A Multiscale Approach to Modelling Planar Cell Polarity in Drosophila Wing using Hierarchically Coloured Petri Nets

Qian Gao *
School of Information Systems, Computing and Mathematics
Brunel University
Uxbridge, UK
qian.gao@brunel.ac.uk

Fei Liu *
Computer Science Department
Brandeburg University of Technology Cottus
Germany
liu@informatik.tu-cottus.de

David Gilbert
School of Information Systems, Computing and Mathematics
Brunel University
Uxbridge, UK
david.gilbert@brunel.ac.uk

Monika Heiner †
School of Information Systems, Computing and Mathematics
Brunel University
Uxbridge, UK
monika.heiner@brunel.ac.uk

David Tree
School of Health Sciences and Social Care
Brunel University
Uxbridge, UK
david.tree@brunel.ac.uk

ABSTRACT

Modelling across multiple scales is a current challenge in Systems Biology, especially when applied to multicellular organisms. In this paper we present an approach to model at different spatial scales, using the new concept of hierarchically coloured Petri Nets (HCPN). We apply 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) signalling in Drosophila wing. We illustrate different levels of abstraction that can be used in order to assist the systematic modelling of such a complex system involving intra- and inter-cellular signalling mechanisms, and we provide a design pattern for similar modelling problems. Our initial model describes normal, wild-type PCP signalling, and we illustrate the power of our approach by easily adapting it to various tissue sizes and to describe the phenotype of a well-documented genetic mutation in Drosophila. We have performed a series of analyses on our models which require computational experiments over very large underlying models. All results are reproducible.

Keywords

Hierarchically coloured Petri nets, multiscale modelling, continuous and stochastic simulation, planar cell polarity.

*Contributed equally to this paper.
†On sabbatical leave from Brandenburg University of Technology, Cottbus, Germany.

CMSB '11, September 21-23, 2011 Paris, FR

1. INTRODUCTION

With the rapid growth of data being generated in the biological field, it has become necessary to organise the data into coherent models that describe system behaviour, which are subsequently used for simulation, analysis or prediction. Modelling biological systems beyond one spatial scale introduces a series of challenges which should be addressed. These include:

Repetition of components – for example the need to describe multiple cells each of which has a similar definition.

Variation of components – sets of similar components with defined variations, e.g. mutants.

Organisation of components – for example how cells are organised into regular or irregular patterns over spatial networks in one, two or three dimensions.

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.

Mobility – for example transport of components within a system, or actively motile cells.

Hierarchical organisation – enabling the description of (possibly repeated) components which contain repeated sub-components. For example, cells containing several compartments. This feature enables the use of abstraction regarding the level of detail used to describe components.

In this paper we have chosen the biological example of Planar Cell Polarity (PCP) signalling in Drosophila wing, which illustrates several of these issues. The epithelial cells in this organ 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 (transmembrane) signalling produces a global effect over the entire organ. Our model of PCP includes the repetition of components in a two-level hierarchy, permitting abstraction at the level of cells, with a static two-dimensional organisation which is different at each level. The first level is that of cells in a rectangular honeycomb matrix, representing the epithelium tissue, and the second level being intra-cellular organisation represented by logical compartments within one cell in a rectangular lattice. Moreover we model variations among cells in the form of patches of mutant cells which lack a specific signalling protein.

Modelling in biology tends to emphasise molecular details. However in biological networks that involve more than a few components the typical situation is that many details are unknown, and it is imperative to devise an approach that can be insightful and predictive even in the absence of complete knowledge. Our strategy was based on building first an abstract model of PCP which attempts to identify the key biological aspects (e.g. formation of intercellular complexes), and then constructing a more detailed, but still simplified model which parameterises the many unknowns.

A large variety of modelling approaches have already been applied to model a wide array of biological systems (see [12] for a review). Among them, Petri nets are particularly suitable for representing and modelling the concurrent and asynchronous behaviour of biological systems. However, standard Petri nets do not readily scale to meet the challenges addressed above, and current attempts to simulate biological systems by standard Petri nets have been mainly restricted so far to relatively small models.

In contrast, Coloured Petri Nets (CPN) overcome the constraints of standard Petri nets by allowing the modelling of large-scale systems in a compact, parameterised and scalable way. Thus coloured Petri nets should be suitable to address the challenges in the modelling of PCP, motivating us to use them to realise our modelling methodology.

In this paper, we introduce a general modelling principle demonstrated by a specific example which illustrates many of the spatial multiscale problems in biological systems, and for this purpose we have developed the new modelling concept of Hierarchically Coloured Petri Nets (HCPN). Our approach is illustrated by a complex and challenging case study, which requires computational experiments over very large underlying models.

This paper is structured as follows: in Section 2 we introduce the biological background of planar cell polarity, followed by Section 3 on related work. Section 4 briefly describes coloured Petri Nets and in Section 5 we present our modelling approach by means of the PCP case study, followed by Section 6 on the analysis of our PCP model, and finally the conclusion.

2. PLANAR CELL POLARITY

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 [30]). The signalling mechanisms underlying PCP have been studied most extensively in the epithelia of the fruit fly

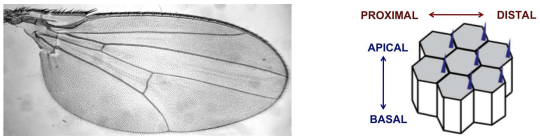

Figure 1: Drosophila: whole wing (left); scheme of hexagonal cells with hairs (right).

Drosophila melanogaster including the wing, the abdomen, the eye, and the bristles of the thorax. 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 [37]). 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 models of PCP signalling.

The adult Drosophila wing comprises about 300,000 hexagonal cells each of which contains a single hair pointing in an invariant distal direction, see Figure 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. Preceding this 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 is found at both the proximal and distal cell membrane, Fz and Dsh are found exclusively at the distal cell membrane and Vang and Pk 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. 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. 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 [31]). The result is a polarisation of individual cells and local alignment of polarity between neighbouring cells. A proposed feedback loop mechanism was introduced by Tree et al. [32] which provides a possible insight into how these proteins achieve their asymmetric distribution while distally localized Fz and Dsh designating the prehair formation. It is believed that this feedback loop mediates a competition between proximal and distal proteins between adjacent surfaces of neighbouring cells. More details of the cellular machinery can be found in the latest review paper [3]. Thus we will use Flamingo-Frizzled-Dishevelled complexes (FFD) as a proxy for hairs in our research.

In this paper, we use HCPN to model PCP signalling in a generic setting that encompasses a broad class of specific models, ranging from a single cell model to a multi-cellular model. To this end, we have developed a model for the

generation of PCP to investigate the signalling by simulating the dynamic behaviour of the PCP proteins and complexes.

Initially we wish to recapitulate the phenotypes of all known mutant conditions, both loss and gain of function. We hope to be able to address questions about the mechanisms underlying the polarising signal, the dynamics of signalling by the individual components and the signals downstream of the PCP proteins which orchestrate the ultimate morphological manifestation of planar polarity. Ultimately we hope our model will make predictions about the mechanisms underlying the PCP signalling process which will be testable in a biological laboratory.

3. RELATED WORK

Modelling 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. [1] (extended in [26]) 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 [20] (applied to the Drosophila eye in [19]). Schamberg et al. [29] built two models by applying reaction-diffusion equations to study the influence of feedback loops, 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. Burak et al. [6] 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. See [3] for an overall review of this field.

These models lack 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 us to systematically construct large scale mathematically tractable models in which cell geometry is clearly formalised. Compared with ODEs and PDEs, hierarchically coloured Petri nets are more intuitive for users who do not have much knowledge about modelling. In addition, we include new aspects of signalling, such as polarised transport of molecules, which have not been considered in previous models.

Coloured Petri nets in Systems Biology. While there is a lot of reported work on the application of different classes of standard Petri nets to a variety of biochemical networks, see [5] for a recent review, there are only a few which take advantage of the additional power and ease of modelling offered by CPN.

Existing studies usually resort to `Design/CPN` [7] or its successor `CPN Tools` [17] in order to model and analyse biological systems. However neither tool was not specifically designed with the requirements of Systems Biology in mind. Thus they are not suitable for many applications, e.g. they do not directly support stochastic or continuous modelling, nor the simulative analysis of the models by stochastic or deterministic simulation. The approach taken in [9, 21] to overcome this problem is to encode the concentration of species as coloured tokens in order to implement continuous simulation in the given net annotation language ML. While this is a nice exercise in demonstrating the power of the annotation language, the burden to implement standard simulation algorithms is left to the modeller.

Coloured Petri nets have been used for qualitative modelling and analysis in [25] to predict pathological phenotypes based on genetic mutations and in [33] to model signal transduction networks. Here, colours encode mutations of the modelled molecules or distinguish between different molecules via their identifiers (colours). Colours have also been used to discriminate metabolites which follow different T-invariants (elementary flux modes) [14, 35, 28]. Benefits of coloured Petri nets in a stochastic setting were first demonstrated in [4] using a very simple epidemic model.

In our work, we use colour not only to express repetition, but also to encode (spatial) locality. We will show that the standard colouring concept can be further enhanced by hierarchically organised colours, which are able to directly reflect the hierarchical organisation of the objects modelled. We demonstrate our approach by a case study which yields a design pattern for similar modelling problems.

4. MODELLING LANGUAGE

4.1 Coloured Petri nets

Coloured Petri nets (CPNs) [15, 16] 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 we are going to exploit in this paper.

Syntax. CPNs are – as are standard Petri nets – made of places, transitions, and arcs. In systems biology, places typically represent species (chemical compounds), while transitions represent any kind of chemical reactions or transport steps. In this paper 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. 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.

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 of the ex-

isting 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 net define the colour sets of the coloured 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 actual net structure of the underlying reaction network.

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 [16]. 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.

Formal definition. The central CPN notion are the colour sets. These are in our case discrete finite data types, which are built starting from some pre-defined basic types, such as Boolean and integer, using some type constructors, such as subset or product. If A is a non-empty (colour) set, then a multiset over A is denoted as A_{MS}.

Moreover, there are different types of expressions, arc expressions, guards and expressions for defining initial markings. An expression is built up from constants, variables, and operation symbols. It is not only associated with a particular colour set, but also written in terms of a predefined syntax. In the following, we denote by EXP the set of expressions that comply with a predefined syntax. Then the formal definition for coloured Petri nets is as follows [15, 16].

Definition 1. A *coloured Petri net* is a tuple $N = [P, T, F, \Sigma, c, g, f, m_0]$, where:

- P is a finite, non-empty set of places.
- T is a finite, non-empty set of transitions.
- F is a finite, non-empty set of directed arcs.
- Σ is a finite, non-empty set of colour sets.
- $c : P \rightarrow \Sigma$ is a colour function that assigns to each place $p \in P$ a colour set $c(p) \in \Sigma$.
- $g : T \rightarrow EXP$ is a guard function that assigns to each transition $t \in T$ a guard expression of Boolean type.
- $f : F \rightarrow EXP$ is an arc function that assigns to each arc $a \in F$ an arc expression of a multiset type $c(p)_{MS}$, where p is the place connected to the arc a.
- $m_0 : P \rightarrow EXP$ is an initialisation function that assigns to each place $p \in P$ an initialisation expression of a multiset type $c(p)_{MS}$.

Furthermore, we allow special arcs, e.g. read arcs or inhibitor arcs. If transitions are additionally associated with stochastic (deterministic) firing rates, as discussed in [10], we will get coloured stochastic (continuous) Petri nets. 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 [22, 23].

4.2 Hierarchically coloured Petri nets

CPNs 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, we can represent one cell 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.

For this purpose, we adopt the notation that 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 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 we may embed a component in a honeycomb (hexagonal) lattice by imposing the requisite constraints over a rectangular grid. Guards over transitions permit the description of the patterns of connection allowed between cells.

Furthermore if we want to describe regular organisation within a cell, we can extend this concept 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, we get a sequence of tuples, each tuple referring to the address within a certain level. We call this concept hierarchically structured colours, and the corresponding net class Hierarchically Coloured Petri Nets (HCPN).

In order to support the description of hierarchically organised systems, we extend the notation for colour sets 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, we will give its position 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, we can denote each level 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 $wing_{cs} \cdot cell_{cs} \cdot subcellular\text{-}location_{cs}$. We further expand our notation for the guards so

that they are associated with the level at which they operate.

In summary, hierarchically structured colours are useful to express repetition and (spatial) locality as we will demonstrate in our case study.

5. MODELLING APPROACH

Overview. We first model each cell as a (standard) Petri net, initially at a highly abstract level with virtual compartments to detect PCP asymmetry. Because Drosophila wing cells form a regular honeycomb lattice, we impose this organisation at the top level of the model as a hexagonal array of cells, see Figure 2. We get the HCPN model by step-wise colouring this spatial information. Finally we create a refined description of a cell and reuse the colouring pattern established by means of the abstract model.

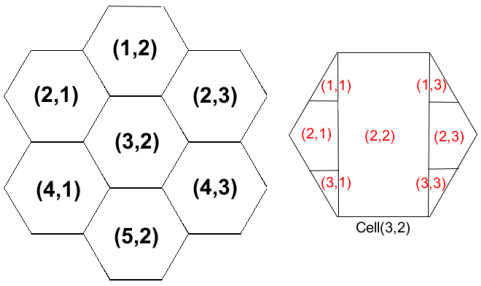

Figure 2: Drosophila wing epithelial cells. (a) A fragment of the wing tissue; the coordinates (labelled in black) in each cell represent the position in the honeycomb lattice; (b) Cell (3,2) with 7 virtual compartments.

Abstract Petri net model for a single cell. Our initial model of the wing epithelial cell, illustrated in Figure 3, is a high-level representation of the network in order to establish the colour sets required.

We first sub-divide each cell into four spatial regions (Figure 3): (1) the extracellular space (labelled as communication), where the intercellular complexes form, (2) the proximal cell margin (left-hand side of each cell) in order to process intercellular signal between two neighbouring cells, (3) production (read arcs denoting an infinite supply) and intracellular transport of core species, and (4) the distal cell margin (right-hand side of each cell).

In order to facilitate the detection of PCP asymmetry, we then partition each cell into seven *virtual compartments*, three each for the proximal and distal membrane edges, and one compartment for the cytosol, see cell $(3,2)$ in Figure 2, right. It should be noted that our abstract model is an extremely simplified version of PCP to begin with, which only includes essential components and structure and eliminates the duplication of molecular species (places) at the distal and proximal edges of a cell. Thus, in this model, the polarity will arise by the asymmetrical distribution of proteins at the distal and proximal edges of each cell together with the intercellular communication.

HCPN model for honeycomb lattice of cells. First, we define two constants M, N and a two-dimensional colour set ($CS1$) representing a rectangular $M \times N$ grid, and select that subset which denotes the coordinates of the hexagonally packed cells (CS_Cell), see Figure 2. At this level of hierarchy (wing tissue comprising folded cells) we obtain a HCPN model, which has a similar structure to that of Figure 3, but each place has assigned the colour set CS_Cell.

Table 1: Declarations for the abstract and refined HCPN models. Here "cs" denotes "color set", "var" "variable", "con" "constant" and "fun" "function".

Type	Declaration
con	$int : M = 15, N = 15, R = 3, C = 3;$
cs	Row = int with $1 - M$;
cs	$Column$ = int with $1 - N$;
cs	$CS1$ = product with $Row \times Column$;
cs	$CS\_Cell = CS1$ with $x\%2 = 1\&y\%2 = 0 \vert x\%2 = 0\&y\%2 = 1$;
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	$CSdistal = CS2$ with $b = 3$;
cs	$CSproximal = CS2$ with $b = 1$;
cs	$CSmiddle = CS2$ with $a = 2\&b = 2$;
cs	CSI = int with $1 - 2$;
cs	$CS3$ = product with $CSproximal \times CSI$;
cs	$CSproximalInter = CS3$ with $r = 2\&a = 2 \vert r = 1$;
var	$x : Row, y : Column,$ $a : ComR, b : ComC, r : CSI$;
fun	$CSproximal\ NW$ $(Row\ x, Column\ y, ComR\ a, ComC\ b, CSI\ r)$ {[(!(x=1|y=1))&(r=1&a=1&b=1|r=2&a=2&b=1)] ((x-1,y-1),(a+1,b+2));}
fun	$CSproximal\ SW$ $(Row\ x, Column\ y, ComR\ a, ComC\ b, CSI\ r)$ {[(!(x=M|y=1))&(r=2&a=2&b=1|r=1&a=3&b=1)] ((x+1,y-1),(a-1,b+2));}
fun	$bool\ MutReg(Row\ x, Column\ y)$ {x>=4&x<=8&y>=3&y<=7;}

Next we assign a colour to each of the seven virtual compartments of a cell. We do this by using a 3×3 matrix (CS_ComP) and ignoring colours (2,1) and (2,3) so that the proximal compartments are (1,2), (2,2) and (2,3), the middle compartment is (2,2), and the distal compartments are (3,1), (3,2) and (3,3). We combine information about cell and compartment locality by defining $CS2$, and we introduce three subsets $CSdistal$, $CSproximal$ and $CSmiddle$ of $CS2$ to more clearly represent the components of a cell in a specific region, i.e. the distal, proximal or middle region. The colour sets we define are hierarchical, so we can locate each place in terms of the coordinates $((x, y), (a, b))$, where (x, y) denotes the position of a cell and (a, b) denotes the position of a compartment within that cell.

We continue folding using these colours to obtain a more compact HCPN model (a tissue of cells comprising virtual compartments). This is achieved by folding the three proximal compartments into one, similarly for the three distal compartments, by assigning the colour sets $CSproximal$ and $CSdistal$ to its places. The central compartment (2,2) is denoted by the colourset $CSmiddle$.

Finally, we fold the two similar communication components (the red transitions) in compartment $(2, 1)$ into one.

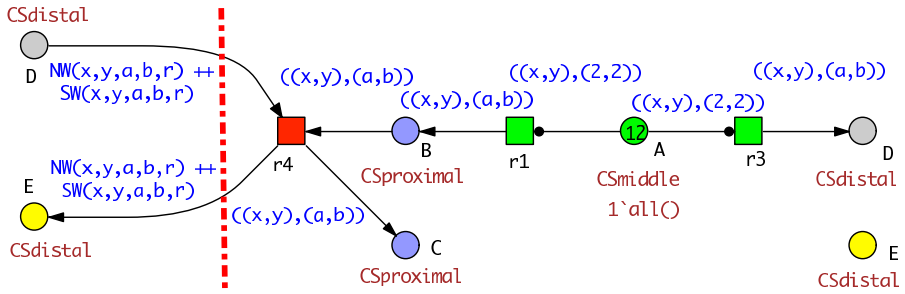

Figure 4: Abstract HCPN model describing cells with seven compartments in a 2-D matrix. Places D and E are logical nodes which are in the distal compartments of each cells. See Table 1 for all declarations. Dotted red line indicates proximal cell boundary.

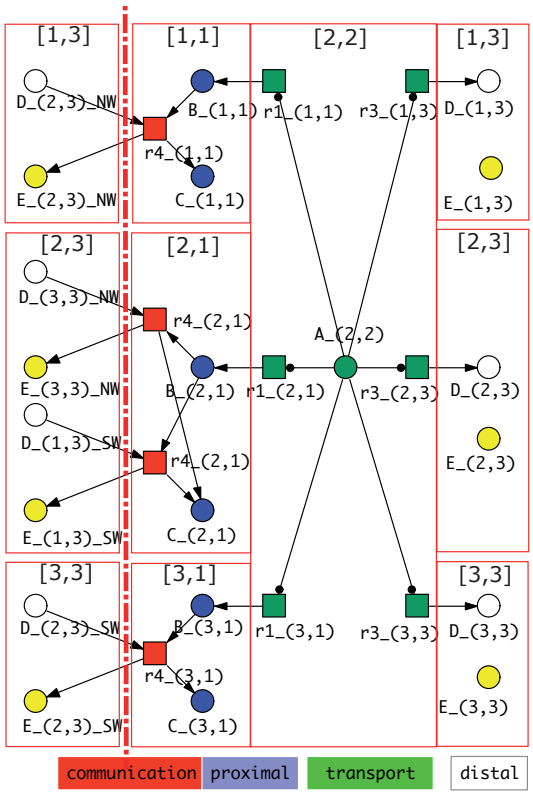

Figure 3: Abstract PN for a single cell, illustrating the four spatial regions (coloured for illustration purposes only): communication, proximal, transport and distal, and the seven virtual compartments (labelled $[1,1]$, $[2,1]$,..., $[3,3]$). Each place or transition belongs to a specific compartment, indicated by a tuple of coordinates given as a suffix in place or transition names. NW and SW denote the two left neighbours of the current cell. Dotted red line indicates proximal cell boundary.

For this we define a colour set CSI. Then we define a product colour set $CS3$ based on CSI and $CSproximal$, and one of its subsets $CSproximalInter$.

In the following, we describe the necessary steps to construct the compact HCPN model. Having defined the colour sets, we create variables that are used in transition guards and arc expressions. In the six virtual compartments for the proximal and distal cell edges, each arc is assigned an expression which includes two tuples of coordinates $((x,y),(a,b))$, meaning that the arc links the associated place to a particular transition in the (a,b) compartment of the (x,y) cell. In the middle virtual compartment, the arc expression changes to $((x,y),(2,2))$. In order to represent the neighbourhood between adjacent cells, we define two neighbour functions, NW and SW, denoting two left neighbours of a cell. Finally, we get the HCPN model for PCP, illustrated in Figure 4. See Table 1 for all declarations of this coloured model.

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.

In summary, the procedure to construct a HCPN model for a multi-cellular tissue with 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.

A refined model of PCP. We develop a more detailed PCP model by refining the model of Figure 4, which is available at `http://people.brunel.ac.uk/~cspgqqg/`. The refined model of PCP signalling begins with a set of six core proteins (Fz, Dsh, Vang, Pk, Fmi and Ld, the experimentally undefined directional signal that biases PCP signalling). In total each cell comprises 102 transitions and 76 places. It is based on by models Amonlirdviman et al. [1] and Le Garrec et al. [20], extended by polarised transport of proteins. In our model, we include cells, membranes, nuclei, cytoplasm and the extracellular space. We also include the intracellular inhibitory loops which amplifies the cell polarity as a result of which each cell accumulates high levels of Fz on one side of the cell and high levels of Vang on the opposite side [18].

To construct the refined model we use the same declarations given in Table 1 as those used for Figure 4; thus, we do not need to start from scratch. We group transitions and places into different regions and virtual compartments, and then assign the same colour sets to each region or compartment as done for Figure 4, likewise for arc expressions. Additionally, we define a function denoted as $MutReg$ (see

Table 1) which will be used to localise mutant cells on the tissue to model biological mutant clones. Thus, with this refined model we can not only perform simulations of PCP signalling in normal, wild-type cells but also on patches of mutant cells in a wild-type background.

6. ANALYSIS

HCPNs enjoy a large variety of analysis techniques, ranging from informal animation to formal static or dynamic analysis techniques. In the following, we confine ourselves to simulative methods to analyse quantitative counterparts of our refined PCP model, which we get by assigning to each transition a rate function following mass-action kinetics. The kinetic parameters have been optimised by using simulated annealing with wet-lab time series target data [2]. It takes approximately an hour to run 100,000 iterations using our optimisation program. Each simulation begins with five of the key proteins homogeneously distributed on the membranes. To simulate wild-type and different mutant conditions, we consider individual marking sets, function sets, and parameter sets (which are maintained within one model file).

This quantitative model can be equally read as a stochastic or continuous model, with appropriate scaling of the kinetic constants, see [36], which includes in our case scaling the value of tokens to the number of levels used, see [13].

To limit the computational expense, we use our generic HCPN model to generate an in-silico tissue based on a 15*15 honeycomb grid consisting of 112 cells in total. The underlying unfolded model comprises 8,624 species (places) and 9,184 reactions (transitions), and thus the ODEs system to be analysed consists of 8,624 equations, because each species requires one equation; see Table 2. Our modelling approach can be easily applied to larger in-silico tissues (by changing two constants in the model), provided sufficient computing power is available. We conduct our analysis as a proof of principle that our model has indeed the ability to capture the given biological phenomena and make sensible predictions. All simulations run over a total time span of 0-180 reported at 100 time points. The time span represents 33 hours of PCP signalling. For performance measures see Table 2.

We start off with simulating and analysing the wild-type, before considering the effect of a patch of cells lacking Frizzled on neighbouring wild-type cells.

Experiment 1: Stochastic simulation. We use the Gillespie stochastic simulation algorithm (SSA) [11] to simulate our PCP model and observe that the model behaviour in the stochastic setting approaches the continuous behaviour for increasing level numbers and/or increasing number of averaged simulation runs. See Fig. 5 for an example; it shows an average of the stochastic behaviour of Flamingo-Frizzled-Dishevelled complexes (FFD) at the three distal compartments (i.e. $(1,3)$, $(2,3)$, and $(3,3)$) in an arbitrarily chosen cell, cell $(3,4)$.

We take this as justification to confine ourselves in the following to the continuous setting.

Experiment 2: Single cell, time plot. Current biological models of PCP signalling postulate that as a result of the intercellular communication between two neighbouring cells, Fmi, Fz and Dsh accumulate on the distal side of each cell, designating it as the future site for prehair formation, while Fmi, Vang and Pk accumulate on the proximal side of its neighbouring cell. Thus, we first look at the behaviour of a single cell in our in-silico tissue via time series data of one of the products of the intercellular communication which mediates the mechanism of PCP. We expect the FFD complexes to accumulate at the distal edges of each cell rather than the proximal edge.

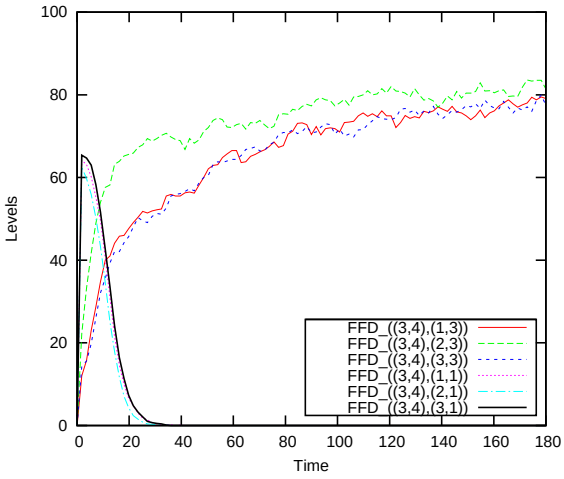

Figure 5: Stochastic simulation result for 100 levels, averaged over 10 runs.

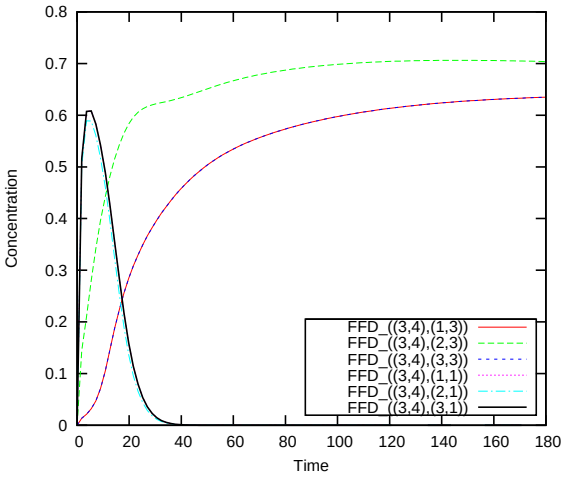

Figure 6: Continuous simulation result. Note that plots for (1,1) and (3,1) are identical, as are (1,3) and (3,3).

The time series plot in Figure 6 displays the dynamic behaviour of FFD in both the proximal $((1,1)$, $(2,1)$ and $(3,1))$ and distal $((1,3)$, $(2,3)$ and $(3,3))$ compartments in cell$(3,4)$. It clearly shows that at the end of signalling FFD accumulates at the distal edge of the cell rather than at the proximal edge, with a very slight peak in the middle distal compartment $(2,3)$. FFD accumulates and reaches a stable state after PCP signalling is activated. Moreover, we find that the concentration of FFD in different distal compartments reaches a similar level if simulating the model over a sufficiently long time.

Experiment 3: Single cell, model checking. We ap-

Table 2: The size of the PCP model and runtime[a)] for unfolding and continuous simulation.

Size				Time (seconds)			
Grid($M \times N$)	Cells	Places	Transitions	Unfolding	Unfolding/Cells	Simulation	Simulation/Cells
5×5	12	924	984	0.665	0.0546	1.192	0.0993
10×10	50	3,850	4,100	2.528	0.0506	5.152	0.1030
15×15	112	8,624	9,184	6.108	0.0545	11.952	0.1067
20×20	200	15,400	16,400	11.705	0.0585	25.272	0.1264
50×50	1,250	96,250	102,500	139.594	0.1117	220.735	0.1766

[a)] done on PC Windows Vista (TM) business Intel(R) Xeon(R), CPN 2.83GHz, Memory(RAM) 4.00 GB;

ply PLTL model checking [8] to analyse traces from the continuous PCP model in order to check whether the property observed in the two former experiments holds in all cells in the considered tissue. We focus on the distal concentrations of species FFD in each cell, and exploit the following two types of properties.

(1) For each cell (x,y) in the honeycomb, we want to check the property: After some time (specified by the constant *init_phase*), FFD in the middle distal logical compartment (i.e. position (2,3)) is always (i.e. during the observed simulation time) greater than all other logical compartments (i.e. positions (1,3) and (3,3)). For this, we form the following query, where A, B and C denote FFD_(x,y)_(1,3), FFD_(x,y)_(2,3) and FFD_(x,y)_(3,3), respectively:

$P_{=?}[G(time > init\_phase \rightarrow ([B] > [A] \& [B] > [C]))]$

We simulate the continuous model for the 15*15 honeycomb grid consisting of 112 cells in total over 180 time units, and use this query with *init_phase* set to 0 to check all cells. The query holds for all these cells except the cells in the last column, i.e. cells (2,15), (4,15) ... (14,15).

(2) For FFD in each distal compartment of each cell, we also want to check how many peaks exist in their traces. For this, we define the following queries (take FFD_(x,y)_(2,3) as an example):

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

Here we use the function $d(species)$ to get the derivative of the concentration of the species at each time point. Using the same traces as above we get the following results.

- For the cells in Row 1, Row 15 and Column 15, FFD_(x,y)_(2,3) has no peak.
- Except for these boundary cells, other cells have exactly one (slight) peak for FFD_(x,y)_(2,3).
- For FFD in other distal compartments, there are no peaks.

Experiment 4: Wild-type tissue. This experiment aims to show how the signalling flows through the whole in-silico tissue via a network of communicating cells to produce the pattern of the whole tissue of wild-type cells.

We selected the time point when the signal between the distal and proximal edges in a cell is the greatest, which indicates that the signal is strongly effective. In our model, we choose time step 60 as such a time point to compute the overall concentration of FFD at the distal edge of each cell. We sum the concentration of FFD in the three distal compartment to give the overall concentration.

In order to analyse the orientation of hairs, we display the distribution of FFD within each cell in the in-silico tissue in the following way. We assume that

- if the FFD level at the distal edge of a cell is higher than that at the proximal edge, the hair points distally (denoted by red right-arrows);
- if the level of FFD at the distal edge of a cell is lower than that at the proximal edge, the hair points proximally (denoted by green left-arrows);
- if the level of FFD is equal at the distal and proximal side, the hair's orientation is decided randomly (denoted by black circle).

For the wild-type tissue, we get a figure similar to Figure 7, left, but trivially with all hairs pointing distally (not given in this paper).

Experiment 5: Mutated tissue. We have modelled the effect of a patch of cells lacking the key signalling molecule Frizzled in an otherwise wild-type field of cells by completely knocking out the concentration and transport of the corresponding places in our model. For demonstration purpose, we present the results of Fz mutant clone as an example [34]. Using the $MutReg$ function (see Table 1), we produce a mutant clone of Fz- inside our in-silico tissue, comprising seven cells ($(4,5)$,$(5,4)$, $(5,6)$, $(6,5)$, $(7,4)$, $(7,6)$, $(8,5)$). 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 [18]. Regarding the capability of our current model, we expect: (1) cells in the clone have incorrect polarity (FFD does not form); (2) wild-type cells distal to the clone have FFD accumulated at the proximal rather than distal edge of cell. We apply the same approach as described in the analysis of the wild-type tissue.

The result shows the impact of a clone of Fz- mutant cells on their distal neighbouring cells (see Figure 7, bottom right). In addition, we find that the level of FFD at the distal edge of wild-type cells adjacent to the Fz- clone has been reduced compared to wild-type tissue as shown from our simulation. Thus, our model has made a predication that can be tested in a biological experiment.

Discussion. We have shown that our model recapitulates the signalling phenomena known to occur in wild-type and in and around Fz-clones as a first attempt to demonstrate the basic principles of implementing of CPN to PCP signalling and show the usefulness of CPN to multi-scale modelling of a multi-cellular system. We will reproduce the phenotype of other mutations in our future research. We intend to continue elaborating a more sophisticated model which will be able to answer the questions addressed at the end of Section 2.

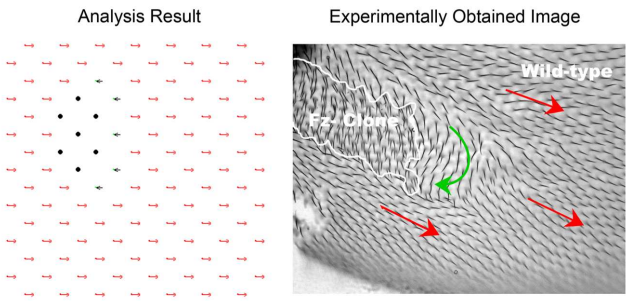

Figure 7: Analysis result from our simulation vs Fz- clone in wild-type background: (a) Analysis result: red right-arrows represents that the overall FFD at the distal edge is higher than the proximal edge, green left-arrows represents the overall FFD at the distal edge is lower than the proximal edge, black circles represent the overall FFD is equal at distal and proximal edges (Fz- clone causes non-production of FFD in the mutated cells in our model); (b) The biological image: the red arrows show the orientation of most wild-type cells surround the Fz- clone and the green arrow follows the orientation of hairs in cells from the distal neighbourhood of the clone.

Another crucial point is how many cells we can simulate in terms of current computational capabilities, i.e. what tissue size can we actually analyse. For this, we have to address three technical key problems: unfolding, ODEs construction (in the continuous case), and simulation. The coloured (stochastic or continuous) Petri nets are automatically unfolded which can be considered as a kind of compilation. In the continuous case, for each place the corresponding ordinary differential equation needs to be constructed. Finally, the established model has to be simulated using an appropriate continuous or stochastic simulation algorithm.

Using the refined coloured Petri net model, we performed a simple test by increasing the number of cells in the tissue. We report the unfolding/simulation time for different size of the PCP model in Table 2. From the ratio of unfolding/simulation time to number of cells, we can see that both the unfolding and simulation time increase approximately linearly.

7. REPRODUCIBILITY

All Petri net models in this paper were constructed with Snoopy [27], recently extended to support coloured Petri nets [22], which can be obtained from `http://www-dssz.informatik.tu-cottbus.de/DSSZ/Software/Snoopy`.

Simulations were done with Snoopy's built-in stochastic and continuous simulators, simulation traces have been written to csv files, which have then been further processed by MC2 [8] for model checking and Matlab®7.11.0 [24], specifically to derive Figure 7, left.

The models in Snoopy format and high resolution diagrams, and the simulation data can be found at `http://people.brunel.ac.uk/~cspgqqg/`.

Thus, all our results can be easily reproduced by the interested reader.

8. CONCLUSION

In this paper, we have presented our current work applying Petri net techniques to construct a generic computational model in order to explore the mechanisms that drive Planar Cell Polarity in Drosophila wing tissue. Our approach has involved developing a sophisticated pattern of communicating sub-models. We have demonstrated the usefulness of coloured Petri nets in this scenario and identified the need for hierarchically organised colours. Our prime motivation in this paper has been to develop the modelling methodology based on the new concept of hierarchically coloured Petri nets, and to demonstrate its usefulness.

Our model has allowed us to generate behaviours as a first step to explaining the complex behaviours observed in the biological system under study, and to explore the effects of mutations by introducing variations in patches of cells in our computational model. Our analysis confirms that the behaviour of the model correctly demonstrates that the major accumulation of actin (from which the hairs are formed) occurs in the most distal part of wild-type cells, corresponding to the location of the prehair formation in wing cells of Drosophila. Moreover our model confirms that the introduction of mutant clones disrupts the pattern of actin accumulation and hence hair orientation in wild-type cells on the distal side of the clone.

The construction of sufficiently large models of Drosophila wing epithelial tissue has required computational experiments over several thousands of ordinary differential equations. Our HCPN model and the software required to simulate and analyse it are freely available, thus ensuring that our results are reproducible by the scientific community.

Our current model serves as a proof of concept. However, its the ability to make predictions and provide an accurate picture of PCP signalling is limited by its lack of sufficient biological detail. In ongoing work we are using CPN to construct more realistic PCP models which incorporate more detailed descriptions of the signalling pathways involved. Our overall aim is to facilitate a better understanding of the mechanisms that drive PCP, and to make predictions about the behaviour of the system when it is perturbed by the loss of specific signalling components. Furthermore we are also challenging our modelling framework by other case studies including motile elements.

9. ACKNOWLEDGMENTS

Qian Gao is supported by a bursary from Brunel University, and Fei Liu is supported by the Modeling of Pain Switch (MOPS) program of Federal Ministry of Education and Research (Funding Number: 0315449H).

10. REFERENCES

[1] K. Amonlirdviman, N. Khare, D. Tree, W. Chen, J. Axelrod, and C. Tomlin. Mathematical modeling of planar cell polarity to understand domineering nonautonomy. *Science*, 307:423–426, 2005.

[2] J. Axelrod. Unipolar membrane association of dishevelled mediates frizzled planar cell polarity signaling. *Genes Dev*, 15:1182–1187, 2001.

[3] J. Axelrod and C. Tomlin. Modeling the control of planar cell polarity. *WIREs Syst Biol Med*, 2011.

[4] N. Bahi-Jaber and D. Pontier. Modeling transmission of directly transmitted infectious diseases using

colored stochastic Petri nets. *Mathematical Biosciences*, 185:1–13, 2003.

[5] P. Baldan, N. Cocco, A. Marin, and M. Simeoni. Petri nets for modelling metabolic pathways: a survey. *Natural Computing*, 9(4):955–989, 2010.

[6] Y. Burak and B. I. Shraiman. Order and stochastic dynamic in drosophila planar cell polarity. *PLoS Comput. Biol.*, 5(12):e1000628, 2009.

[7] S. Christensen, J. B. Jørgensen, and L. M. Kristensen. Design/CPN - A Computer Tool for Coloured Petri Nets, 1997.

[8] R. Donaldson and D. Gilbert. A model checking approach to the parameter estimation of biochemical pathways. *LNCS/LNBI*, 5307:269–287, 2008.

[9] H. Genrich, R. Küffner, and K. Voss. Executable Petri net models for the analysis of metabolic pathways. *International Journal on Software Tools for Technology Transfer*, 3(4):394–404, 2001.

[10] D. Gilbert, M. Heiner, and S. Lehrack. A unifying framework for modelling and analysing biochemical pathways using Petri nets. *LNCS/LNBI*, 4695:200–216, 2007.

[11] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[12] A. P. Heath and L. E. Kavraki. Computational challenges in systems biology. *Computer Science Review*, 3(1):1–17, 2009.

[13] M. Heiner, D. Gilbert, and R. Donaldson. Petri nets in systems and synthetic biology, 2008.

[14] M. Heiner, I. Koch, and K. Voss. Analysis and simulation of steady states in metabolic pathways with Petri nets, 2001.

[15] K. Jensen. Coloured Petri Nets and the Invariant-Method. *Theor. Comput*, 14:317–336, 1981.

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

[17] K. Jensen, L. M. Kristensen, and L. M. Wells. Coloured Petri Nets and CPN Tools for Modelling and Validation of Concurrent Systems. *International Journal on Software Tools for Technology Transfer*, 9(3/4):213–254, 2007.

[18] P. A. Lawrence, G. Struhl, and J. Casal. Planar cell polarity: A bridge too far? *Curr Biol.*, 18(20):959–961, 2010.

[19] J. 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. *Dev. Genes Evol.*, 218:413–426, 2008.

[20] J. Le Garrec, P. Lopez, and M. Kerszberg. Establishment and maintenance of planar epithelial cell polarity by asymmetric cadherin bridges: A computer model. *Dev. Dyn.*, 235:235–246, 2006.

[21] D. Lee, R. Zimmer, S. Lee, and S. Park. Colored Petri net modeling and simulation of signal transduction pathways. *Metabolic Engineering*, 8:112–122, 2006.

[22] F. Liu and M. Heiner. Colored Petri nets to model and simulate biological systems, June 2010.

[23] F. Liu and M. Heiner. The Manual for QPN^c/SPN^c/CPN^c. Technical report, Brandenburg University of Technology, 2010.

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

[25] M. Peleg, I. S. Gabashvili, and R. B. Altman. Qualitative models of molecular function: linking genetic polymorphisms of trna to their functional sequelae. *Proc. of the IEEE*, 90(12):1875–1886, 2002.

[26] R. Raffard, K. Amonlirdviman, J. Axelrod, and C. Tomlin. An adjoint-based parameter identification algorithm applied to planar cell polarity signaling. *IEEE Trans. Automat. Contr.*, 53:109–121, 2008.

[27] C. Rohr, W. Marwan, and M. Heiner. Snoopy - a unifying Petri net framework to investigate biomolecular networks. *Bioinformatics*, 26(7):974–975, 2010.

[28] T. Runge. Application of coloured Petri nets in systems biology, 2004.

[29] N. M. S. Schamberg, P. Houston and M. R. Owen. Modelling and analysis of planar cell polarity. *Bulletin of Mathematical Biology*, 72:645–680, 2010.

[30] M. Simons and M. Mlodzik. Planar cell polarity signaling: From fly development to human disease. *Annu. Rev. Genet*, 42:517–540, 2008.

[31] D. Strutt. Asymmetric localization of frizzled and the establishment of cell polarity in the drosophila wing. *Mol. Cell*, 7:367–375, 2002.

[32] D. R. P. 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*, 109:371–381, 2002.

[33] C. Täubner, B. Mathiak, A. Kupfer, N. Fleischer, and S. Eckstein. Modelling and simulation of the TLR4 pathway with coloured Petri nets, 2006.

[34] C. R. Vinson and P. N. Adler. Directional non-cell autonomy and the transmission of polarity information by the frizzled gene of drosophila. *Nature*, 14(329):549–551, 1987.

[35] K. Voss, M. Heiner, and I. Koch. Steady state analysis of metabolic pathways using Petri nets. *In Silico Biology*, 3:0031, 2003.

[36] D. J. Wilkinson. *Stochastic Modelling for Systems Biology*. Chapman & Hall/CRC, 2006.

[37] L. Wong and P. Adler. Tissue polarity genes of drosophila regulate the subcellular location for prehair initiation in pupal wing cells. *J. Cell Biol.*, 123:209–220, 1993.