dubitzky. "Computational methodologies for modelling, analysis and simulation of signalling networks". *Briefings in Bioinformatics*, Vol. 7, No. 4, pp. 339–353, 2006.

[Gilb 07a] D. Gilbert, M. Heiner, and S. Lehrack. "A unifying framework for modelling and analysing biochemical pathways using Petri nets". *LNCS/LNBI*, Vol. 4695, pp. 200–216, 2007.

[Gilb 07b] D. Gilbert, M. Heiner, and S. Lehrack. "A Unifying Framework for Modelling and Analysing Biochemical Pathways Using Petri Nets". In: *Proc. 5th International Conference on Computational Methods in Systems Biology (CMSB 2007), Edinburgh*, pp. 200–216, Springer, 2007. This paper is a short version of Techn. Report I-02/2007, BTU Cottbus.

[Gilb 09] D. R. Gilbert, R. Breitling, M. Heiner, and R. A. Donaldson. "An introduction to BioModel engineering, illustrated for signal transduction pathways". In: *Membrane Computing*, pp. 13–28, Springer, 2009.

[Gilb 10] D. Gilbert, M. Heiner, R. Breitling, and R. Orton. "Computational Modelling of Kinase Signalling Cascades". In: R. Seger, Ed., *MAP Kinase Signaling Protocols*, pp. 369–384, Humana Press, 2010.

[Gill 77] D. T. Gillespie. “Exact stochastic simulation of coupled chemical reactions”. *Journal of Physical Chemistry*, Vol. 81, No. 25, pp. 2340–2361, 1977.

[Gonz 03] P. Gonzalez, M. Cardenas, D. Camacho, A. Franyuti, O. Rosas, and J. Lagunez-Otero. “Cellulat: an agent-based intracellular signalling model”. *Biiosystems*, Vol. 68, No. 3, pp. 171–185, 2003.

[Gonz 06] P. Gonzalez, R. Smallwood, E. Qwarnstrom, and M. Holcombe. “Formal agent- based modelling of intracellular chemical interactions”. *Biiosystems*, Vol. 85, No. 1, pp. 37–45, 2006.

[Goss 11] E. Gosse. *Encyclopedia Britannica*, Chap. Ode. Vol. 20, Cambridge University Press, 11th ed. Ed., 1911.

[Gran 06] M. Grant, K. Mostov, T. Tlsty, and C. Hunt. “Simulating Properties of in vitro Epithelial Cell Morphogenesis”. *PLoS Computaional Biology*, Vol. 2, No. 10, p. e129, 2006.

[Grim 06] V. Grimm and S. F. Railsback. “Agent-Based Models in Ecology: Patterns and Alternative Theories of Adaptive Behaviour”. In: F. Billari, T. Fent, A. Prskawetz, and J. Scheffran, Eds., *Agent-Based Computational Modelling*, pp. 139–152, Physica-Verlag HD, 2006.

[Gubb 82] D. Gubb and A. Garcia-Bellido. “A genetic analysis of the determination of cuticular polarity during development in *Drosophila melanogaster*”. *Journal of Embryology and Experimental Morphology*, Vol. 68, pp. 37–57, April 1982.

[Gubb 99] D. Gubb, C. Green, D. Huen, D. Coulson, G. Johnson, D. Tree, S. Collier, and J. Roote. “The balance between isoforms of the Prickle LIM domain protein is critical for planar polarity in Drosophila imaginal discs”. *Genes & Development*, Vol. 13, No. 17, pp. 2315–2327, 1999.

[Halk 08] M. Halkidi and M. Vazirgiannis. “A density-based cluster validity approach using multi-representatives”. *Pattern Recognition Letters*, Vol. 29, No. 6, pp. 773–786, april 2008.

[Hann 02] M. Hannus, F. Feiguin, C.-P. Heisenberg, and S. Eaton. “Planar cell polarization requires Widerborst, a B’ regulatory subunit of protein phosphatase 2A”. *Development*, Vol. 129, No. 14, pp. 3493–3503, July 2002.

[Haru 10] T. Harumoto, M. Ito, Y. Shimada, T. J. Kobayashi, H. R. Ueda, B. Lu, and T. Uemura. “Atypical cadherins dachsous and fat control dynamics of noncentrosomal microtubules in planar cell polarity”. *Developmental Cell*, Vol. 19, No. 3, pp. 389–401, 2010.

[Hast 09] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*, Chap. 14.3.12, pp. 520–528. *2nd ed*, New York: Springer, October 2009.

[Heat 09] A. P. Heath and L. E. Kavraki. “Computational challenges in systems biology”. *Computer Science Review*, Vol. 3, No. 1, pp. 1–17, 2009.

[Hein 06] M. Heinemann and S. Panke. “Synthetic biology - putting engineering into biology”. *Bioinformatics*, Vol. 22, No. 22, 2006.

[Hein 08] M. Heiner, D. Gilbert, and R. Donaldson. “Petri nets in Systems and Synthetic Biology”. In: *Schools on Formal Methods (SFM)*, pp. 215–264, LNCS 5016, Springer, 2008.

[Hein 11] M. Heiner and D. Gilbert. “How might Petri nets enhance your systems biology toolkit”. In: *Proceedings of PETRI NETS*, pp. 17–37, Springer, 2011.

[Hein 12a] M. Heiner and D. Gilbert. “BioModel Enginneering for Multiscale Systems Biology”. *Progress in Biophysics and Molecular Biology*, Vol. In press, 2012.

[Hein 12b] M. Heiner, M. Herajy, F. Liu, C. Rohr, and M. Schwarick. “Snoopy a unifying Petri net tool”. In: *Proc. PETRI NETS 2012*, p. 398407, Springer, June 2012.

[Hoek 07] A. G. Hoekstra, E. Lorenz, J.-L. Falcone, and B. Chopard. “Toward a Complex Automata Formalism for MultiScale Modeling”. *International Journal for Multiscale Computational Engineering*, Vol. 5, pp. 491–502, 2007.

[Hoek 08] A. Hoekstra, J.-L. Falcone, A. Caiazzo, and B. Chopard. “Multi-scale Modeling with Cellular Automata: The Complex Automata Approach”. In: H. Umeo, S. Morishita, K. Nishinari, T. Komatsuzaki, and S. Bandini, Eds., *Cellular Automata*, pp. 192–199, Springer Berlin Heidelberg, 2008.

[Hood 03] L. Hood. “Systems biology: integrating technology, biology and computation”. *Mechanisms of Aging and Development*, Vol. 124, No. 1, pp. 9–16, 2003.

[Hoop 06] S. Hoops, S. Sahle, R. Gauges, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes, and U. Kummer. “COPASI - a COmplex PAthway SImulator”. *Bioinformatics*, Vol. 22, pp. 3067–3074, 2006.

[Horn 72] F. Horn and R. Jackson. “General mass action kinetics”. *Archive for Rational Mechanics and Analysis*, Vol. 47, No. 2, pp. 81–116, 1972.

[Huck 03] M. Hucka, A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano, and the rest of the SBML Forum. “The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models”. *Bioinformatics*, Vol. 19, pp. 524–531, 2003.

[Huck 12] M. Hucka, F. Bergmann, S. Hoops, S. Keating, S. Sahle, J. Schaff, L. Smith, and D. Wilkinson. “The Systems Biology Markup Language (SBML): Language Specification for Level 3 Version 1 Core”. October 2012.

[Hunt 95] T. Hunter. "Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling". *Cell*, Vol. 80, No. 2, pp. 225–236, 1995.

[Idek 01] T. Ideker, T. Galitski, and L. Hood. "A new approach to decode life: systems biology". *Annual Review of Genomics and Human Genetics*, Vol. 2, pp. 343–372, 2001.

[Jenn 03] A. Jenny, R. S. Darken, P. A. Wilson, and M. Mlodzik. "Prickle and Strabismus form a functional complex to generate a correct axis during planar cell polarity signaling". *EMBO Journal*, Vol. 22, pp. 4409–4420, September 2003.

[Jenn 05] A. Jenny, J. Reynolds-Kenneally, G. Das, M. Burnett, and M. Mlodzik. "Diego and Prickle regulate Frizzled planar cell polarity signalling by competing for Dishevelled binding". *Nature Cell Biology*, Vol. 7, pp. 691–697, 2005.

[Jens 09] K. Jensen and L. M. Kristensen. *Coloured Petri nets*. Springer, 2009.

[Jens 81] K. Jensen. "Coloured Petri nets and the Invariant-Method". *Theoretical Computer Science*, Vol. 14, pp. 317–336, 1981.

[Joll 86] I. T. Jolliffe. *Principal Component Analysis*. Springer-Verlag, New York, New York, 1986.

[Jone 96] K. H. Jones, J. Liu, and P. N. Adler. "Molecular analysis of EMS-Induced *frizzled* mutations in *Drosophila melanogaster*". *Genetics*, Vol. 142, No. 1, pp. 205–215, 1996.

[Jong 04] H. D. Jong, J. Gouze, C. Hernandez, M. Page, T. Sari, and J. Geiselmann. "Qualitative simulation of genetic regulatory networks using piecewise-linear models". *Bulletin of Mathematical Biology*, Vol. 66, No. 2, 2004.

[Kacs 73] H. Kacser and J. A. Burns. "The control of flux". *Symposia of the Society of Experimental Biology*, Vol. 27, pp. 65–104, 1973.

[Kauf 69] S. Kauffman. "Metabolic stability and epigenesis in randomly constructed genetic nets". *Journal of Theoretical Biology*, Vol. 22, No. 3, 1969.

[Kier 09] L. B. Kier and P. G. Seybold. "Cellular Automata Modeling of Complex Biochemical Systems". In: R. A. Meyers, Ed., *Encyclopedia of Complexity and Systems Science*, pp. 848–865, Springer, 2009.

[Kili 94] J. Kilingensmith and R. Nusse. "Signaling by wingless in Drosophila". *Developmental Biology*, Vol. 166, No. 2, pp. 396–414, December 1994.

[Kirk 83] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. "Optimization by Simulated Annealing". *Science*, Vol. 220, No. 4598, pp. 671–680, 1983.

[Kita 02a] H. Kitano. "Computational systems biology". *Nature*, Vol. 420, No. 6912, pp. 206–210, November 2002.

[Kita 02b] H. Kitano. "Systems biology: a brief overview". *Science*, Vol. 295, No. 5560, pp. 1662–1664, March 2002.

[Kita 05] H. Kitano, A. Funahashi, Y. Matsuoka, and K. Oda. "Using process diagrams for the graphical representation of biological networks". *Nat Biotechnol*, Vol. 23, No. 8, pp. 961–6, 2005.

[Klan 11] M. Klann, A. Lapin, and M. Reuss. "Agent-based simulation of reactions in the crowded and structured intracellular environment: Influence of mobility and location of the reactants". *BMC Systems Biology*, Vol. 5, No. 1, p. 71, 2011.

[Kolc 00] W. Kolch. "Meaningful relationships: the regulation of the Ras/Raf/MEK/ERK pathway by protein interactions". *Journal of Biochemistry*, Vol. 351, pp. 289–305, 2000.

[Komp 84] D. Kompala, D. Ramkrishna, and G. Tsao. “Cybernetic modeling of microbial growth on multiple substrates”. *Biotechnology and Bioengineering*, Vol. 26, No. 1, pp. 1272–1281, 1984.

[Kono 71] R. J. Konopka and S. Benzer. “Clock Mutants of *Drosophila melanogaster*”. *Proceedings of the National Academy of Sciences*, Vol. 68, No. 9, pp. 2112–2116, 1971.

[Kras 94] R. E. Krasnow and P. N. Adler. “A single *frizzled* protein has a dual function in tissue polarity”. *Development*, Vol. 120, No. 7, pp. 1883–893, 1994.

[Kras 95] R. E. Krasnow, L. L. Wong, and P. N. Adler. “*dishevelled* is a component of the *frizzled* signaling pathway in Drosophila”. *Development*, Vol. 121, No. 12, pp. 4095–4102, 1995.

[Kurt 72] T. G. Kurtz. “The relationship between stochastic and deterministic models for chemical reactions”. *Journal of Chemical Physics*, Vol. 57, No. 7, pp. 2976–2978, October 1972.

[Lawr 02] P. A. Lawrence, J. Casal, and G. Struhl. “Towards a model of the organisation of planar polarity and pattern in the *Drosophila* abdomen”. *Development*, Vol. 129, No. 11, pp. 2749–2760, 2002.

[Lawr 04] P. A. Lawrence, J. Casal, and G. Struhl. “Cell interactions and planar polarity in the abdominal epidermis of *Drosophila*”. *Development*, Vol. 131, No. 19, pp. 4651–4664, 2004.

[Lawr 07] P. A. Lawrence, G. Struhl, and J. Casal. “Planar Cell Polarity: one or two pathways?”. *Nature Reviews Genetics*, Vol. 8, No. 7, pp. 555–563, 2007.

[Lawr 10] P. A. Lawrence, G. Struhl, and J. Casal. “Planar Cell Polarity: A Bridge Too Far?”. *Current Biology*, Vol. 18, No. 20, pp. 959–961, 2010.

[Lawr 66] P. A. Lawrence. "The hormonal control of the development of hairs and bristles in the milkweed bug, *Oncopeltus fasciatus*, Dall". *Journal of Experimental Biology*, Vol. 44, pp. 507–522, June 1966.

[Lawr 99] P. A. Lawrence, J. Casal, and G. Struhl. "The Hedgehog morphogen and gradients of cell affinity in the abdomen of *Drosophila*". *Development*, Vol. 126, No. 11, pp. 2441–2449, 1999.

[Le G 06] J. F. Le Garrec, P. Lopez, and M. Kerszberg. "Establishment and maintenance of planar epithelial cell polarity by asymmetric cadherin bridges: A computer model". *Development Dynamics*, Vol. 235, pp. 235–246, 2006.

[Le G 08] J. F. Le Garrec and M. Kerszberg. "Modeling polarity buildup and cell fate decision in the fly eye: Insight into the connection between the PCP and Notch pathways". *Development Genes & Evolution*, Vol. 218, pp. 413–426, 2008.

[Li 08] N. Li, K. Verdolini, G. Clermont, Q. Mi, E. Rubinstein, P. Hebda, and Y. Vodovotz. "A patient-specific in silico model of inflammation and healing tested in acute vocal fold injury". *PLoS ONE*, Vol. 3, No. 7, p. e2789, 2008.

[Liu 08] X. Liu, J. Jiang, O. Ajayi, X. Gu, and D. Gilbert. "BioNessie(G) - A grid enabled biiochemical networks simulation environment". *Studies in Health Techonology and Informatics*, Vol. 138, pp. 147–157, 2008.

[Liu 10] F. Liu and M. Heiner. "Colored Petri nets to model and simulate biological systems". In: *International Workshop on Biological Processes and Petri nets (BioPPN), satellite event of Petri nets 2010*, June 2010.

[Liu 11] F. Liu and M. Heiner. "Manual for Colored Petri Nets in Snoopy". Tech. Rep., Brandenburg University of Technology Cottbus, Department of Computer Science, July 2011.

[Liu 12] F. Liu. *Colored Petri nets for Systems Biology*. PhD thesis, Brandenburg University of Technology Cottbus, Department of Computer Science, January 2012.

[Loll 06] P. Lollini, S. Motta, and F. Pappalardo. "Discovery of cancer vaccination protocols with a genetic algorithm driving an agent based simulator". *BMC Bioinformatics*, Vol. 7, No. 1, p. 352, 2006.

[Lu 99] B. Lu, T. Usui, T. Uemura, L. Jan, and Y. N. Jan. "Flamingo controls the planar polarity of sensory bristles and asymmetric division of sensory organ precursors in Drosophila". *Current Biology*, Vol. 9, No. 21, pp. 1247–1250, November 1999.

[Ma 03] D. Ma, C. H. Yang, H. McNeill, and M. A. Simon. "Fideility in planar cell polarity in *Drosophila*". *Nature*, Vol. 421, No. 6922, pp. 543–547, 2003.

[Ma 08] D. Ma, K. Amonlirviman, R. L. Raffard, A. Abate, C. J. Tomlin, and J. D. Axelrod. "Cell packing influences planar cell polarity signaling". *Proceedings of the National Academy of Sciences of the United State of America(PNAS)*, Vol. 105, No. 48, pp. 18800–18805, December 2008.

[Macc 12] D. Maccagnola, E. Messina, Q. Gao, and D. Gilbert. "A machine learning approach for generating temporal logic classifications of complex model behaviours". In: O. Rose and A. M. Uhrmacher, Eds., *Winter Simulation Conference, WSC '12, Berlin, Germany, December 9-12, 2012*, p. 294, WSC, 2012.

[MacQ 65] J. MacQueen. "Some methods for classification and analysis of multivariate observations". In: *Proc. of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–297, University of California Press, Berkeley, CA, 1965.

[Maly 96] T. Maly and L. R. Petzold. “Numberical methods and software for sensitivity analysis of differential-algebraic systems”. *Applied Numberical Mathematics*, Vol. 20, pp. 57–79, 1996.

[Mata 04] H. Matakatsu and S. S. Blair. “Interactions between Fat and Dachsous and the regulation of planar cell polarity in the drosophila wing”. *Development*, Vol. 131, No. 15, pp. 3785–3794, August 2004.

[MATL 11] MATLAB. *version 7.11.0*. The MathWorks Inc., Natick, Massachusetts, 2011.

[McAd 97] H. H. McAdams and A. Arkin. “Stochastic mechanisms in gene expression”. *Proceedings of the National Academy of Sciences*, Vol. 94, No. 3, pp. 814–819, February 1997.

[McNe 97] H. McNeil, C. H. Yang, M. Brodsky, J. Ungos, and M. A. Simon. “*mirror* encodes a novel PBX-class homeoprotein that functions in the definition of the dorsal-ventral border in the Drosophila eye”. *Genes & Development*, Vol. 11, No. 8, pp. 1073–1082, April 1997.

[Meho 11] A. Mehonic. *The role of physics in epithelial hoeostasis and development*. PhD thesis, Department of Physics and Astronomy, University College London, August 2011.

[Mend 09a] P. Mendes, S. Hoops, S. Sahle, R. Gauges, J. Dada, and U. Kummer. “Computational Modeling of Biochemical Networks Using COPASI”. In: I. V. Maly, Ed., *Systems Biology*, pp. 17–59, Humana Press, 2009.

[Mend 09b] P. Mendes, H. Messiha, N. Malys, and S. Hoops. “Chapter 22 Enzyme Kinetics and Computational Modeling for Systems Biology”. In: M. L. J. L. Brand, Ed., *Methods in Enzymology*, pp. 583 – 599, Academic Press, 2009.

[Ment 13] L. Menten and M. I. Michaelis. “Die Kinetik der Invertinwirkung”. *Biochemistry Z*, Vol. 49, pp. 333–369, 1913.

[Mere 07] E. Merelli, G. Armano, N. Cannata, F. Corradini, M. d’Inverno, A. Doms, P. Lord, A. Martin, L. Milanesi, and S. Moller. “Agents in bioinformatics, computational and systems biology”. *Brief Bioinfrmatics*, Vol. 8, No. 1, pp. 45–59, 2007.

[Mill 10] A. K. Miller, J. Marsh, A. Reeve, A. Garny, R. Britten, M. Halstead, J. Cooper, D. P. Nickerson, and P. F. Nielsen. “An overview of the CellML API and its implementation”. *BMC Bioinformatics*, Apr. 08 2010.

[Miln 80] R. Milner. “A Calculus of Communicating Systems”. *LNCS*, Vol. 92, 1980.

[Milo 02] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. “Network motifs: Simple building blocks of complex networks”. *Science*, Vol. 298, No. 5594, pp. 824–827, October 2002.

[Mith 08] F. Mitha, T. Lucas, F. Feng, T. Kepler, and C. Chan. “The Multiscale Systems Immunology project: software for cell-based immunological simulation”. *Source Code Biol Med*, Vol. 3, 2008.

[Mlod 02] M. Mlodzik. “Planar Cell polarization: do the same mechanisms regulate Drosophila tissue polarity and vertebrate gastrulation?”. *Trends in Genetics*, Vol. 18, No. 11, pp. 564–571, November 2002.

[Mura 89] T. Murata. “Petri Nets: Properties, Analysis and Applications”. In: *Proceedings of the IEEE*, April 1989.

[Nets 12] P. Nets. “Petri Nets World”. 2012.

[Neum 66] J. von Neumann. *Theory of self-reproducing automata*. University of Illinois Press, Urbana, 1966.

[Nick 05] D. Nickerson, N. Smith, and P. Hunter. "New developments in a strongly coupled cardiac electromechanical model". *Europace*, Vol. 7, No. Suppl 2, pp. 118–127, 2005.

[Nobl 03] D. Noble. "The future: putting Humpty-Dumpty together again". *Biochemical Society Transations*, Vol. 31, pp. 156–158, February 2003.

[Nuss 92] R. Nusse and H. E. Varmus. "Wnt genes". *Cell*, Vol. 69, No. 7, pp. 1073–1087, June 1992.

[Olgu 10] P. Olguin and M. Mlodzik. "A new spin on planar cell polarity". *Cell*, Vol. 142, No. 5, p. 6740676, 2010.

[Orto 05] R. J. Orton, O. E. Sturm, V. Vyshemirsky, M. Calder, D. R. Gilbert, and W. Kolch. "Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway". *Journal of Biochemistry*, Vol. 392, pp. 249–261, 2005.

[Park 02] M. Park and R. T. Moon. "The planar cell polarity gene stbm regulates cell behaviour and cell fate in vertebrate embryos". *Nature Cell Biology*, Vol. 4, pp. 20–25, January 2002.

[Park 96] W. Park, J. Liu, E. J. Sharp, and P. N. Adler. "The Drosophila tissue polarity gene inturned acts cell autonomously and encodes a novel protein". *Development*, Vol. 122, pp. 961–969, 1996.

[Paun 05] G. Păun and M. J. Pérez-jiménez. "Membrane computing: Brief introduction, recent results and applications". 2005.

[Pear 88] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan-Kaufmann Publishers, INC., 1988.

[Petr 62] C. A. Petri. *Kommunikation mit Automaten*. PhD thesis, Institut f´ur Instrumentelle Mathematik, Schriften des IIM Nr 2, Bonn, Apri 1962.

[Pett 05] A. Pettinen, T. Aho, O. Smolander, T. Manninen, A. Saarimem, K. Taattola, O. Yli-Harja, and M. Linne. "Simulation tools for biochemical networks: evaluation of performance and usabiligy". *Bioinformatics*, Vol. 21, pp. 357–363, 2005.

[Pitt 09] J. Pitt-Francis, P. Pathmanathan, M. O. Bernabeu, R. Bordas, J. Cooper, A. G. Fletcher, G. R. Mirams, P. Murray, J. M. Osborne, A. Walter, S. J. Chapman, A. Garny, I. M. van Leeuwen, P. K. Maini, B. Rodr?guez, S. L. Waters, J. P. Whiteley, H. M. Byrne, and D. J. Gavaghan. "Chaste: A test-driven approach to software development for biological modelling". *Computer Physics Communications*, Vol. 180, No. 12, pp. 2452 – 2471, 2009.

[Pogs 08] M. Pogson, M. Holcombe, R. Smallwood, and E. Qwarnstrom. "Introducing spatial information into predictive NF-kappa B modelling - an agent-based approach". *PLoS One*, june 2008.

[Pove 05] M. Povelones and R. Nusse. "The role of the cysteine-rich domain of Frizzled in Wingless-Armadillo signaling". *EMBO Journal*, Vol. 24, pp. 3493–3503, 2005.

[Puau 00] G. Păun. "Computing with Membranes". *JCSS: Journal of Computer and System Sciences*, Vol. 61, 2000.

[Puau 01] G. Păun. "Computing with Membranes (P Systems): An Introduction". In: *Current Trends in Theoretical Computer Science*, pp. 845–866, World Scientific, 2001.

[Puau 98] G. Păun. "Computing with membranes". Tech. Rep. TR 208, Turku Centre for Computer Science TUCS, Nov. 1998.

[Puau 99] G. Păun. "Computing with Membranes: An Introduction". *Bulletin of the EATCS*, Vol. 67, pp. 139–152, 1999.

[Raff 08] R. L. Raffard, K. Amonlirdviman, J. D. Axelrod, and C. J. Tomlin. "An adjoint-based parameter identification algorithm applied to planar cell polarity signaling". *IEEE Transaction on Automatic Control*, Vol. 53, pp. 109–121, 2008.

[Rawl 03] A. S. Rawls and T. Wolff. "Strabismus requires Flamingo and Prickle function to regulate tissue polarity in the Drosophila eye". *Development*, Vol. 130, No. 9, pp. 1877–1887, May 2003.

[Rizz 97] M. Rizzi, M. Baltes, U. Theobald, and M. Reuss. "In vivo analysis of metabolic dynamics in Saccharomyces cerevisiae: II. Mathematical model". *Biotechnology and Bioengineering*, Vol. 55, No. 4, pp. 592–608, August 1997.

[Roeg 01a] F. Roegiers, S. Younger-Shepherd, L. Y. Jan, and Y. N. Jan. "Bazooka is required for localization of determinants and controlling proliferation in the sensory organ precursor cell lineage in Drosophila". *Proceedings of the National Academy of Science*, Vol. 98, No. 25, pp. 144469–14474, December 2001.

[Roeg 01b] F. Roegiers, S. Younger-Shepherd, L. Y. Jan, and Y. N. Jan. "Two types of asymmetric divisions in the Drosophila sensory organ precursor cell lineage". *Nature Cell Biology*, Vol. 3, No. 1, pp. 58–67, January 2001.

[Roge 98] Y. Roger. "The green fluorescent protein". *Annual Review of Biochemistry*, Vol. 67, pp. 509–544, 1998.

[Rohr 10] C. Rohr, W. Marwan, and M. Heiner. "Snoopy - a unifying Petri net framework to investigate biomolecular networks". *Bioinformatics*, Vol. 26, No. 7, pp. 974–975, 2010.

[Rome 09a] F. J. Romero-Campero, J. Twycross, C. Hongqing, J. Blakes, and N. Krasnogor. "A multiscale modelling framework based on P systems". *Lecture Notes in Computer Science*, Vol. 5391, pp. 63–77, 01/2009 2009.

[Rome 09b] F. J. Romero-Campero, J. Twycross, M. Cámara, M. Bennett, M. Gheorghe, and N. Krasnogor. "Modular Assembly of Cell Systems Biology Models Using P Systems". *Int. J. Found. Comput. Sci*, Vol. 20, No. 3, pp. 427–442, 2009.

[Sahl 06] S. Sahle, R. Gauges, J. Pahle, N. Simus, U. Kummer, S. Hoops, C. Lee, M. Singhal, L. Xu, and P. Mendes. "Simulation of biochemical networks using COPASI: a complex pathway simulator". In: *Proceedings of the 38th conference on Winter simulation*, pp. 1698–1706, Winter Simulation Conference, 2006.

[Sant 01] R. M. Z. dos Santos and S. Coutinho. "Dynamics of HIV infection: A cellular automata approach". *P H Y S I C A L R E V I E W L E T T E R S*, Vol. 87, No. 16, p. 168102, October 2001.

[Sava 69a] M. A. Savageau. "Biochemical systems analysis: I. Some mathematical properties of the rate law for the component enzymatic reactions". *Journal of Theorectical Biology*, Vol. 25, No. 3, pp. 365–369, December 1969.

[Sava 69b] M. A. Savageau. "Biochemical systems analysis: II. The steady-state solutions for an n-pool system using a power-law approximation". *Journal of Theorectical Biology*, Vol. 25, No. 3, pp. 370–379, 1969.

[Sava 70] M. A. Savageau. “Biochemical systems analysis: III. Dynamic solutions using a power-law approximation”. *Journal of Theorectical Biology*, Vol. 26, No. 2, pp. 215–226, 1970.

[SBML 12] SBML-Team. “The Systems Biology Markup Language (SBML)”. September 2012.

[Scha 10] S. Schamberg, P. Houston, N. A. M. Monk, and M. R. Owen. “Modelling and analysis of Planar Cell Polarity”. *Bulletin of Mathematical Biology*, Vol. 72, pp. 645–680, 2010.

[Scha 99] J. Schaff and L. M. Loew. “The virtual cell”. *The Pacific Symposium on Biocomputing*, Vol. 4, pp. 228–239, 1999.

[Schn 07] S. Schnell, R. Grima, and P. Maini. “Multiscale Modeling in Biology”. *American Scientist*, Vol. 95, No. 2, p. 134, March-April 2007.

[Schw 11] M. Schwarick, C. Rohr, and M. Heiner. “MARCIE - Model checking And Reachability analysis done effiCIEntly”. In: *Proc. 8th International Conference on Quantitative Evaluation of SysTems (QEST 2011)*, pp. 91–100, IEEE CS Press, September 2011.

[Sego 04] J. Segovia-Juarez, S. Ganguli, and D. Kirschner. “Identifying control mechanisms of granuloma formation during M. tuberculosis infection using an agent-based model”. *Journal of Theoretical Biology*, Vol. 231, No. 3, pp. 357–376, 2004.

[Sham 97] L. F. Shampine and M. W. Reichelt. “The Matlab ODE suite”. *SIAM Journal on Scientific COmputing*, Vol. 18, pp. 1–22, 1997.

[Sham 99] L. F. Shampine, M. W. Reichelt, and J. A. Kierzenka. “Solving index-1 DAEs in MATLAB and Simulink”. *SIAM Review*, Vol. 41, pp. 538–552, 1999.

[Shih 92] J. Shih and R. Keller. "Patterns of cell motility in the organizer and dorsal mesoderm of Xenopus laevis". *Development*, Vol. 116, No. 4, pp. 915–930, December 1992.

[Shim 01] Y. Shimada, T. Usui, S. ichi Yangawa, M. Takeichi, and T. Uemura. "Asymmetric colocalization of Flamingo, a seven-pass transmembrane cadherin, and Dishevelled in planar cell polarization". *Current Biology*, Vol. 11, No. 11, pp. 859–863, June 2001.

[Shim 06] Y. Shimada, S. Yonemura, H. Ohkura, D. Strutt, and T. Uemura. "Polarized transport of frizzled along the planar microtubule arrays in *Drosophila* wing epithelium". *Developmental Cell*, Vol. 10, pp. 209–222, 2006.

[Shor 72] B. Shorrocks. *Drosophila*. Ginn & Company Limited, 18 Bedford Row, London, WC1R 4EJ, ed. 1 Ed., 1972.

[Shul 98] J. M. Shulman, N. Perrimon, and J. D. Axelrod. "Frizzled signaling and the developmental control of cell polarity". *Trends in Genetics*, Vol. 14, No. 11, pp. 452–458, November 1998.

[Simo 08] M. Simons and M. Mlodzik. "Planar cell polarity signaling: From fly development to human disease". *Annual Review of Genetics*, Vol. 42, pp. 517–540, 2008.

[Sout 08] J. Southern, J. Pitt-Francis, J. Whiteley, D. Stokeley, H. Kobashi, R. Nobes, Y. Kadooka, and D. Gavaghan. "Multi-scale computational modelling in biology and physiology". *Progress in Biophysics and and Molecular Biology*, Vol. 96, pp. 60–89, 2008.

[Sozi 05] O. Sozinova, Y. Jiang, D. Kaiser, and M. Alber. "A three-dimensional model of myxobacterial aggregation by contact-mediated interactions". In:

Proceedings of the National Academy of Sciences of the United States of America, pp. 11308–11312, August 2005.

[Stru 01] D. I. Strutt. "Asymmetric localization of Frizzled and the establishment of cell polarity in the *Drosophila* wing". *Molecular Cell*, Vol. 7, pp. 367–375, 2001.

[Stru 02a] D. Strutt, R. Johnson, K. Cooper, and S. Bray. "Asymmetric localization of Frizzled and the determination of Notch-dependent cell fate in the *Drosophila* eye". *Current Biology*, Vol. 12, No. 10, pp. 813–824, 2002.

[Stru 02b] D. I. Strutt. "Asymmetric localization of frizzled and the establishment of cell polarity in the *Drosophila* wing". *Molecular Cell*, Vol. 7, pp. 367–375, 2002.

[Stru 02c] D. I. Strutt. "The asymmetric subcellular localisation of components of the planar polarity pathway". *Seminars in Cell & Developmental Biology*, Vol. 13, No. 3, pp. 225–231, June 2002.

[Stru 02d] H. Strutt and D. Strutt. "Nonautonomous planar polarity patterning in *Drosophila*: Disheveleed-independent functions of Frizzled". *Developmental Cell*, Vol. 3, pp. 851–863, 2002.

[Stru 04] H. Strutt, J. Mundy, K. Hofstra, and D. Strutt. "Cleavage and secretion is not required for Four-jointed function in Drosophila patterning". *Development*, Vol. 131, No. 4, pp. 881–890, February 2004.

[Stru 08] D. Strutt and S. J. Warrington. "Planar polarity genes in the Drosophila wing regulate the localisation of the FH3-domain protein Multiple Wing Hairs to control the site of hair production". *Development*, Vol. 135, pp. 3103–3111, 2008.

[Stru 97a] G. Struhl, D. A. Barbash, and P. A. Lawrence. "Hedgehog acts by distinct gradient and signal relay mechanisms to organise cell type and cell polarity

in the*Drosophila* abdomen". *Development*, Vol. 124, No. 11, pp. 2155–2165, 1997.

[Stru 97b] D. I. Strutt, U. Weber, and M. Mlodzik. "The role of RhoA in tissue polarity and Frizzled signalling". *Nature*, Vol. 387, pp. 292–295, 1997.

[Stru 99] H. Strutt and D. Strutt. "Polarity determination in the Drosophila eye". *Current Opinion in Genetics & Development*, Vol. 9, No. 4, pp. 442–446, August 1999.

[Take 03] M. Takeuchi, J. Nakabayashi, T. Sakaguchi, T. S. Yamamoto, H. Takahashi, H. Takeda, and N. Ueno. "The *prickle*-related gene in vertebrates is essential for gastrulation cell movements". *Current Biology*, Vol. 13, No. 8, pp. 674–679, April 2003.

[Tayl 98] J. Taylor, N. Abramova, J. Charlton, and P. N. Adler. "Van Gogh: A new drosophila tissue polarity gene". *Genetics*, Vol. 150, No. 1, pp. 199–210, September 1998.

[Thes 94] H. Thesisen, J. Purcell, M. Bennett, D. Kansagara, A. Syed, and J. L. Marsh. "*dishevelled* is required during wingless signaling to establish both cell polarity and cell identity". *Development*, Vol. 120, No. 2, pp. 347–360, 1994.

[Thor 8] B. Thorne, A. Bailey, and S. Peirce. "Combining experiments with multi-cell agent- based modeling to study biological tissue patterning". *Brief Bioinformatics*, Vol. 4, No. 245-257, 8.

[Tomi 99] M. Tomita, K. Hashimoto, K. Takahashi, T. S. Shimizu, Y. Matsuzaki, F. Miyoshi, K. Saito, S. Tanida, K. Yugi, J. C. Venter, and C. A. H. III. "E-CELL: software environment for whole-cell simulation". *Bioinformatics*, Vol. 15, No. 1, pp. 72–84, 1999.

[Tree 02a] D. R. Tree, J. M. Shulman, R. Rousset, M. Scott, D. Gubb, and J. D. Axelrod. "Prickle mediates feedback amplification to generate asymmetric planar cell polarity signaling". *Cell*, Vol. 109, No. 3, pp. 371–381, 2002.

[Tree 02b] D. R. Tree, D. Ma, and J. D. Axelrod. "A three-tiered mechanism for regulation of planar cell polarity". *Seminars in Cell & Developmental Biology*, Vol. 13, No. 3, pp. 217–224, June 2002.

[Tsan 93] E. Tsang. *Foundations of Constraint Satisfaction*. Academic Press, London and San Diego, 1993.

[Tsie 98] R. Y. Tsien. "The green fluorescent protein". *Annual Review of Biochemistry*, Vol. 67, pp. 509–544, 1998.

[Turi 52] A. M. Turing. "The chemical basis of morphogenesis". *Philosophical Transactions of the Royal Society of London Series B*, Vol. 237, No. 641, pp. 37–72, August 1952.

[Turn 04] T. Turner, S. Schnell, and K. Burrage. "Stochastic approaches for modelling in vivo reactions". *Computational Biology and Chemistry*, Vol. 28, No. 3, 2004.

[Turn 95] C. M. Turner and P. N. Adler. "Morphogenesis of Drosophila pupal wings in vitro". *Mechanisms of Development*, Vol. 52, No. 2-3, pp. 247–255, August 1995.

[Twyc 10] J. Twycross, L. R. Band, M. J. Bennett, J. R. King, and N. Krasnogor. "Stochastic and deterministic multiscale models for systems biology: an auxin-transport case study". *BMC Systems Biology*, Vol. 4, No. 34, March 2010.

[Tyso 03] J. Tyson, K. Chen, and B. Novak. "Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell". *Current Opinion in Cell Biology*, Vol. 15, No. 2, pp. 221–231, 2003.

[Usui 99] T. Usui, Y. Shima, Y. Shimada, S. Hirano, R. W. Burgess, T. L. Schwarz, M. Takeichi, and T. Uemura. "Flamingo, a seven-pass transmembrane cadherin, regulates planar cell polarity under the control of Frizzled". *Cell*, Vol. 98, No. 5, pp. 585–595, 1999.

[Usyk 03] T. P. Usyk and A. D. McCulloch. "Relationship between regional shortening and asynchronous electrical activation in a three-dimensional model of ventricular electromechanics". *Journal of Cardiovascular Electrophysiology*, Vol. 14, pp. S196–202, 2003.

[Valk 98] R. Valk. "Petri Nets as Toke Objects: An Introduction to Elementary Object Nets". In: J. Desel and M. Silva, Eds., *Application and Theory of Petri Nets 1998*, pp. 1–25, Springer-Verlag, Berlin, 1998.

[Veem 03] M. T. Veeman, D. C. Slusarski, A. Kaykas, S. H. Louie, and R. T. Moon. "Zebrafish Prickle, a modulator of noncanonical Wnt/Fz signaling, regulates gastrulation movements". *Current Biology*, Vol. 13, No. 8, pp. 680–685, April 2003.

[Vill 95] J. L. Villano and F. N. Katz. "*four-jointed* is required for intermediate growth in the proximal-distal axis in Drosophila". *Development*, Vol. 121, No. 9, pp. 2767–2777, September 1995.

[Vins 87] C. R. Vinson and P. N. Adler. "Directional non-cell autonomy and the transmission of polarity information by the frizzled gene of *Drosophila*". *Nature*, Vol. 14, No. 329, pp. 549–551, 1987.

[Vins 89] C. R. Vinson, S. Conover, and P. N. Adler. "A Drosophila tissue polarity locus encodes a protein containing seven potential transmembrane domains". *Nature*, Vol. 338, No. 6212, pp. 263–264, March 1989.

[Vohr 01] J. Vohradsky. "Neural network model of gene expression". *FASEB Journal*, Vol. 15, No. 3, pp. 846–854, 2001.

[Walk 09] D. C. Walker and J. Southgate. "The virtual cell - a candidate co-ordinator for 'middle-out' modelling of biological systems". *Briefings in Bioinformatics*, Vol. 10, No. 4, pp. 450–461, July 2009.

[Wall 01] J. B. Wallingford and R. M. Harland. "Xenopus dishevelled signaling regulates both neural and mesodermal convergent extension: parallel forces elongating the body axis". *Development*, Vol. 128, No. 13, pp. 2581–2592, July 2001.

[Wall 02] J. B. Wallingford, S. E. Fraser, and R. M. Harland. "Convergent extension: The molecular control of polarized cell movement during embryonic development". *Developmental Cell*, Vol. 2, No. 6, pp. 695–706, June 2002.

[Wang 09] Z. Wang, C. M. Birch, J. A. Sagotsky, and T. S. Deisboeck. "Cross-scale, cross-pathway evaluation using an agent-based non-small cell lung cancer model". *Bioinformatics*, Vol. 25, No. 18, pp. 2389–2396, 2009.

[Wata 04] H. Watanabe, S. Sugiura, H. Kafuku, and T. Hisada. "Multiphysics Simulation of Left Ventricular Filling Dynamics Using Fluid-Structure Interaction Finite Element Method". *Biophysical Journal*, Vol. 87, No. 3, pp. 2074–2085, 2004.

[Webe 13] M. Weber, S. Henkel, S. Vlaic, R. Guthke, E. van Zoelen, and D. Driesch. "Inference of dynamical gene-regulatory networks based on time-resolved multi-stimuli multi-experiment data applying NetGenerator V2.0". *BMC Systems Biology*, Vol. 7, No. 1, p. 1, 2013.

[Wehr 98] M. Wehrli and A. Tomlinson. "Independent regulation of anterior/posterior and equatorial/polar polarity in the *Drospholia* eye; evidence for the involvement of Wnt signaling in the equatorial/polar axis". *Development*, Vol. 125, No. 8, pp. 1421–1432, 1998.

[Weim 02] J. Weimar. "Cellular Automata Approaches to Enzymatic Reaction Networks". In: *International Conference on Cellular Automata for Research and Industry, ACRI, LNCS*, 2002.

[Weng 99] G. Weng, U. S. Bhalla, and R. Iyengar. "Complexity in biological signaling systems". *Science*, Vol. 284, No. 5411, pp. 92–96, April 1999.

[Wint 01] C. G. Winter, B. Wang, A. Ballew, A. Royou, R. Karess, J. D. Axelrod, and L. Luo. "Drosophila Rho-associated kinase (Drok) links Frizzled-mediated planar cell polarity signaling to the actin cytoskeleton". *Cell*, Vol. 105, pp. 81–91, 2001.

[Wish 05] D. S. Wishart, R. Yang, D. Arndt, P. Tang, and J. Cruz. "Dynamic cellular automata: an alternative approach to cellular simulation". *In Silico Biololgy*, Vol. 5, No. 2, pp. 139–161, 2005.

[Wolf 98] T. Wolff and G. M. Rubin. "*strabismus*, a novel gene that regulates tissue polarity and cell fate decisions in Drosophila". *Development*, Vol. 125, No. 6, pp. 1149–1159, March 1998.

[Wong 03] H.-C. Wong, A. Bourdelas, A. Krauss, H.-J. Lee, Y. Shao, D. Wu, M. Mlodzik, D.-L. Shi, and J. Zheng. "Direct binding of the pdz domain of *dishevelled* to a conserved internal sequence in the c-terminal region of *fizzled*". *Molecular Cell*, Vol. 12, No. 5, pp. 1251–1260, 2003.

[Wong 93] L. L. Wong and P. N. Adler. "Tissue polarity genes of *Drosophila* regulate the subcellular location for prehair initiation in pupal wing cells". *Journal of Cell Biology*, Vol. 123, pp. 209–220, 1993.

[Wu 09] J. Wu and M. Mlodzik. "A quest for the mechanism regulating global planar cell polarity of tissues". *Trends of Cell Biology*, Vol. 19, No. 7, pp. 295–305, 2009.

[Wurt 00] J. U. Wurthner, A. K. Mukhopadhyay, and C.-J. Peimann. "A cellular automaton model of cellular signal transduction". *Comp. in Bio. and Med*, Vol. 30, No. 1, pp. 1–21, 2000.

[Yang 02] C. Yang, J. D. Axelrod, and M. A. Simon. "Regulation of Frizzled by Fat-like cadherins during planar polarity signaling in the *Drosophila* compound eye". *Cell*, Vol. 108, pp. 675–688, 2002.

[yosh 99] N. yoshida and M. C. Liberman. "Stereociliary anomaly in the guinea pig: effects of hair bundle rotation on cochlear sensitivity". *Hearing Research*, Vol. 131, No. 1-2, pp. 29–38, May 1999.

[Zall 07] J. A. Zallen. "Planar polarity and tissue morphogenesis". *Cell*, Vol. 129, pp. 1051–1063, 2007.

[Zeid 00] M. P. Zeidler, N. Perrimon, and D. I. Strutt. "Multiple roles for *four-jointed* in planar polarity and limb patterning". *Developmental Biology*, Vol. 228, No. 2, pp. 181–196, December 2000.

[Zeid 99] M. P. Zeidler, N. Perrimo, and D. I. Strutt. "Polarity determination in the *Drosophila* eye: a novel role for Unparied and JAK/STAT signaling". *Genes and Development*, Vol. 13, No. 10, pp. 1342–1353, 1999.

[Zhan 07] L. Zhang, C. Athale, and T. Deisboeck. "Development of a three-dimensional multiscale agent-based tumor model: simulating gene-protein interaction

profiles, cell phenotypes and multicellular patterns in brain cancer". *Journal of Theoretical Biology*, Vol. 244, No. 1, pp. 96–107, 2007.

[Zhan 09] L. Zhang, C. G. Strouthos, Z. Wang, and T. S. Deisboeck. "Simulating brain tumor heterogeneity with a multiscale agent-based model: Linking molecular signatures, phenotypes and expansion rate". *Mathematical and Computer Modeling*, Vol. 49, pp. 307–319, 2009.

[Zhen 95] L. Zheng, J. Zhang, and R. W. Carthew. "*frizzled* regulates mirror-symmetric pattern formation in the *Drosophila* eye". *Development*, Vol. 121, No. 9, pp. 3045–3055, September 1995.

Appendix A

Details of the ODE models

Parameter	Value	Parameter	Value	Parameter	Value
$[Prot]_0$	2	$[Fz]_0$	1	$[FmiVang]_0$	0
$[Fmi_t]_0$	0	$[Dsh]_0$	1	$[FFD]_0$	0
$[Fz_t]_0$	0	$[Pk]_0$	1	$[FVP]_0$	0
$[Dsh_t]_0$	0	$[Vang]_0$	1	$[complex]_0$	0
$[Pk_t]_0$	0	$[Ld]_0$	1	$[Fz^{\dagger}]_0$	0
$[Vang_t]_0$	0	$[FmiFmi]_0$	0	$[mutant]_0$	0
$[Fmi]_0$	1	$[FmiFz]_0$	0		

Table A.1: Initial concentration set for the ODE models.

Parameter	Value	Parameter	Value	Parameter	Value
k_1	5.47	k_{f1}	3.9e-1	k_{p5}	1.3e-3
k_{d1}	9.7e-3	k_{f2}	8.5e-1	k_{t1}	9.6e-3
k_2	9.65	k_{f3}	8.0	k_{t2}	8.0e-3
k_3	9.65	k_{f4}	4.1	k_{t3}	7.7e-3
k_4	9.59	k_{p1}	9.6e-3	k_{t4}	6.6e-3
k_{d4}	6.6e-2	k_{p2}	8.0e-3	k_{t5}	1.3e-3
k_5	9.57	k_{p3}	7.7e-3	k_x	5.0
k_{d5}	4.9e-2	k_{p4}	6.6e-3	k^a_{mutant}	0

Table A.2: Selected parameter set for the ODE models underlying BFX model.
[a] the rate for knock-out reactions.

Model	Description
Unbiased	includes no biasing mechanisms, such as feedback loop, factor X and polarised vesicle transport. Proteins are symmetrically distributed along the cell membrane [Stru 01]
BF	includes feedback loop, which mediates the proximal-distal competition among signalling components, as a result cells tend to align their polarity such that each cell accumulates high levels of Fz on the same side of the cell and high levels of Vang on the opposite side [Axel 01, Axel 11, Gubb 82, Gubb 99, Tree 02a]
BX	includes a global directional clue (factor X), which was reported to be required to orient cell polarity with respect to the tissue axes [Adle 00, Fant 03, Lawr 99, Axel 11]
BT	includes polarised transport, where the preferential transport of Fmi, Fz and Dsh towards the distal cell membrane have been reported in [Shim 06]
BFX	includes both feedback loop and factor X
BFT	includes both feedback loop and polarised transport
BFXT	includes feedback loop, factor X and polarised transport
BXT	includes both factor X and polarised transport

Table A.3: Details of the HCPN models

Mutant clone	Description and implementations
fmi-	cells in the clone lack of Fmi. In the model, $[Fmi]_0$ equals zero and the transport of Fmi is set to zero
fz-	cells in the clone lack of Fz. In the model, $[Fz]_0$ equals zero and the transport of Fz is set to zero
dsh-	cells in the clone lack of Dsh. In the model, $[Dsh]_0$ equals zero and the transport of Dsh is set to zero
vang-	cells in the clone lack of Vang. In the model, $[Vang]_0$ equals zero and the transport of Vang is set to zero
pk-	cells in the clone lack of Pk. In the model, $[Pk]_0$ equals zero and the transport of Pk is set to zero
pk+	cells in the clone consist of overexpressed Pk. In the model, $[Fmi]_0$ is set to 10 times of the original value used in other models

Table A.4: Incorporating the mutant clones into the models

Appendix B

Detailed results of BF and BFXT models in one-to-two neighbourhood

The following provides the simulation and clustering results for BF and BFXT models as reference to the comparison of different HCPN models discussed in Chapter 5.

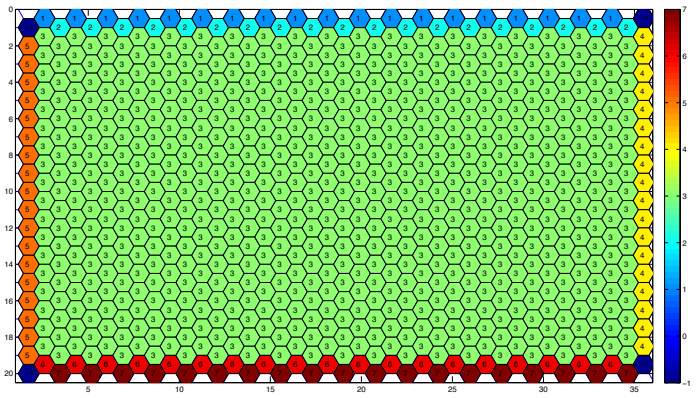

Figure B.1: Clustering for continuous simulation of BFWt and BFXTWt models.

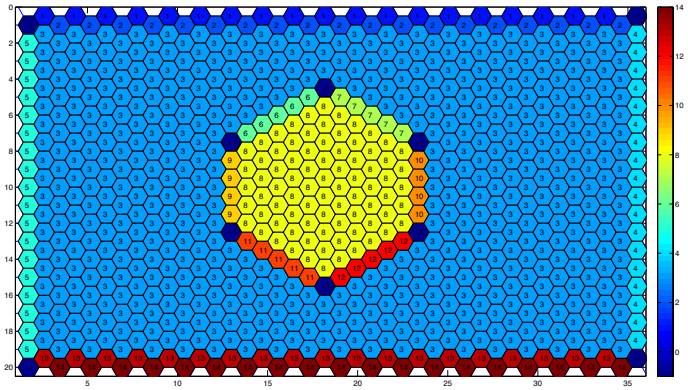

Figure B.2: Clustering for continuous simulation of BFFz and BFXTFz models.

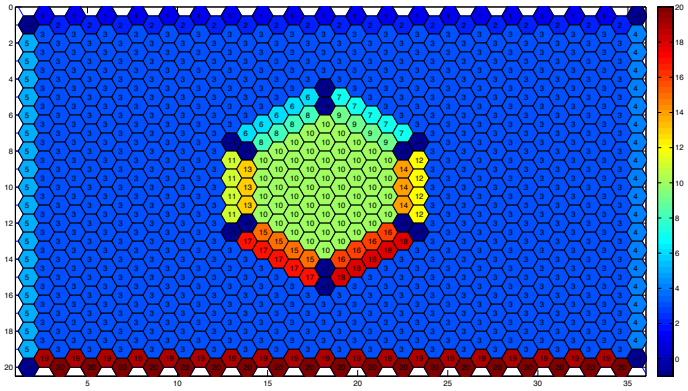

Figure B.3: Clustering for continuous simulation of BFVang and BFXTVang models.

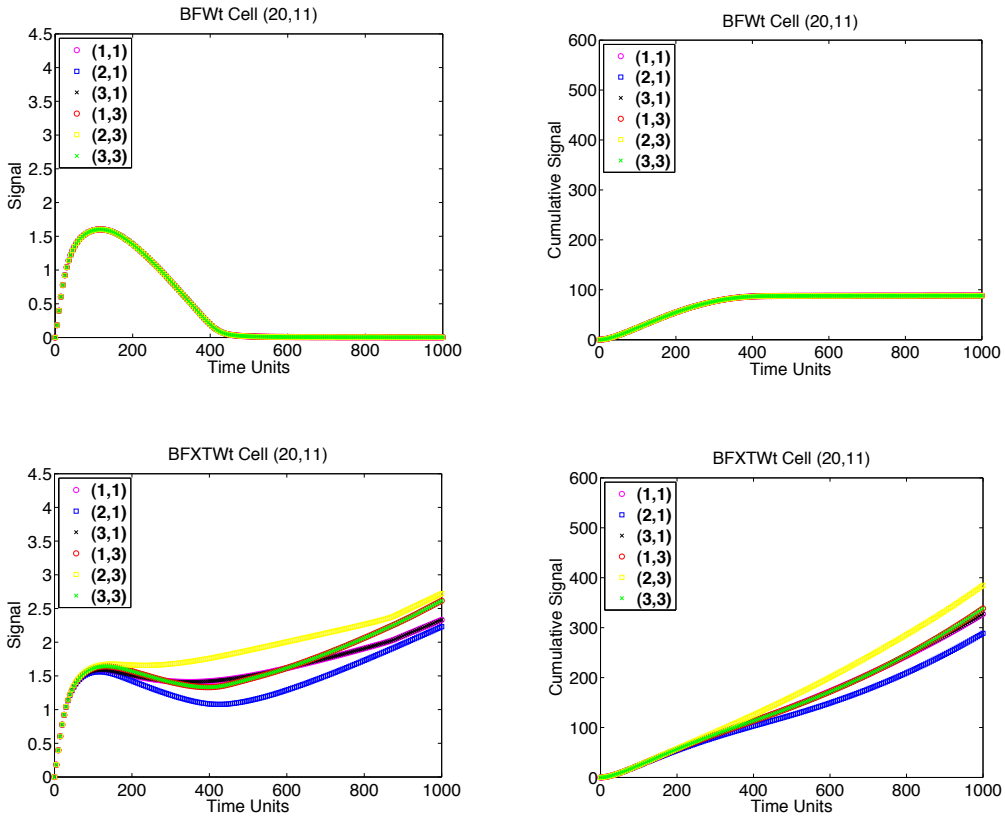

Figure B.4: BFWt and BFXTWt models: continuous simulation for a representative of wild-type cells cells, BFWt model (top left); cumulative signal for a representative of wild-type cells cells, BFWt model (top right); continuous simulation for a representative of wild-type cells cells, BFXTWt model (bottom left); cumulative signal for a representative of wild-type cells cells, BFXTWt model (bottom right). Cluster 3 in Figure B.1.

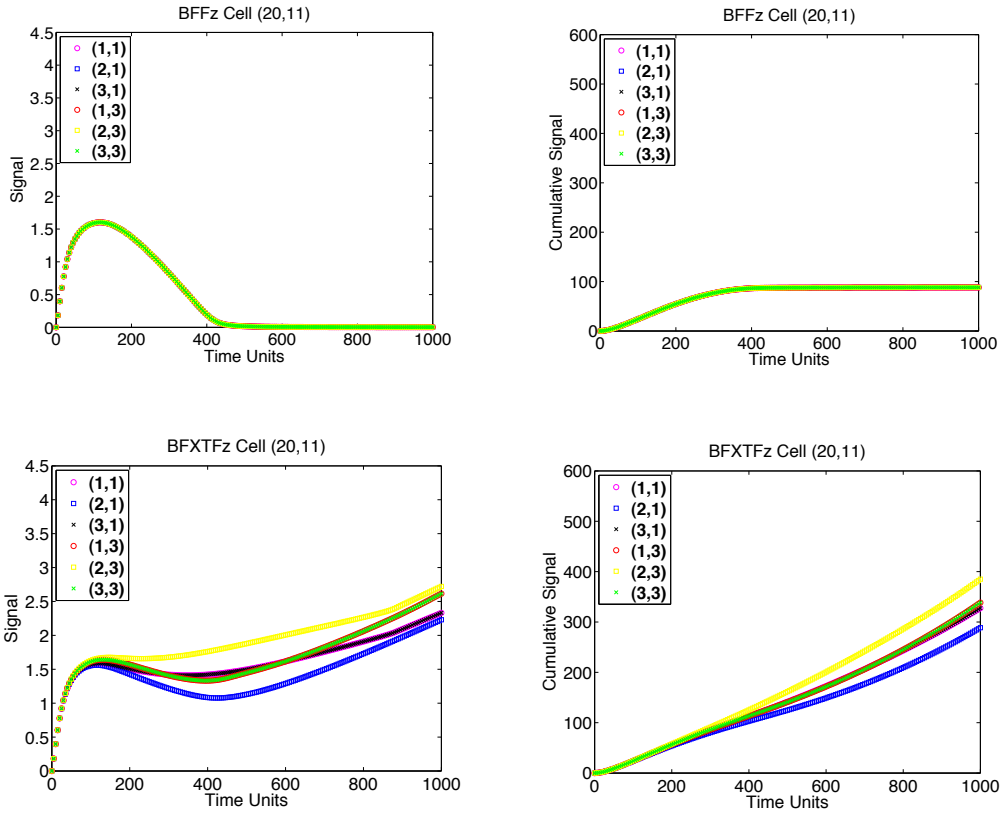

Figure B.5: BFFz and BFXTFz models: continuous simulation for a representative of wild-type cells away from the clone, BFFz model (top left); cumulative signal for a representative of wild-type cells away from the clone, BFFz model (top right); continuous simulation for a representative of wild-type cells away from the clone, BFXTFz model (bottom left); cumulative signal for a representative of wild-type cells away from the clone, BFXTFz model (bottom right). Cluster 3 in Figure B.2.

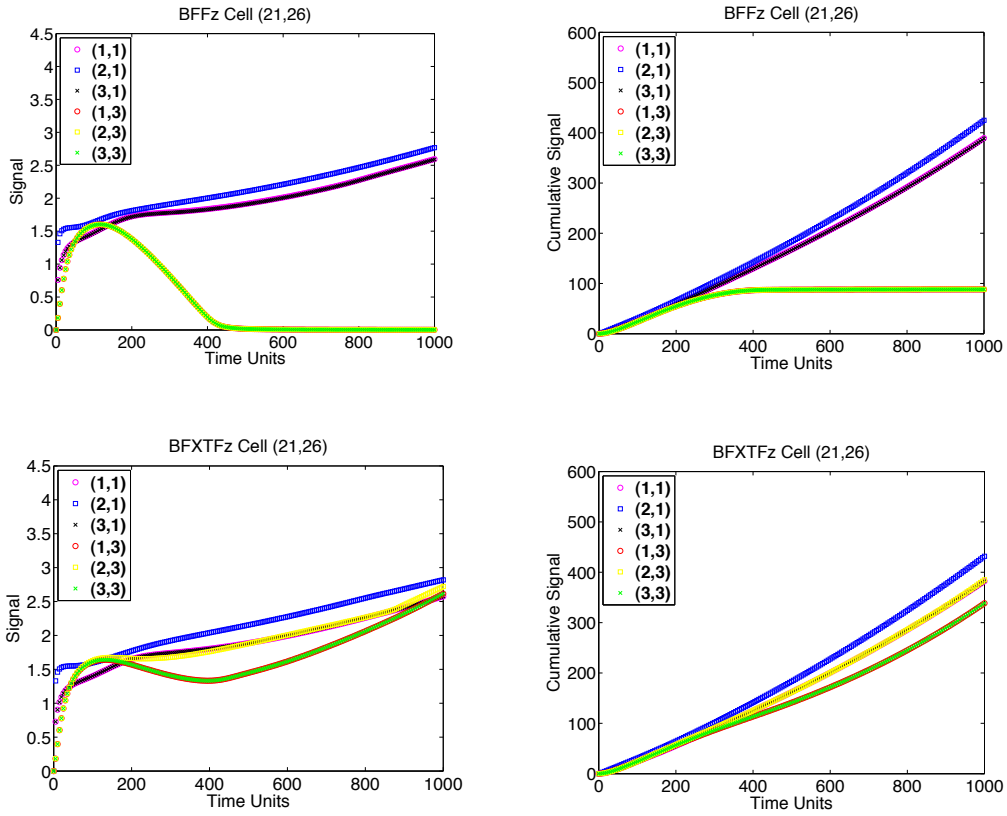

Figure B.6: BFFz and BFXTFz models: continuous simulation for a representative of cells distally adjacent to the clone, BFFz model (top left); cumulative signal for a representative of cells distally adjacent to the clone, BFFz model (top right); continuous simulation for a representative of cells distally adjacent to the clone, BFXTFz model (bottom left); cumulative signal for a representative of cells distally adjacent to the clone, BFXTFz model (bottom right). Cluster 10 in Figure B.2.

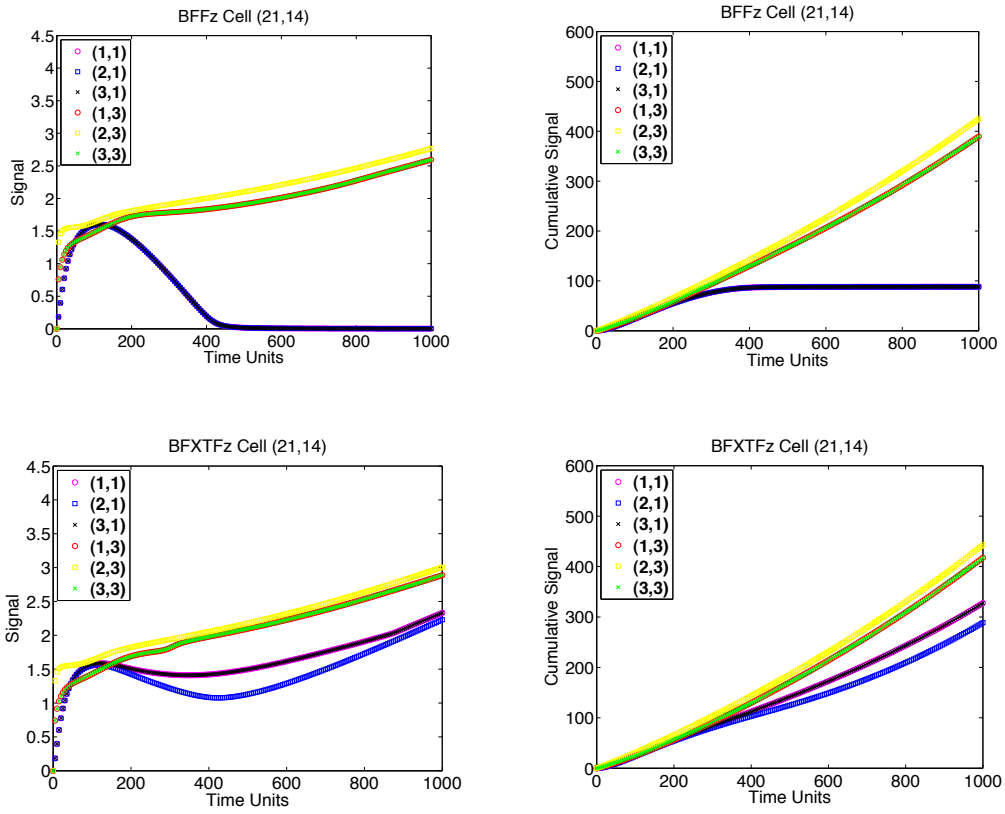

Figure B.7: BFFz and BFXTFz models: continuous simulation for a representative of cells proximally adjacent to the clone, BFFz model (top left); cumulative signal for a representative of cells proximally adjacent to the clone, BFFz model (top right); continuous simulation for a representative of cells proximally adjacent to the clone, BFXTFz model (bottom left); cumulative signal for a representative of cells proximally adjacent to the clone, BFXTFz model (bottom right). Cluster 9 in Figure B.2.

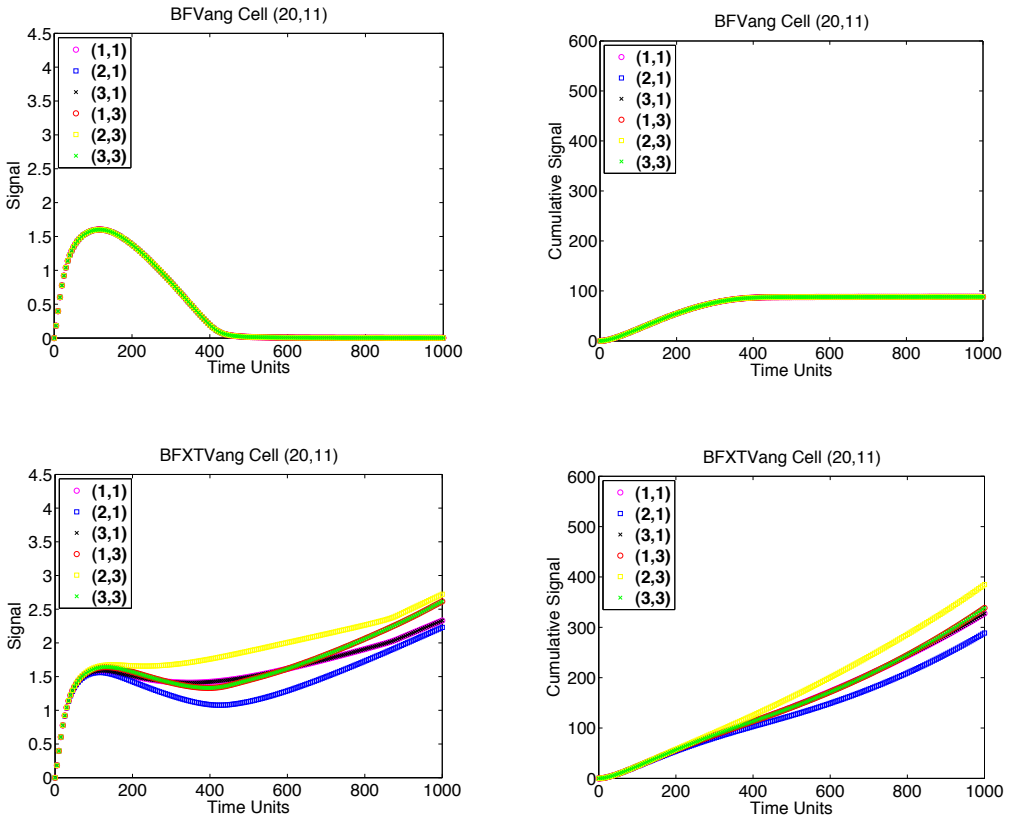

Figure B.8: BFVang and BFXTVang models: continuous simulation for a representative of wild-type cells away from the clone, BFVang model (top left); cumulative signal for a representative of wild-type cells away from the clone, BFVang model (top right); continuous simulation for a representative of wild-type cells away from the clone, BFXTVang model (bottom left); cumulative signal for a representative of wild-type cells away from the clone, BFXTVang model (bottom right). Cluster 3 in Figure B.3.

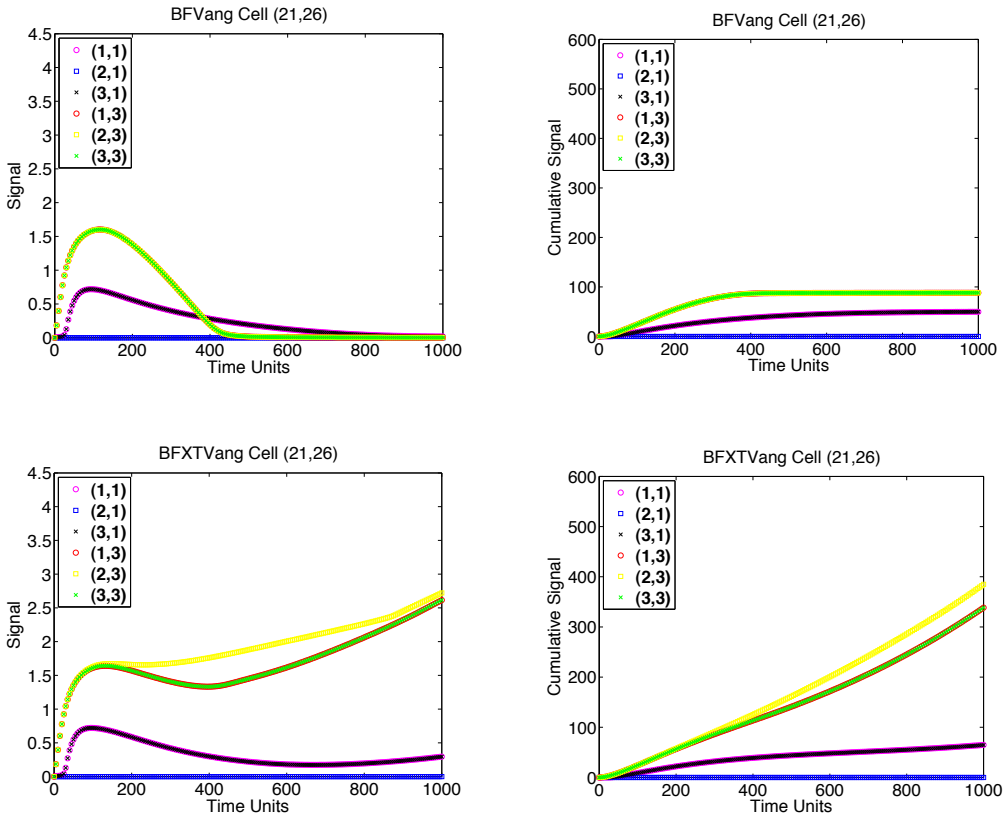

Figure B.9: BFVang and BFXTVang models: continuous simulation for a representative of cells distally adjacent to the clone, BFVang model (top left); cumulative signal for a representative of cells distally adjacent to the clone, BFVang model (top right); continuous simulation for a representative of cells distally adjacent to the clone, BFXTVang model (bottom left); cumulative signal for a representative of cells distally adjacent to the clone, BFXTVang model (bottom right). Cluster 12 in Figure B.3.

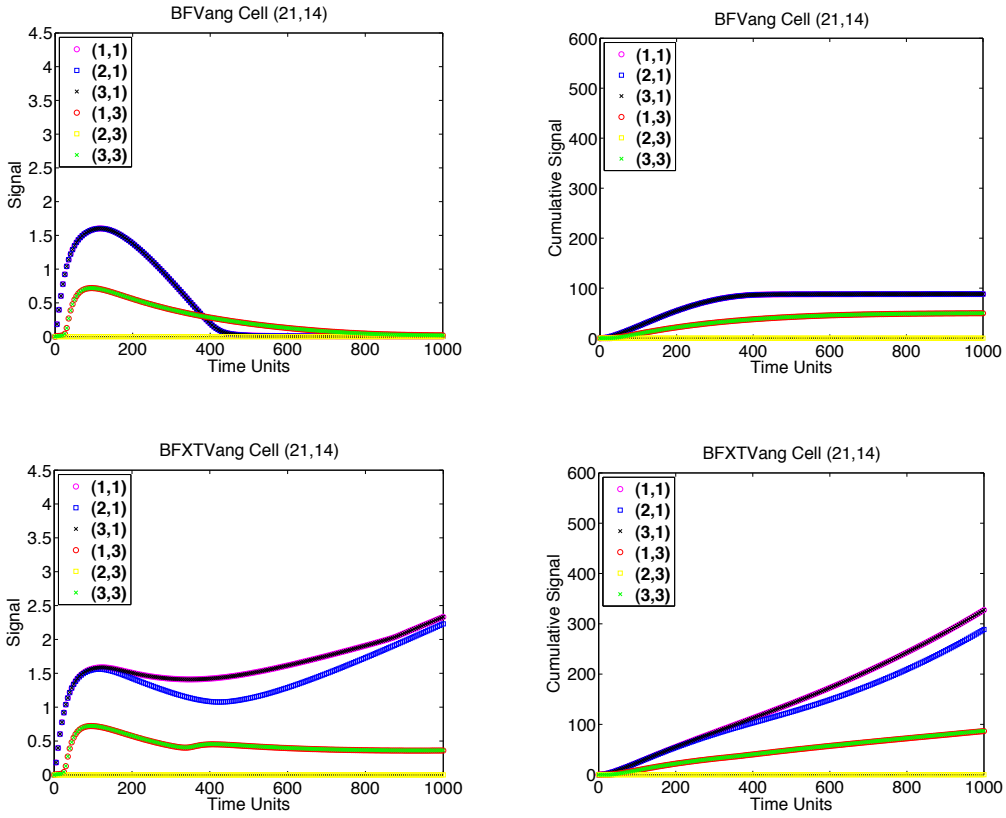

Figure B.10: BFVang and BFXTVang models: continuous simulation for a representative of cells proximally adjacent to the clone, BFVang model (top left); cumulative signal for a representative of cells proximally adjacent to the clone, BFVang model (top right); continuous simulation for a representative of cells proximally adjacent to the clone, BFXTVang model (bottom left); cumulative signal for a representative of cells proximally adjacent to the clone, BFXTVang model (bottom right). Cluster 11 in Figure B.3.

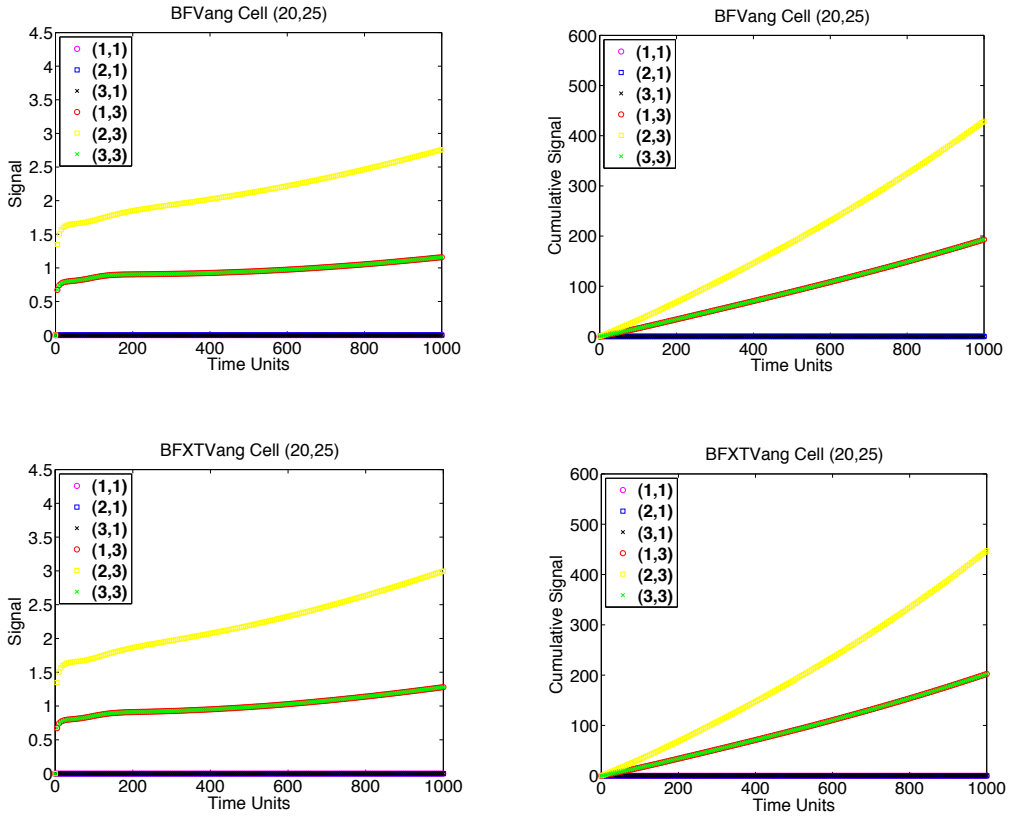

Figure B.11: BFVang and BFXTVang models: continuous simulation for a representative of mutant cells at the distal boundary of the clone, BFVang model (top left); cumulative signal for a representative of mutant cells at the distal boundary of the clone, BFVang model (top right); continuous simulation for a representative of mutant cells at the distal boundary of the clone, BFXTVang model (bottom left); cumulative signal for a representative of mutant cells at the distal boundary of the clone, BFXTVang model (bottom right). Cluster 14 in Figure B.3.

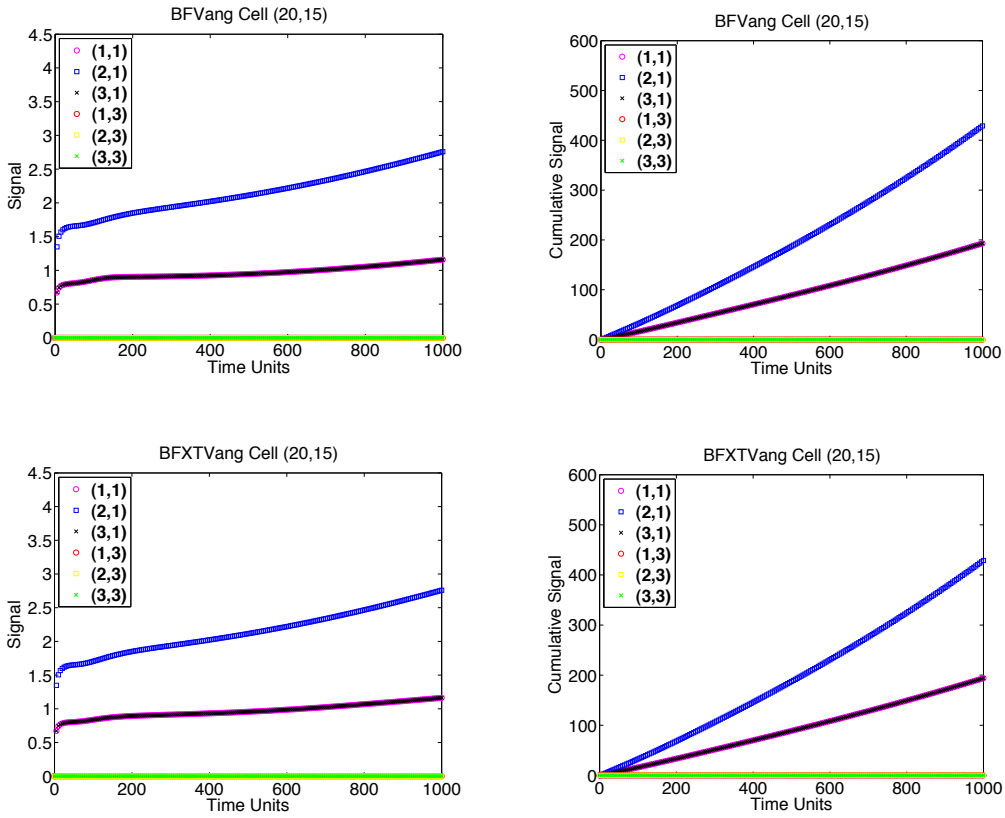

Figure B.12: BFVang and BFXTVang models: continuous simulation for a representative of mutant cells at the proximal boundary of the clone, BFVang model (top left); cumulative signal for a representative of mutant cells at the proximal boundary of the clone, BFVang model (top right); continuous simulation for a representative of mutant cells at the proximal boundary of the clone, BFXTVang model (bottom left); cumulative signal for a representative of mutant cells at the proximal boundary of the clone, BFXTVang model (bottom right). Cluster 13 in Figure B.3.