Agent Based Modelling of Auxin Transport Canalisation

Philip Richard Garnett

A Thesis submitted for the degree of Doctor of Philosophy

University of York

Department of Biology

September 2010

Abstract

The transport of the plant hormone auxin is central to many aspects of plant development. This thesis describes work developing agent-based simulations of auxin transport canalisation. Auxin transport canalisation is the dominant theory describing how auxin promotes its own transport through a tissue from a site of auxin biosynthesis to an auxin sink. Directional transport of auxin through plant tissues is mediated by the PIN family of membrane transporters. The exact mechanisms of PIN localisation that facilitate this directional transport are not known. Part of the aim of this work is to test theories of PIN localisation.

The development of the agent-based simulations has been assisted by the CoSMoS process. Using the CoSMoS modelling process, a series of models are produced that assist in the development of simulations of PIN localisation hypotheses. The CoSMoS process describes a systematic method for the development of simulations of complex systems. The process helps to ensure that the assumptions made during the modelling process are understood, and known.

Two generations of simulator have been developed. The first has successfully tested two hypotheses of PIN localisation in rectangular cells. One, the Diligent Worker Hypothesis is able to produce auxin transport canals in tissues. These canals are an emergent property of the hypothesis and form in a variety of situations. The CoSMoS process was then used to enhance the existing model to test the Diligent Worker Hypothesis in irregular cells of a more natural shape, and to lay the foundation for a well constructed 3D model. The enhanced simulator is flexible and could be used in the future for the testing of additional localisation hypotheses, or auxin homoeostasis in cells.

Contents

1 Introduction **1**

2 Auxin Transport **3**
2.1 Auxin . . . 3
2.2 Auxin's Roles in Plant Development and Patterning . . . 5
2.2.1 Auxin Transport Canalisation . . . 6
2.2.2 Shoot Branching . . . 6
2.3 Modelling Auxin Transport . . . 7
2.3.1 Canalisation Models and Venation Patterns . . . 8
2.3.2 Phyllotaxis Models . . . 10
2.3.3 Static PIN Models . . . 13
2.3.4 PIN polarisation mechanisms . . . 13
2.3.5 Whole Plant Models . . . 16
2.4 Modelling Methods . . . 17
2.4.1 Biology and UML . . . 17
2.4.2 Other Model Engineering Systems . . . 19
2.5 Executable Models . . . 20
2.6 Emergence . . . 21
2.7 Conclusions . . . 21
2.8 The Context of this Work . . . 22

3 First Gen. Development **25**
3.1 Overview of the modelling process . . . 26
3.2 Domain Model: auxin transport . . . 28
3.2.1 The biological background . . . 28
3.2.2 Modelling . . . 28
3.2.3 Domain Model Use Cases . . . 29
3.2.4 Domain Model class diagram . . . 29
3.2.5 Domain Model state diagrams . . . 31
3.3 Software Model . . . 36
3.3.1 Software Model Use Cases . . . 36
3.3.2 Software Model Class diagram . . . 37
3.3.3 Software Model State diagrams . . . 38
3.4 Implementation Model . . . 39
3.4.1 Implementation Model class diagram . . . 39
3.4.2 Implementation Model state diagrams . . . 40

3.5 Auxin canalisation Simulator 40
3.5.1 Molecule diffusion 40
3.5.2 User interface 43
3.5.3 Visual Output of Simulations 44
3.5.4 2D and 3D simulations 45
3.6 Diligent Worker Hypothesis 46
3.7 Internal Auxin hypothesis: Measuring Internal Auxin Gradients 48
3.8 General Simulation Parameters 49
3.9 Conclusion 52

4 First Gen. Results 53
4.1 Parameter Search Methods 53
4.1.1 Genetic Algorithm 54
4.1.2 Diligent Worker Hypothesis Parametrisation 54
4.1.3 Successful Parameters 56
4.2 Internal Auxin hypothesis: Measuring Internal Auxin Gradients 84
4.3 Experimental Results 85
4.3.1 Canal Emergence from Background Auxin 92
4.4 Conclusion 93

5 Second Gen. Development 95
5.1 CoSMoS Process: The modelling lifecycle 95
5.1.1 Modelling biology and simulations with UML 97
5.2 The Research Context 97
5.3 The Domain Model: auxin transport canalisation 98
5.3.1 Domain Model UML 98
5.4 Modelling Space 99
5.5 Platform Model 100
5.5.1 Platform Model UML 100
5.5.2 The Division of Space 101
5.5.3 Agents in Space 105
5.5.4 Space from Templates 105
5.6 Conclusion 107

6 Second Gen. Results 109
6.1 Improved Simulation architecture allows performance improvements 109
6.2 Comparison with first generation model 110
6.3 Irregular Cells 113
6.4 Vacuole Size 118
6.5 Conclusion 124

7 3D Model 125
7.0.1 3D Space 126
7.1 Conclusion 128

8 Discussion **129**
8.1 Model Development . . . 129
8.1.1 Development of First Model . . . 130
8.1.2 Enhancement of First Model . . . 131
8.2 Simulator Results . . . 132
8.2.1 PINs: The Diligent Workers of Auxin Transport Canalisation . . . 132
8.2.2 Comparison with Other Canalisation Models . . . 135
8.2.3 Vacuole Size . . . 138
8.2.4 Cell Shape . . . 138
8.2.5 Localisation by Cytoplasmic Auxin Gradients . . . 139
8.3 Simulation Run Time . . . 140
8.4 Simulator Interfaces . . . 141
8.5 Support Tools . . . 141
8.6 Other Areas of Future Work . . . 142
8.6.1 Experiments with different arrangements of Sources and Sinks . . . 142
8.6.2 Experiments with Single Cells . . . 143
8.6.3 Simulator Extensions . . . 144
8.7 Concluding Remarks . . . 144

A Supplementary Data **147**
A.1 Starting Parameters Used for GA . . . 147
A.2 Starting Parameters Used for Internal Auxin GA . . . 149

B Accessory Programs **151**
B.1 Check pointing . . . 151
B.2 DataMining . . . 151
B.3 ChromoServe . . . 152
B.3.1 Analysis of Rough Cluster Data . . . 152
B.3.2 ChromoServe2 . . . 153
B.4 SourceSink . . . 153
B.5 Template Tools . . . 153
B.6 Time Series . . . 153

List of Figures

2.1 The biology of auxin transport through cells. 4
2.2 Membrane localisation of PIN. 15

3.1 The CoSMoS LifeCycle 27
3.2 The Domain Model Use Case diagram. This covers the high level biological aspects that need to be included in the simulator, with the Plant itself as the actor. 29
3.3 The Domain Model class diagram. The biological components of the simulation are modelled as interacting objects defined by classes. 30
3.4 State diagram for the Cell. 31
3.5 The state diagram for Auxin. 32
3.6 State diagram for the PIN Efflux Protein. 34
3.7 State diagram of an alternative hypothesis for the PIN Efflux Protein. 35
3.8 The Software Model Use Case diagram, with the Simulation user actor. 37
3.9 The Software Model class diagram, showing only the biological part of the model. 38
3.10 A simplified implementation class diagram. 39
3.11 Production and movement of auxin agents in the simulations. . 42
3.12 Production and movement of AUX/LAX agents in the simulations. 43
3.13 Visual output from the first generation simulator. 44
3.14 The diligent worker hypothesis. 47

4.1 The effect of changing the probability of PIN dissociating from the membrane. 59
4.2 PIN dissociation probability of 10^{-3}. 60
4.3 PIN dissociation probability of 10^{-7}. 60
4.4 PIN dissociation probability of 10^{-11}. 61

4.5 Interaction of PIN with the auxin of the Apoplast 63
4.6 Interaction of PIN with the auxin of the Apoplast 64
4.7 Plot showing the probability of PIN dissociating from the membrane due to apoplastic auxin concentration. 65
4.8 Simulation run with only half of the Diligent Worker Hypothesis present 66
4.9 Simulation run with only half of the Diligent Worker Hypothesis present 67
4.10 PPC: A parameter in the Hill equation that controls the production of PIN in response to increasing auxin concentration . . 69
4.11 PHC: A parameter in the Hill equation that controls the production of PIN in response to increasing auxin concentration . . 70
4.12 μ: The parameter that determines the highest possible concentration of PIN in the cells. 71
4.13 Plot showing the changing concentration of PIN protein as the concentration of auxin increases. 72
4.14 Changing the probability of PIN degrading in the cytoplasm of the cell. 74
4.15 With a probability of PIN degradation of 10^{-2} canals are still able to form but qualitatively not as good as those formed with a probability of 10^{-4} (figure 4.17). 75
4.16 With a probability of PIN degradation of 10^{-6} canals form between the auxin source and the sink. 75
4.17 With a probability of PIN degradation of 10^{-4} narrow canals form between the auxin source and the sink. 76
4.18 Time series data showing the concentration of PIN in a cell that is part of a canal. 78
4.19 Time series data showing the concentration of PIN in a cell that is on the edge of a simulated tissue and not part of a canal. . . . 79
4.20 Plot showing the effect of increasing AUX/LAX concentration. . 81
4.21 The effect of an AUX/LAX concentration of 0. 82
4.22 The effect of an AUX/LAX concentration of 0.04. 83
4.23 Internal Auxin Hypothesis 86
4.24 Internal Auxin hypothesis 87
4.25 Still showing a section of a simulated tissue where one cell has become a high concentration auxin sink. 87
4.26 A diagonal test of the diligent worker hypothesis. 88

4.27 A test of the diligent worker hypothesis with a single blockage in the tissue. 89
4.28 A test of the diligent worker hypothesis with a two blockages in the tissue. 90
4.29 Figures from Sachs'[103] paper showing the formation of vascular tissue in response to applied auxin. The canals which form are able to go around corners introduced into the tissue. 91
4.30 Emergent Canals produced by the diligent worker hypothesis in a tissue. 92

5.1 The components of the updated CoSMoS process. 96
5.2 Platform Model class diagram. 102
5.3 Output from the first generation simulator. 102
5.4 Implementation level Platform Model class diagram of the initial simulator. 103
5.5 Implementation level Platform Model class diagram of the enhanced simulator. 104
5.6 Processing sections of plants into the model. 106

6.1 Comparison of simulator performance between generations. . . . 110
6.2 Basic test with regular cells. The canal forms between the source 'S' and the Sink 'Sn'. 111
6.3 Comparison of the fitness of a canals formed through a 5x5 grid of cells. 112
6.4 Basic diagonal test with regular cells. The canal forms between the source 'S' and the Sink 'Sn'. 112
6.5 Comparison of the first and second generation models. 114
6.6 Plot showing the average number of Apoplastic neighbours a Cell Membrane area has. 115
6.7 Straight canal through irregular cells 116
6.8 Canal going around a block through irregular cells 117
6.9 Effect of vacuole size in smaller cells. 118
6.10 Effect of vacuole size on PIN localisation in larger cells. 119
6.11 Early PIN distribution on the membrane. 121
6.12 The patchy period of PIN distribution on the membrane. 122
6.13 Polar PIN distribution on the membrane. 123

7.1 Comparison of possible paths of auxin molecules (or other agents) in 2D or 3D cells. 127

List of Tables

3.1 Important assumptions concerning the behaviour of auxin in the simulations. 33
3.2 Important assumptions concerning the behaviour of PIN in the simulations. 35
3.3 General Model Parameters. These parameters remain the same for all hypotheses tested. 50
3.4 General Agent Parameters. These parameters remain the same for all hypotheses tested. 51

4.1 Diligent Worker Parameters. These parameters are specific to the Diligent Worker Hypothesis. 57
4.2 Most successful Internal Auxin Hypothesis parameters. 84

A.1 Diligent Worker Parameters used at the start of the search carried out with the Genetic Algorithm 148
A.2 Starting parameters used in for Internal Auxin hypothesis parameter search. 149

Acknowledgements

I would like to thank my supervisors Ottoline Leyser and Susan Stepney, and the members of my thesis advisory panel, Leo Caves (The Dude) and Harv Isaacs, for their support, advice and guidance. Przemyslaw Prusinkiewicz for discussion and ideas. I would also like to thank Lauren Shipley for her support and four years of expert proof reading. Finally I would like to thank Alastair, Gemma, Lynne, Scott, Jenny, Nick, Jo, Adam, Vera, Ed and all of YCCSA and the Leyser Lab for being good friends and colleagues.

Declaration

The work presented in Chapter 2 is based on work published in, "Computer Simulation: the imaginary friend of auxin transport biology", BioEssays 32(9), pages 828–835, September 2010, Philip Garnett, Arno Steinacher, Susan Stepney, Richard Clayton, Ottoline Leyser. The work presented in Chapter 3 is based on work published in, "Towards an executable model of auxin transport canalisation", CoSMoS workshop, York, UK, September 2008, pp63-91. Luniver Press 2008, Philip Garnett, Susan Stepney, Ottoline Leyser. The work presented in Chapter 2 is based on work published in, "Using the CoSMoS Process to Enhance an Executable Model of Auxin Transport Canalisation", CoSMoS workshop, Odense, Denmark, August 2010. pp9-32, Luniver Press 2010, Philip Garnett, Susan Stepney, Francesca Day, Ottoline Leyser.

The program DataMining, described in Appendix B.2, and some associated scripts for 3D template generation, described in Appendix B.5, is developed from initial work done by Francesca Day during a summer project at the University of York.

Chapter 1

Introduction

The simulation of biological systems presents a significant challenge. Knowledge from all branches of science is often required in order to capture all aspects of the biological, chemical and physical processes occurring. The challenge is made more difficult by the complex nature of biology; it is often very difficult to make good assumptions about how a particular process is regulated. The quality of the data available is also very important, as the simplest solution based on the information at hand may not be representative of what is happening in the system being studied, but more a reflection of how the current data are insufficient. Or the data may be excellent, but missing a part of the picture altogether. Another significant difficultly is that the connectivity of processes in biology is often very high, therefore the question of the level of abstraction and simulation complexity is important. If the abstraction level is too high we risk ruling out simulations producing emergent behaviour; too low, and the simulations produced could be difficult to work with and validate. These factors make the developmental decisions made when producing a simulation important, as a balance must be sought between having enough of the system represented to allow for unexplected interactions between processes to produce emergent behavours, without developing a simulation that is potentially as difficult to understand as the system being modelled.

This thesis presents work done to develop abstract models of cells using the CoSMoS[1] process assisted by UML (Unified Modelling Language)[81], developed with the UML-based software development plug-in for the Netbeans integrated development environment[84]. The resulting executable agent-based simulator is implemented in the object-oriented (OO) programming language Java[83]. The CoSMoS process with UML was used to enhance the cross discipline communicability of the models; and the combination of using UML

and the OO programming paradigm for producing biological models produces simulations that map intuitively back to the biology.

The modelling framework developed allows for the testing of biological hypotheses to investigate the process of PIN polarisation that occurs during auxin transport canalisation. Chapter 2 provides an overview of the biology of auxin transport, section 2.3 introduces existing models of auxin transport, and section 2.4.2 introduces other frameworks used to assist the development of biological simulations. The simulator went through two main phases of development, producing two generations of simulation software. Chapter 3 discusses the development using the CoSMoS process of the first generation simulator. The results from this simulator are discussed in Chapter 4 and Chapter 4.3. Chapter 5 discusses how the first generation simulator was enhanced using the CoSMoS process to produce a more capable second generation simulator. The results from the second generation simulator are discussed in Chapter 6. Chapter 7 discusses the 3D capability of the enhanced simulator that could be exploited in the future. Finally Chapter 8 evaluates the simulator in the context of the existing work in the field and also discusses possible future work that could be carried out.

Chapter 2

Auxin Transport

The transport of the plant hormone auxin is central to numerous different patterning and developmental processes that occur in plants. This work is focused primarily on the process known as auxin transport canalisation, specifically looking at the localisation of the auxin transport protein PIN in the membranes of cells. The mechanism of PIN localisation is likely to be central to other auxin transport mediated processes such as phyllotactic patterning, embryogenesis, and shoot branching regulation. In section 2.1 the hormone auxin and its transporters are introduced. Section 2.2 discusses some of auxin transport's many roles in plant development. With subsection 2.2.1 discussing the biology of auxin transport canalisation in more detail, and 2.2.2 discussing shoot branching. An overview of current modelling of auxin transport is given in section 2.3, and section 2.4 discusses methods used to develop computational models and simulations.

2.1 Auxin

The chemistry of auxin plays an important part in how it regulates patterning and development in plants. The most abundant natural auxin is indole-3-acetic acid (IAA). In plant issues IAA is in equilibrium between its protonated and deprotonated (ionised) forms. At the low pH (approximately pH5.5) of the extracellular space (the apoplast), a significant fraction of the auxin is protonated, and thus able to cross the plasma membrane and enter cells passively. However, the cytoplasmic pH is much higher (approximately pH7.0), so once inside the cell auxin is almost entirely ionised and thus unable to leave without active protein-mediated transport (figure 2.1). This phenomenon is known as the acid trap or the chemiosmotic theory and is summarised by Rubery &

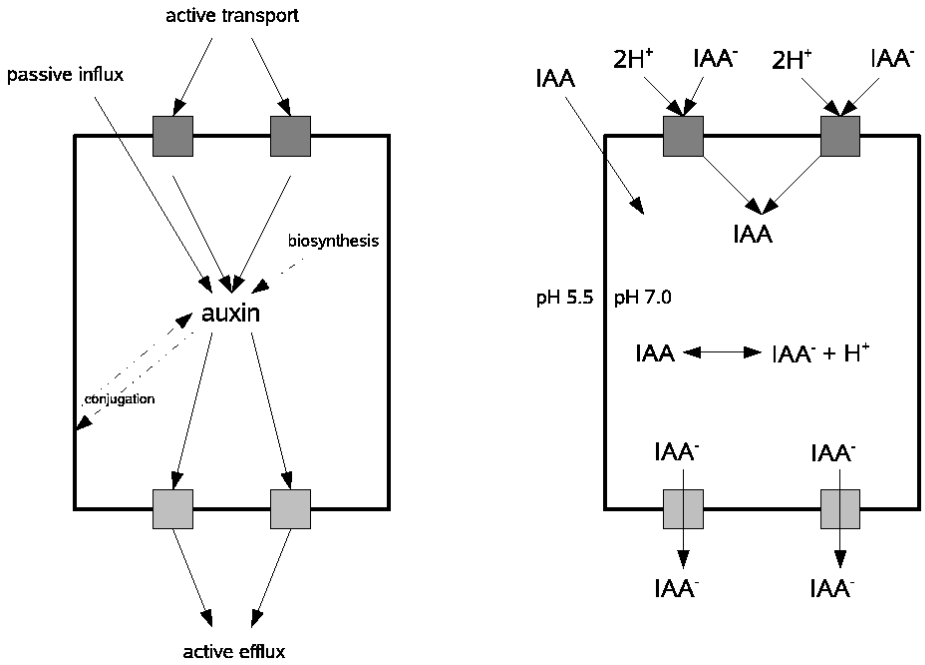


Figure 2.1: The biology of auxin transport through cells. Auxin (IAA) can enter cells passively, or it can be actively transported by AUX/LAX proteins. Once in the cell its only method of escape is active transport by proteins like PIN[76].

Sheldrake[99], and Raven[93].

Synthesised in the cytoplasm of cells the auxin is trapped inside, reliant on membrane localised efflux proteins to get across the plasma membrane out into the apoplast (figure 2.1). There are several protein families that have been shown to act in auxin efflux. Prominent among these are members of the ABCB[34, 4] and PIN families of membrane localised transporters[29, 88, 120, 115]. A great deal of attention has focused on the PIN protein family. PIN proteins are are often found distributed asymmetrically around the membrane of a cell, and it is this asymmetric localisation of PIN transporters that provides directionality to auxin transport. Experimental evidence supports the hypothesis that the direction of auxin transport across tissues is determined by the co-ordinated positioning of PIN proteins[18]. The mechanisms of PIN localisation on membranes are a primary focus of this work and are discussed in more detail in section 2.3.4. Auxin can also be actively transported into cells via members of the AUX/LAX influx protein family[86].

2.2 Auxin's Roles in Plant Development and Patterning

During the early stages of embryogenesis the apical-basal axis is defined by dynamic patterns of auxin accumulation. Auxin accumulates first at the shoot pole and then later the pattern reverses and auxin accumulates at the root pole. Auxin accumulation at these sites is required to establish the shoot apical meristem (SAM) at the shoot pole, and the root apical meristem at the root pole[6]. During post-embryonic development these small groups of stem cells are responsible for the production of the shoot and root systems respectively. The reversal of auxin distribution in the apical/basal axis during embryogenesis is correlated with the redistribution of PIN proteins from the apical face to the basal face of embryo cells[29, 30]. The importance of these events to embryonic patterning is demonstrated by the fact that the establishment of the embryonic axis is disrupted in pin mutants, and in wild-type plants to which pharmacological inhibitors of auxin transport have been applied.

After germination the SAM goes on to initiate the successive organs of the shoot. During this process, known as phyllotactic patterning, auxin directs the positioning of primordia for lateral organs, such as leaves, in a regular pattern around the shoot. This process has long fascinated plant scientists and mathematicians because of the geometry of the patterns produced. One common pattern is spiral phyllotaxis, in which organs are produced in a spiral around the shoot axis separated by approximately 137.5 degrees, the golden angle derived from the ratio of consecutive numbers in the Fibonacci sequence. During phyllotactic patterning, PIN protein localisation in cells of the epidermis of the SAM apparently directs local sites of auxin accumulation and these sites correspond to the position at which the organ primordia develop[55, 97, 96]. This process requires the continuous repositioning of PIN proteins to create the sequence of sites of auxin accumulation needed to trigger the initiation of successive organs. Models of Phyllotactic Patterning are in many ways closely related to those of canalisation and are discussed in section 2.3.2.

Auxin is also very important for the patterning of the root tip. Developing leaves are a strong source of auxin in a growing plant[65, 122, 11]. Auxin is transported through the expanding leaves and exported out into the main stem. Once in the main stem it is transported downwards in files of cells associated with the vascular bundles. These cells express high levels of basally localised PIN proteins[39, 50]. The root system of a plant acts as a sink for

the auxin. In the root the auxin is transported in the central vascular cylinder towards the root tip. At the tip, a complex pattern of PIN protein localisation results in the auxin being transported upwards from the tip in the outer cell layers of the root, where is recycled back into the central vascular cylinder and transported back down towards the root tip[9, 39]. This is often referred to as a reverse fountain.

2.2.1 Auxin Transport Canalisation

Auxin plays a key role in the development of the vascular network which extends throughout the plant. The vascular bundles of the main stem and it's branches link the intricate patterns of veins in leaves to the central vascular core of the root[106]. Canalisation theory is the most prominent theory for the mode of action of auxin in vascular patterning. Canalisation describes a positive feedback between auxin, auxin transport, and its polarisation. It is thought that auxin produced at a source is transported through the tissue by PIN proteins. Auxin promotes its own transport through the tissue as the amount of PIN available to transport auxin increases with auxin concentration[85]. The exact details of how PINs become localised in unknown, but they become polarised within the cell to transport the auxin towards a sink elsewhere in the tissue. Thus the more auxin in a cell, the more it can transport, and that transport is directed towards the auxin sink. The initially wide auxin transport canal narrows as it develops which results in the establishment of canals of auxin transporting cell files connecting an auxin source to an auxin sink. These canals later differentiate into vascular tissue. The biology of how a number of these events occur in the cells is unknown and therefore modelling can provide a platform for the testing of hypotheses.

2.2.2 Shoot Branching

The basipetal flux of auxin down the plant from the shoot to the root is called the polar transport stream and provides a systemic source of information about the health and activity of shoot apices[62]. One role for this transport stream is in the regulation of shoot branching. Shoot branches arise from secondary SAMs established in the axial of each leaf formed by the primary shoot meristem. These axillary meristems may remain dormant, or they may activate to produce a shoot branch. Auxin, produced by active shoot apices with young expanding leaves, is exported into the stem and has an inhibitory influence

on the activity of the axillary meristems below, a phenomenon known as apical dominance[113, 13]. Removal of the active shoots results in release of the inhibited axillary meristems, which can be prevented by application of auxin to the decapitated stump. The auxin in the polar transport stream inhibits bud growth indirectly, since it does not enter the bud[89, 10]. Several non-exclusive mechanisms have been proposed to account for this. These include: down-regulation of the synthesis of the plant hormone cytokinin, which is a promoter of axillary bud activation[112]; up-regulation of the synthesis of a third hormone likely to be a strigolactone or related compound that inhibits bud activation[114, 37]; and prevention of auxin transport canalisation out of the axillary bud by reducing main stem sink strength[63, 7, 90]. The possible link with auxin transport canalisation is of most interest to this work and a model of auxin transport mediated branching regulation is discussed in section 2.3.5.

The examples above illustrate that the auxin transport network is critical for regulating plant development. Particularly striking are the complex positional dynamics of the PIN protein family of auxin transporters. This highlights the question of how PIN protein localisation is regulated[21]. Relatively little is known about this regulation, but it is clear that auxin itself plays a central role. Auxin is known to regulate PINs at multiple stages, affecting the transcription of PIN genes[116], the cycling of PIN proteins between the plasma membrane and endomembrane compartments and the stability of PINs[17, 35, 47, 85]. Since the amount of auxin both inside and outside cells is affected by PINs, auxins and their transporters are intimately linked in complex feedback loops. The complexities of these feedback loops and the demanding nature of the cell biology and biochemistry required to probe the mechanisms of PIN localisation have created a fertile ground for computational modelling to test the plausibility of hypotheses and prioritise wet experiments.

2.3 Modelling Auxin Transport

Due to complexity of the auxin regulated patterning and development processes such as phyllotaxis and canalisation, modelling has become popular as a way to assist in progressing the field's understanding. There is therefore a significant body of existing modelling work. However, the development of models of auxin transport is not without significant challenges. The measurement of the change in concentration of auxin and PIN over time in order to get

accurate parameters for the models is very challenging. Efforts have been made to investigate careful parameter choice and validate models[51]. However, validation of hypotheses generated by models and often the models themselves has been mostly at a more macroscopic level, involving successful reproduction of observable plant systems behaviour *in silico* (e.g. by reproducing laser ablation of tissue parts in the model[39, 97]).

A particular trend has been to build models based on alternative unsupported assumptions such as postulating a hypothetical sensory mechanism[50, 5, 45]. If such models are able to reproduce complex plant behaviours then this indicates that the assumptions are plausible and can guide future wet experiments aimed at testing their validity. The work presented in this thesis was to develop a framework for the the testing of PIN localisation hypotheses in the context of auxin transport canalisation. The success of the simulations will be validated against biology, and other models, and therefore the examples below are presented to allow comparison. Section 2.3.1 describes existing models of auxin transport canalisation. Section 2.3.2 discusses models of phyllotaxis. Phyllotaxis models normally differ in the mechanism used to polarise PIN, however attempts have been made to unify phyllotaxis and canalisation, and these are discussed in section 2.3.2.

2.3.1 Canalisation Models and Venation Patterns

Auxin transport is central to the regulation of the development of plant architecture. From the polarisation of the embryo, to positioning of leaves and the development of vascular tissue. Canalisation is a particularly intellectually attractive target for modelling because it is a self-organising process, where feedback between auxin and its transport can result in emergent patterns of auxin transport pathways. As described above, the venation patterns observed in leaves, and their connection with the vascular tissue of the main stem, are hypothesised to be established by canalisation. It is also possible that canalisation might be part of the regulatory system controlling shoot branching. It is not yet fully understood how this single mechanism is able to generate the wide range of venation patterns observed in nature, such as parallel veins, branching veins and reticulated networks[105]. Computer modelling and simulation plays a central role in progressing the fields understanding of the regulation of auxin transport canalisation.

Although the first mathematical models of polar auxin transport date back to 1966[60], it is the computer models developed by G. J. Mitchison[73, 74, 75],

informed by the experimental work of Tsvi Sachs[100, 102], that are the foundation of canalisation modelling today. Using pea stem segments, Sachs conducted a number of experiments with different arrangements of auxin sources and sinks. He was able to show that vascular strands differentiate in a narrow path connecting auxin sources, such as exogenously applied auxin, to auxin sinks, such as the auxin transporting tissues of existing vascular strands. Mitchison proposed and modelled two possible mechanisms for canalisation: polar transport and facilitated diffusion. Mitchison's models were executed over a 2D grid of square cells with auxin moving directly from one cell to the next (i.e. no apoplast). This is a common technique used in tissue simulations, where the behaviour of cell membranes and cell walls is simplified into equations that determine the amount of auxin that moves from one cell to its neighbour. In Mitchison's polar transport model, the transporters are recruited to the membrane of a cell in proportion to the flux of auxin crossing that membrane: the greater the flux, the more transporters are recruited, and the more auxin is transported. This positive feedback can result in the formation of files of auxin-transporting cells linking an auxin source and a sink. More complex patterns, similar to leaf venation patterns, can also be generated, such as the linking of two canalised cell files to form closed loops[73].

Mitchison's early canalisation models predicted high auxin flux but low concentration in the canals when compared to the surrounding tissue. The biological evidence, acquired through experiments using reporters such as DR5:GFP, suggests that the canals in fact have high auxin concentrations[78]. Kramer[50] has shown that canals with high flux and high concentration can be formed with the addition of auxin influx transporters that actively pump auxin into cells in the canal from the surrounding apoplast, or efflux transporters that pump auxin out of cells bordering the canal towards the cells of the canal.

High concentration canals can also be generated by varying the way in which transporter allocation is regulated by auxin flux. Based on Mitchison's work, Feugier et al[22] simulated canalisation events in a leaf, represented by a grid of hexagonal cells. Canals of high auxin concentration and high flux can be achieved by having the six faces of the cells compete for PINs from a single available pool within the cell, which may vary in size, instead of PIN being delivered to each face independently from an essentially infinite pool. This competitive method of PIN allocation is able to produce a branched pattern of veins, but without the closed loops seen in leaves. If the cell faces acquired PINs in proportion to their flux, but independently of each other, the resulting

canals showed low auxin concentration (as in Mitchison[73]). The nature of the relationship between auxin flux through a particular cell membrane and the allocation of PINs to that membrane is also an important factor in the formation of canals in Feugier et al. model[22].

Mitchison's second proposed mechanism for auxin transport is facilitated diffusion. The facilitated diffusion idea is controversial because it is very mechanistically obscure. In this model auxin is able to diffuse between cells in either direction with a diffusion constant that increases with flux. Mitchison was able to show that this mechanism was able to produce an open, branching vein system. More recent modelling of this mechanism has shown that in some cases, a facilitated diffusion model is more efficient than a polar transport model at joining a source of auxin to a sink[108]. The two methods were compared on the same layout of 2D cells, with one auxin source and one sink. The polar transport model produced a vein that initially went past the sink but was able to connect to it later on. However, with source and sink in the same position, facilitated diffusion was able to join them along something close to the shortest path.

2.3.2 Phyllotaxis Models

Like canalisation, phyllotaxis is also a self-organising system with dynamic changes in auxin transport at its centre. As such, it has been a strong focus for modelling approaches. Phyllotaxis occurs in the shoot apical meristem (SAM). The SAM has a domed shape and is divided into two areas. The central zone (CZ) is a small group of undifferentiated stem cells that divide to replace themselves while feeding daughter cells into the surrounding peripheral zone (PZ). In the PZ, cells are able to differentiate into leaf founder cells. Leaf specification apparently occurs in the epidermis (the outermost cell layer), triggered by local auxin accumulation[95]. An early event is the formation of the leaf's mid-vein in the underlying tissue, which provides the vascular link between the developing leaf primordium and the existing vasculature in the stem below.

The spacing of successive leaves during phyllotaxis is highly regular. Hofmeister[42] suggested that a new primordium would form as far away as possible from all existing primordia; this led to inhibition models of phyllotaxis, where primordia produce a factor that prevents new primordia forming nearby. Models of phyllotaxis such as those described in [57] have followed this basic principle, with either a lack of activator or presence of an inhibitor, or both, ensuring

that primordia form with the correct spacing. However, auxin-induced patterning is different from reaction-diffusion models since active transport has to be considered[108]. Therefore more recently, with mounting evidence that phyllotaxis is regulated at least in part by PIN-mediated patterns of local auxin accumulation, phyllotaxis models have focused specifically on auxin and its transport[97, 45, 19, 109, 41, 104].

Up-the-gradient models

The fact that PIN proteins orient toward the sites of leaf specifications at which auxin accumulates has suggested that a canalisation-type mechanism, in which auxin fluxes direct PIN localisation, may not be able to account for phyllotaxis. An alternative system for PIN polarisation has therefore been suggested in which PIN proteins are preferentially allocated to the membrane facing the neighbouring cell with highest auxin concentration[41]. These models are here referred to as up-the-gradient models. Once one cell has a slightly higher auxin concentration than its neighbours, for example as a result of stochastic fluctuations in synthesis, degradation or transport, the neighbours preferentially orient their PIN proteins towards it. This creates a positive feedback loop that accumulates auxin in a small number of cells and depletes the surrounding tissue of auxin. Models using this mechanism are able to reproduce the positioning of primordia seen in phyllotaxis on both static and growing 2D tissue grids[45, 19, 104]. These models depend on the ability of cells to report their auxin concentration to their neighbours, by an as yet unknown cell signalling mechanism.

If these models are correct, the PINs of the epidermis are positioned based on auxin concentration. However, the formation of the midvein immediately below the site of auxin accumulation involves PIN orientation towards low auxin concentration in a source-to-sink pattern typical of classical with-the-flux canalisation models. This means that either there are at least two distinct mechanisms for positioning PINs, capable of operating in close proximity, or one of these mechanisms does not occur.

Unifying models

Attempts have been made to develop simulations that can explain both phyllotaxis and canalisation, using only one PIN allocation mechanism. Merks et al.[70] propose the travelling-wave hypothesis, an up-the-gradient based concentration-dependent mechanism where a transport canal in a leaf could

be formed by PIN polarising towards the neighbouring cell of highest auxin concentration. Modelled on a leaf template, an unstable local auxin maximum develops at the leaf margin. The cells in the maximum align their PIN towards their neighbours of highest concentration, which causes the maximum to move, leaving a trail of polarised cells behind it. This model has yet to be rigorously tested against experimental evidence.

On the other hand, Stoma et al.[111] have shown that it is possible to model phyllotaxis using only flux to control the polarisation of PIN. They show that flux-driven PIN polarisation can occur both down and up an auxin concentration gradient, based on the definition of two types of auxin sinks. Weak sinks lead to auxin maximum formation in the L1, while strong sinks induce the formation of provascular strands in the tissue below. During the simulation, auxin moves towards weak sinks. As soon as auxin concentration in a weak sink cell reaches a critical level, canalisation towards the inner layers of the SAM, a strong sink, is initiated. Above a threshold of auxin concentration, sink cells change their identity from epidermal to leaf primordia cells. This identity parameter allows for a developmental switch between phyllotaxis and canalisation. Interestingly, while both mechanisms are flux-driven, the choice the system makes between two types of fluxes does not happen gradually, requiring a binary switch mechanism. The models also predict a transient reduction in auxin concentration at the point of leaf specification. There is currently no evidence to support this model prediction. However, it is worth noting that the mechanism can also produce other developmental patterns, such as the allocation of PIN proteins observed in roots.

An alternative approach was proposed by Bayer et al.[5], who developed a hybrid model with both up-the-gradient and with-the-flux PIN polarisation. The two processes are combined in the simulation so that at low concentrations of auxin the polarisation is up-the-gradient but as the auxin concentration increases, the mechanism changes to with-the-flux. This system is able to simulate primordia formation in the SAM epidermis, which once formed can produce a midvein through the inner layer, joining to an existing sink—the vasculature of the stem. This model predicts a transient polarisation of PIN in the inner tissue towards the epidermal layer, near the site of midvein formation. This prediction was subsequently observed in both *Arabidopsis* and Tomato meristems. One difficulty with this model is that although a midvein can form, it reliably connects to the sink only if the sink produces an additional attractive signal to advertise its location. There is currently no evidence for

the existence of such a signal, although there is a long-running debate in the literature about the role of the stem vasculature in contributing to phyllotactic patterning in the meristem above[59, 58].

2.3.3 Static PIN Models

Not all models allow dynamic PIN polarisation. Some models statically allocate PINs to membranes, often to investigate a particular aspect of auxin transport regulated patterning. Even without feedback regulation on PIN, the complex patterns of localisation observed in plant tissues make predicting resulting auxin distribution non-intuitive, especially in growing tissues. To help understand these systems, models using patterns of PIN localisation that are obtained from immunocytochemistry of real plant tissues have been developed[39, 15]. These PIN accumulation patterns are used as templates by which cell-to-cell transport of auxin is determined. Similar to this, there are also attempts which do not use explicit PIN localisation patterns, assuming homogeneous PIN distribution on cell membranes[64]. One application of these static approaches is to investigate the robustness of auxin transport routes and auxin distributions, such as the stability of an auxin maximum in root tips[39] or the dynamics of auxin accumulation in the SAM[15]. Dynamism in PIN allocation in these systems is limited to PIN redistribution caused by cells changing their fate and thus adopting a different PIN accumulation pattern[39], or by limited auxin redistribution such as occurs during gravitropism[121, 92].

2.3.4 PIN polarisation mechanisms

In cells the trafficking of PIN to the cell membranes appears to be dependant on the protein GNOM[36, 49]. However how PIN becomes polarly localised is not understood. Much discussion about the validity of different auxin transport models centres on the plausibility of the underlying PIN localisation mechanisms. The simplifying and mechanistic assumptions of a model must strike a balance between remaining faithful to the biology and minimising complexity. Often it is the assumptions that can tell us the most about how a model works and may indicate where we need more biological data. Figure 2.2 summarises the important aspects of auxin and PIN biology that need to be captured in a model.

For example, for many canalisation models auxin flux is a key piece of in-

formation required to achieve correct PIN polarisation. Models rely on having information about how much auxin is crossing each membrane of a cell. The membrane that is experiencing the greatest amount of auxin flux accumulates either the largest share of a pool of PINs or accumulates PIN in proportion to the flux across it. This presents a difficulty as currently there is no known biological mechanism for measuring auxin flux.

Kramer[52] showed that it is possible to use the internal gradient of auxin created by PIN activity to polarise the PIN proteins in the direction of an auxin sink. In this model PIN mediated efflux would cause auxin depletion from cytoplasmic areas around the PINs. This would lead to an intracellular gradient of auxin. In Kramers model PIN is preferentially delivered to the membrane at the side of the cell with a lower cytoplasmic auxin concentration, this leads to polarisation towards an already existing flux. Although this model results in PIN allocation proportional to auxin flux, it is also mechanistically difficult to envision because it requires a gradient-sensing molecule in the cytoplasm. A slight variation on this model would be to allow PIN to be randomly delivered to the membrane of the cell, but preferentially removed from membrane areas neighbouring cytoplasm with high auxin concentration. This would allow a polarisation towards and existing flux and may remove the need for a cytoplasmic gradient-sensing molecule. It is possible that transport of auxin across the cell membrane sets up a proton gradient across the membrane. Apoplastic acidification increases auxin fluxes over the membrane which in could be a potential signal for the localisation of PIN[3].

Another possibility for cells to measure flux would depend on tally molecules localised at the membrane[14]. As one auxin molecule moves through a cell, it causes a tally molecule to be moved to the site of its efflux. This would lead to an unequal distribution of tally molecules corresponding to the direction of auxin flux.

In up-the-gradient models the localisation of PIN proteins in a cell is hypothesised to occur preferentially towards neighbours with higher auxin concentrations. Although there are known mechanisms by which overall cellular auxin concentration is sensed, how this information might be transmitted to neighbouring cells is unknown. Two ideas emerging from modelling are that an unknown signalling molecule transfers information about intracellular auxin concentration between cells[45], or that auxin-induced changes in cell expansion generate physical forces that are sensed by neighbours[38, 79].

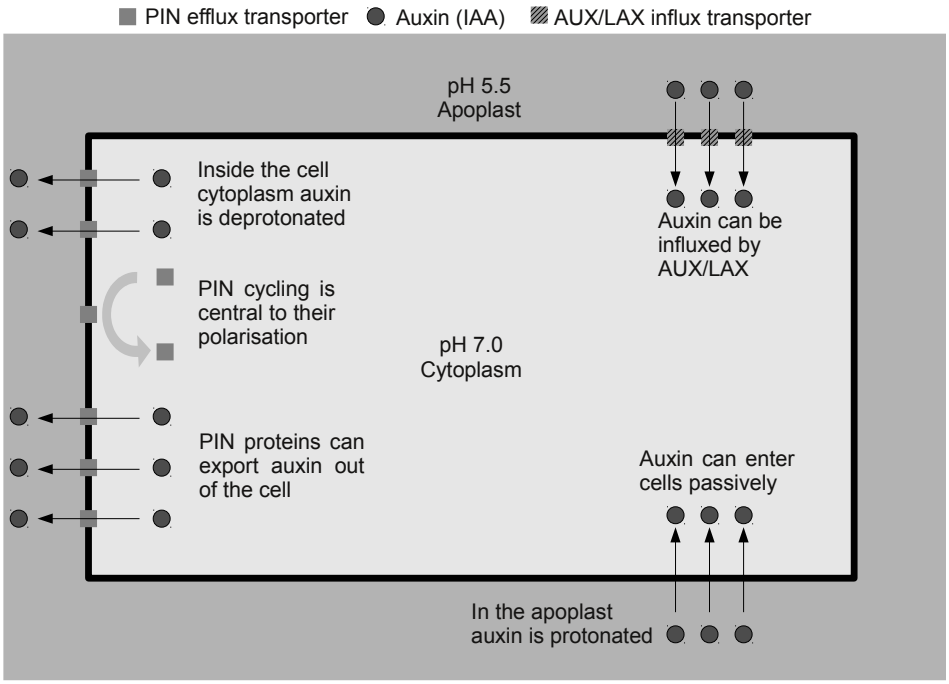


Figure 2.2: Auxin transport into and out of cells is central to canalisation. Protonated auxin in the apoplast is able to enter the cell passively, or to be actively influxed by AUX/LAX transporters. Once inside the cell the majority of auxin is deprotonated and is therefore unable to leave the cell unaided. This is often known as the Acid Trap hypotheses [99, 93]. PIN transporters are important to the efflux of auxin from cells. The regulated cycling of PINs on and off the cell membrane causes them to become localised asymmetrically around the cell membrane. This process is not fully understood, but is critical to the directional transport of auxin in tissues, and the process of canalisation.

2.3.5 Whole Plant Models

Whole plant models have the advantage of allowing us to investigate the role of mechanisms like canalisation in a wider context, including integrating aspects of the plant's external environment. Prusinkiewicz et al.[90] hypothesise that the canalisation described by Mitchison[73] can be implemented at the level of entire shoot segments, allowing efficient modelling of the polar transport stream. As described above, one role for this auxin is the inhibition of shoot branching through prevention of the activation of axillary buds formed at the base each leaf. Experiments with radiolabeled auxin have shown that auxin moving down the stem does not enter the axillary buds to inhibit them, but rather the effect is indirect. Prusinkiewicz et al.[90] propose that this indirect mechanism is mediated by an auxin transport switch. If canalisation of auxin transport out of the bud is required for bud activation, auxin moving in the stem might reduce its strength as a sink, preventing auxin transport canalisation from the bud.

This model makes an assumption that the many cells of the stem can be approximated as a single segment of auxin transporting tissue. This reduces computation time, and thus allows whole plant level modelling. The model uses a number of these segments to represent the stem. Each stem segment has an attached "bud" which is a potential source of auxin. The nature of the transport switch is demonstrated in a small model of two sources, both connected to a stem segment. The stem is connected to a sink. If either one of the sources is activated auxin begins to flow into the stem - there is a net flux from the source into the stem. This flux increases the PIN on the membrane between the source and stem. Increased PIN causes increased flux, and very quickly positive feedback causes polarisation of PIN towards the stem. If the second source is then activated it now has a high auxin concentration. However, the net flux from source to stem is very low because of high auxin in the stem. The positive feedback on the flux does not occur and the source does not polarise. If both sources are activated in very quick succession, both experience a high net flux in the direction of the stem for long enough for the positive feedback to cause polarisation, and both canalise auxin transport into the stem.

Using this simple switch mechanism in larger growing plant models, where new segments with "buds" are added at the apex of the plant, the authors were able to reproduce many aspects of plant branching behaviour, including basipetal and acropetal branching patterns and the response to decapitation.

Furthermore they were able to model the behaviour of a number of *Arabidopsis* branching mutants and double mutants affecting auxin transport by changing single model parameters consistent with the known molecular functions of the genes affected by the mutations.

2.4 Modelling Methods

Building biological models of any type presents a significant challenge. Care must be taken to ensure that enough biological detail is captured without over complicating the models. This is achieved by making well-informed, thoughtful simplifying assumptions, and working at the appropriate level of abstraction. In order to help achieve this goal some model developers use frameworks to support the development process. These include fairly traditional software engineering tools such as UML[81] and also tools developed specifically for developing biological models such as the Systems Biology Markup Language (SMBL)[71].

2.4.1 Modelling Biology and Simulations with UML

The Unified Modelling Language (UML) is a modelling language comprising a suite of diagramming notations originally designed for use during the development of large complex software engineering projects[81]. It was developed to help large groups of programmers work in parallel in smaller sub groups to develop software for the same project by defining how the program parts interact. As long as the interaction of the separately programmed objects is correct then the program parts should work together when they are combined into the finished product.

There are a large number of philosophies used when developing software with UML. The process is normally top down from a set of high level requirements. The process can start by looking at high level interactions in the system being described. For example, this would include how the system interacts with its environment, where the environment is either various users or other systems. This is normally modelled in a Use Case Diagram coupled with written usage scenarios.

The high level model is gradually refined by adding detail. Further diagrams are used to drill down details of how the system is built, what the objects are, what information they exchange and when those exchanges of information need to take place. This eventually results in a code skeleton being

created for the objects with the attributes in place and the methods waiting to be implemented.

Two good examples of software engineering processes that are frequently used with UML are the Evolutionary, and Waterfall models. The Waterfall model is a sequential software development process that starts with the identification of requirements and moves through design, implementation, verification, and maintenance stages. This has many of the requirements for a simulation development process, but the original Waterfall model did not encourage iteration through the process to refine the software being developed. Evolutionary software engineering was developed to take the concept of "evolution" into an engineering paradigm and therefore the modelling process is iterative. This makes this process more suitable for developing biological simulators[110].

As well as classic object-oriented technologies, UML is well suited to agent-based modelling[80] (where an agent can be thought of as an object with its own thread of control, allowing highly parallel systems of multiple agents). Biological 'agents', such as cells and proteins, can be modelled as UML agents. There are many processes in cells acting in parallel. Some of these processes are individually sequential, such as expression of proteins in response signals detected in the cell. The signals might cause a number of events such as protein expression, which then in turn causes more events to occur. This sequential behaviour is often called a cell pathway. The parallel behaviour comes from this type of process occurring in a number of different pathways in one cell of the plant at the same time, and possibly in many cells at the same time throughout the plant.

The combination of UML and object-oriented programming maps naturally to the biological processes that we are modelling. Biological objects, such as proteins and cells, map directly to objects in the UML models, which are then implemented as objects in the program code. The interactions between these biological objects similarly map directly to associations between objects in the UML models, which are then implemented as communications between objects in the program code. UML class diagrams can be used to describe the objects and can show some of the attributes of the objects and type of processes that object might perform. Attributes might include the parameters that define the behaviour of the object, and the processes could be things like the production of another type of object. The combination of attributes and methods describing the biological function of the object, and the links between objects in a class diagram can provide information about where objects interact. For example,

a cell object might have a connection to an auxin object and contain an auxin production method. Indicating that cells produce auxin.

This allows us to build models containing the biological objects that we believe to be involved in canalisation, and then produce simulations that we can use to test various hypotheses about the biological processes of interest. If an hypothesis is correct we should see the correct emergent behaviour when the simulation is run; if not we can then return to the UML models and alter the simulations to test our next hypothesis. This provides a process to assist us in determining if our simulated biology is consistent with the real biology.

Sate diagrams can also provide useful insight into biological processes. State diagrams show how an object changes between different possible internal states. In a simulations the different states can map to different biological processes. An auxin agent for example could be diffusing, and therefore in the diffusion state. State diagrams cannot capture the spacial aspects of biology. If auxin changes its state depending on what environment it is currently in these changes are difficult to capture. None of the UML diagrams are capable of natively capturing space.

Time is another concept that is very important to biological processes and has to be carefully considered in models and simulations. UML can capture the order in which some events happen in sequence or collaboration diagrams. These diagrams capture message passing between objects in a program. They are difficult to apply in a biological context as communication between biological agents has no guarantee of success.

Despite there limitations UML diagrams remain a useful tool for assisting with the development of biological simulations and are relatively accessible to biologists. Allowing them to provide input to the model of the simulation without the need to understand the code. There are a number of published cases where UML has been successfully used to produce biological models[20, 46, 118, 94].

2.4.2 Other Model Engineering Systems

The process of developing biological models with UML is similar to the processes used for development using the Systems Biology Markup Language (SBML)[44, 23, 43, 71]. However, as we are implementing our programs in Java, the ability to produce code skeletons easily and flexibly from UML, with programs such as NetBeans or Rational Rose, is an advantage. UML is thought as being implementation neutral, as the diagrams can be parsed into a wide

variety of different outputs.

Unfortunately most other existing models of auxin transport have either not been developed using a systematic process or the details of how such a process was used have not been published. This is unfortunate as more detailed information about how a model is constructed can improve the reproducibility of models, especially of the original source code is not available. Knowing how a model was constructed can also provide insight to how suitable a model is for reuse. Understanding the assumptions made when developing a model can help modellers decide if a model can be adapted for different use.

2.5 Executable Models

There are many mathematical models of nearly all aspects of plant development [91], and many concerned with auxin transport [51]. Executable models are version of a model (that could be formally described by UML) that can be run on a computer. This work is to develop an executable model of auxin transport. Executable models could prove to be well suited to the modelling of biological systems, and might offer an alternative perspective that yields results [25]. As the models produced by biologists to describe how they believe a process is occurring can be formalised and then turned directly into an executable model. The models could either be written or diagrammatic, or both. Other aspects of biology have been modelled using this technique. These include, for example, some computational models of the nematode worm *C. elegans* [24, 26].

This thesis describes work to develop an executable model of auxin transport canalisation. To date all models of auxin transport have been mathematical models. An executable model could provide additional insight that cannot be gained from mathematical models alone. Agent based models can be targeted at a lower level of abstraction and remain intuitive. The behaviour of the simulated system comes from the interaction of the agents with each other and there environment. Not from an abstract set of equations. There comparing the output from a mathematical model and an simulation might provide insight into if assumptions in equations are effecting the results. This could be particularly useful if the occurrence of rare events is important in the system of study.

2.6 Emergence

An exact definition for what emergence is remains a topic of some debate. One possible simple definition is that a system that exhibits emergence properties is greater than the sum of its parts. Or as Aristotle put it.

> ... the totality is not, as it were, a mere heap, but the whole is something besides the parts ...

Emergence is a difficult part of the modelling process. Normally emergence is something to try to achieve. In the case of auxin transport the desirable emergent property is the development of the transport that you are interested in seeing; either a canal or an area of high auxin concentration in phyllotaxis. During the modelling process it is useful to identify desired emergent properties, but it is important that the emergence is not hard coded into the model. Rather it should be a consequence of the underlying mechanisms that the model is attempting to simulate. If the emergent property is not seen the underlying mechanism should be altered to attempt to produce it.

The approach used in this work in some ways has emergence at its centre. The simulations at their lowest level a collection of interacting agents. Where one agents interactions are with its environment, which is other agents and the space that they are in. All the agents in the system have behaviour encoded into them in the form of a set of instructions that they attempt to carry out. The success of the instructions is at least in part dependant on their environment. The agents have no knowledge of the greater system, and can only *see* their immediate surroundings. This means that the behaviour of the system (all the agents in the entire space) is sum of all the interactions of all the agents. In the auxin transport canalisation simulations the instructions given to the agents should say nothing about how to build an auxin transport canal. The canal should be an emergent property, greater than the sum of the instructions given to the individual agents.

2.7 Conclusions

Auxin transport is central to many aspects of plant developmental biology. Understand to processes involved in its regulation is a difficult challenge. There is a long history of modelling auxin transport, and in many ways the field is a good example of how modelling techniques and biological techniques are progressing together. A feedback is developing between the models and biology

that can only benefit them both. Increasingly the major labs studying auxin transport have full time modellers in the group. The work presented in this thesis has two major aspects to it. The work attempts to apply a new technique to the problem of auxin transport canalisation, executable agent based models. In the hope that these might provide addition insight into auxin transport itself. The work also follows are clearly defined process for developing a simulation from biological data. Following such a process could also have benefits for the wider community as it allows the inner workings and assumptions of a model to be presented clearly.

2.8 The Context of this Work

This thesis describes the development of models of the formation of auxin transport canals, particularly the transport canals that might form in the stem of plants and often go on to differentiate into vascular tissue. The work focuses on auxin transport canals that form between sites of auxin production and auxin removal–sources and sinks–described in section 2.2.1. The models have been developed to test hypothetical mechanisms for the regulation of PIN polarisation, a critical part of the canalisation process. With the longer term goal of investigating how canalisation could be involved in biological processes like shoot branching regulation (see sections 2.2.2 and 2.3.5) and midvein formation (see section 2.3.2).

Initially the work followed a simple approach that could be considered similar to an evolutionary software engineering process where UML is developed, turned into the simulation code and tested. Then to help ensure the models and code were consistent any necessary alterations were fed back into the beginning of the process. This developed into a more rigorous and systematic development approach supported by the CoSMoS process that produced the first generation simulator, presented in Chapter 3. The first generation simulator has been used to successfully test hypotheses of PIN polarisation on abstract cell tissues consisting of grids of rectangular 2D cells in various arrangements. Results are presented in Chapters 4 and 4.3. Chapter 5 describes the process of developing the second generation auxin transport simulator. This is an enhanced simulator that was developed to allow hypotheses to be tested in cells of more natural shapes in 2D. Results are presented in Chapter 6. The second generation model was also developed to allow testing in 3D. Chapter 7 describes the current state of the unfinished 3D model. Finally both

the process of developing the simulators and the results produced are discussed in Chapter 8.

Chapter 3

Development of First Generation Model

Increasingly biology is looking to modelling to help progress understanding. Developing a simulation of a biological process is a challenging task in itself, but doing so can assist with progressing our understanding of complex biological processes. The modelling process requires the builders to go systematically through the information and data about a system, ideally with experts in the field. Simply going through this modelling process can highlight new areas of focus, or problems and gaps in understanding. The resulting simulation and models can also be a tool for the generation and testing of hypotheses, hiding some of the complexity of the real system but capturing enough to allow the study of the process of interest.

The level of abstraction in a model is critical. Too high, and we risk ruling out the possibility that simulations will produce interesting emergent behaviours that are observed in the real system. Too low, and the simulations produced could be difficult to work with, understand and validate. These factors make the design decisions made when producing a simulation important, as they determine the balance between these conflicting requirements.

This chapter of the thesis describes the development of the first generation simulator. The simulator was developed from some preliminary work[32] and other early toy models to be a flexible platform for the testing of PIN localisation mechanisms in the context of auxin transport canalisation. The CoSMoS (Complex Systems Modelling and Simulation)[2] process was used to assist in the development of the simulator.

A simulation must be developed using a rigorous process of design, implementation and validation if it is to be scientifically respectable. Additionally,

a useful simulation will need to be upgraded and enhanced in a principled manner as its requirements change to address new research questions. The CoSMoS process[2] provides a flexible approach designed to support the modelling and analysis of complex systems, including the design and validation of appropriate computer simulations.

3.1 Overview of the modelling process

The CoSMoS program[2] is developing generic modelling tools and simulation techniques for complex systems. The principle aims are to support the modelling, analysis and prediction of complex systems, and to help design and validate complex systems. This process can be used to produce *simulations* of complex (biological) systems, by implementing the abstract computational models to produce simulators. The work presented in this chapter uses the naming convention outlined in Andrews et al. [2].

Fowler [27, p.51] identifies three perspectives that can be used when building models: the *conceptual* perspective, representing the domain under study; the *specification* perspective, representing the software being designed in terms of its interfaces; and the *implementation* perspective, representing the detailed software design. (Later editions collapse some of these perspectives.)

The CoSMoS project is developing a complex systems simulation development process based on this conceptual model and Fowler's three perspectives. The CoSMoS lifecyle used during the development of this first generation simulator is summarised in figure 3.1. The arrows indicate the direction of the information flow, and are not meant to indicate a path through the CoSMoS process. It identifies the following:

Domain Model: a "top down" conceptual model of the real world system to be simulated, derived in this case from the biological data, from the literature, and observations and experiments needed to provide sufficient data for modelling. Some modelling decisions about what to put in and what to leave out are made here. The model may explicitly include various emergent properties, since from a top down perspective it may not be obvious that these are emergent; or, if we are aware of the emergent properties, it may not be obvious what low level processes produce them.

Software Model: a "bottom up" model of how the real world system is cast as a simulation. This includes: a definition of the system boundary (what parts of the Domain Model are being simulated); simplifying assumptions and

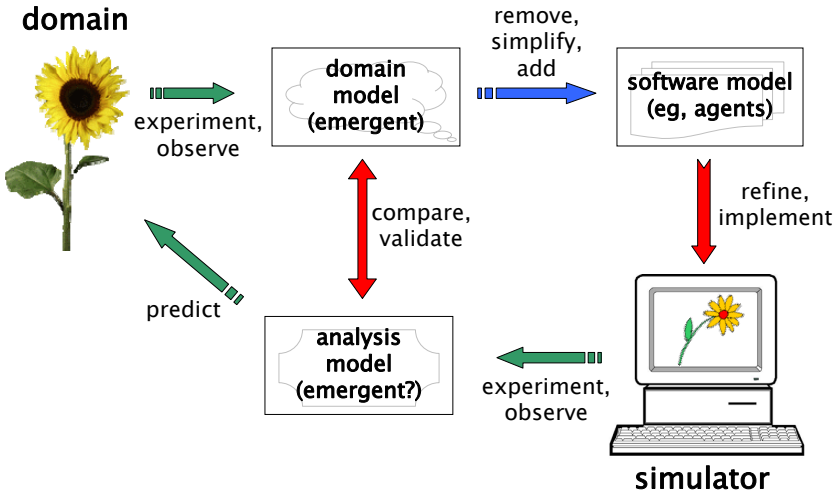


Figure 3.1: The components of the CoSMoS basic lifecycle. The arrows indicate the flow of information between the different steps (updated lifecycle [2]).

abstractions; removal of emergent properties and replacement with the local interactions that are hypothesised to result in them; extra simulation-only concepts, such as "physics" engines to implement real world processes in possibly unnatural ways, user interfaces to view and control the simulator, and "probes" to produce output data for analysis.

Simulator: the executable implementation. The development of the Simulator from the Software Model is a standard software engineering process.

Analysis Model: a "top down" conceptual model of the *simulated* world, derived from observations and experiments on the simulation. The model may explicitly include various observed emergent properties. This model is compared to the Domain Model in order to test various hypotheses, such as the validity of the simplifications used to derive the Software Model. Analysis models have not been produced for the simulator.

This process is neutral in its choice of modelling language(s). For example, it could use a mix of text, biological 'cartoons', Soft Systems' Rich Pictures [12], and mathematical equations to describe the Domain Model, and any standard software engineering technique to define the Implementation Model. This work uses UML supported with text.

This process allows us to separate implementation details from the biology being simulated. This offers a number of advantages; it makes the individual models and accompanying diagrams simpler, as they are focussed on specific perspectives. As we are partly using UML as a communication tool it is advantageous for the diagrams to be as simple as possible. Different groups are

more interested in certain perspectives: for example, biologists are probably on the whole more interested in a clear representation of the biology, rather than how the data I/O and GUI work.

3.2 Domain Model: auxin transport

3.2.1 The biological background

The simulator has been designed as a tool for testing hypotheses of PIN localisation particularly in the context of auxin transport localisation. The detail of the biological background to the simulations has been presented in Chapter 2.

3.2.2 Modelling

The motivation for developing the models is to help progress understanding of the process of auxin transport canalisation, with the longer term goal of investigating its role in auxin shoot branching regulation. The process of auxin transport canalisation is clearly complex, and is not fully understood. This makes it a natural target for modelling, but also a challenge.

A successful executable model of canalisation has been developed by using the CoSMoS approach, in UML[81], starting from the background biology described earlier. The background biology was developed into a UML Domain Model that includes all the necessary biology of the model to function, but keeps the model as simple as possible. The UML Software model was developed from the UML Domain Model. This was then refined to develop the simulator program itself. The refined UML System Model can be used to produce code skeletons to assist with simulator implementation. Using this process helps ensure that the reasons for what is included in the various models, and how it got there, are understood. It is possible to make comparisons with how the model simulates biology compared with how we think the biology works.

It was not necessary to use all of the diagrams available in UML; described are only the ones that have been found to be of greatest use. Most useful are the diagrams that map intuitively to describing biological objects, and how they behave (Class and State diagrams). Less useful are the diagrams that describe how object communicate with each other (Sequence and Collaboration diagrams). As the number of people developing the models is small in this case, it was not necessary to use one of the more rigorously defined UML development processes common in software engineering. The models are imple-

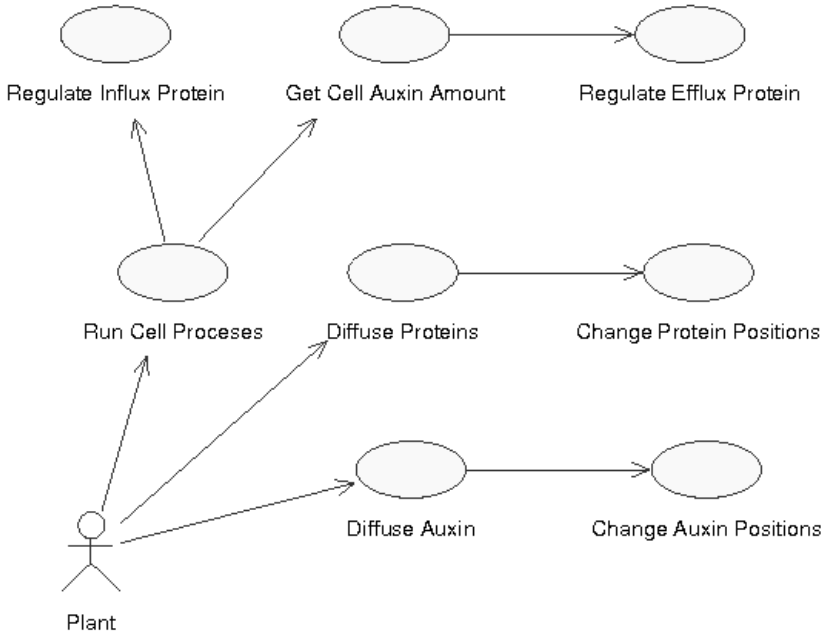


Figure 3.2: The Domain Model Use Case diagram. This covers the high level biological aspects that need to be included in the simulator, with the Plant itself as the actor.

mented in the Object Orientated programming language Java[83]. Using UML and the OO programming paradigm in Java for the production of biological models produces simulations that map intuitively back to the biology.

3.2.3 Domain Model Use Cases

The UML Domain Model is built from the background biological material. Starting with a Use Case diagram to produce a high level representation of what is happening in the plant that needs to be in our model, such as expression of proteins and transport of auxin (figure 3.2). This is the first high level filter, where a decision is made about what is required to be in the model in order for it to fulfil the requirements.

3.2.4 Domain Model class diagram

The biological components of interest are modelled as objects and classes. This approach works well here because much of cellular biology can be thought of as interactions of discrete objects that result in complex behaviours. Therefore we consider the different parts of the cells, such as the cell membrane and vacuoles (cellular compartments), the proteins like PIN and hormones like auxin, all as objects. One of the objects of interest at this level of model is

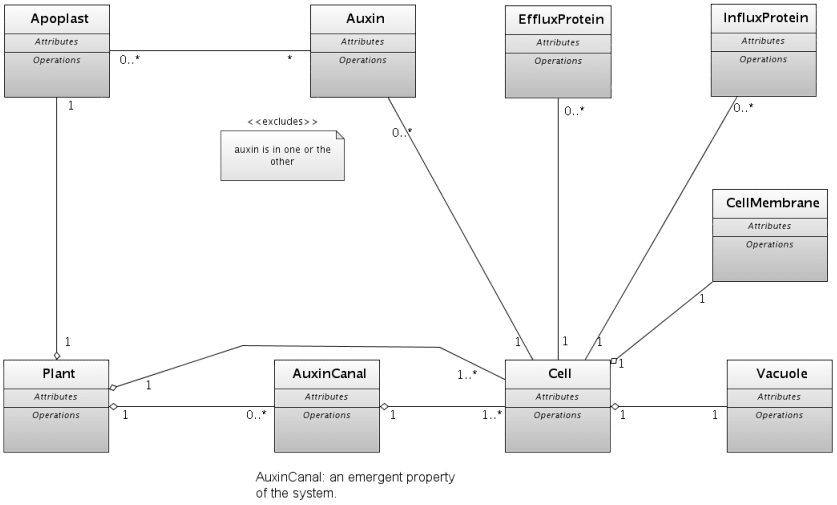


Figure 3.3: The Domain Model class diagram. The biological components of the simulation are modelled as interacting objects defined by classes.

the auxin canal. This is an emergent property of the lower level interactions. We model it explicitly here to capture its biological properties so that later simulation outputs can be related to it.

Figure 3.3 shows our Domain Model class diagram of the biologically relevant parts of the system (deciding what *is* biologically relevant is also part of the modelling process). In detail, it shows the following classes (type of objects) and relationships between the objects:

- A Plant has one Apoplast (the space between cells), one or more Cells, and an optional Auxin Canal. (It also has other components, but these are not being modelled, even at the system level.)

- An Apoplast is part of one Plant, and has zero or more Auxin molecules.

- An Auxin molecule is in the Apoplast or in a Cell. (The relationship lines say that it may be in an Apoplast and it may be in a Cell; the excludes condition says that it is one or the other).

- An Auxin Canal is part of one Plant, and has one or more Cells.

- A Cell may be part of an Auxin Canal; it is part of a Plant. It has zero or more Auxin molecules, zero or more Efflux Proteins, and zero or more Influx Proteins. It has one Membrane and one Vacuole.

- An Efflux Protein is in one Cell; an Influx Protein is in one Cell.

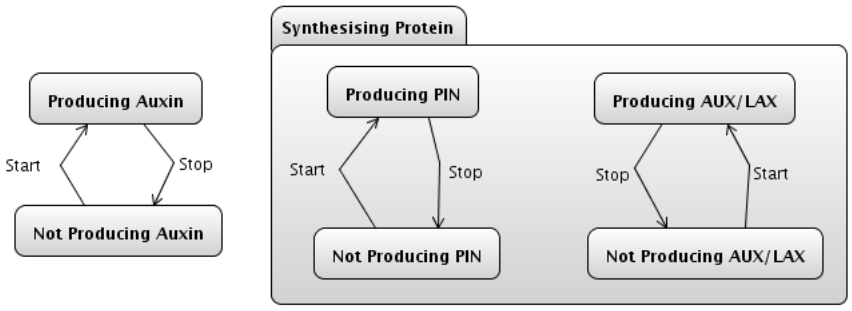


Figure 3.4: State diagram for the Cell. (Since we are not explicitly modelling cell birth or death, no start or end states are needed.)

- A Membrane is part of a Cell; a Vacuole is part of a Cell.
- We impose an extra condition on the loop of relationships containing the Plant, Auxin Canal, and the Cells: Consider a Cell that is part of an Auxin Canal that is part of a Plant, that Cell is also directly part of *the same* Plant. (There is no loop of relationships containing the Plant, Auxin Canal, Cells, Auxin, and Apoplast: the apparent loop is broken by the excludes condition.

3.2.5 Domain Model state diagrams

The class diagrams and use case diagrams help with the organisation of the different objects of the model, but they provide little information about how those objects behave and how they interact. Interactions are often captured in UML using sequence diagrams, and these show the passage of information between objects over time. In biology the order and direction of interactions is less clearly defined: the next step in the interaction sequence might not occur; the process might back up to the previous step in the sequence. This makes capturing timing of events difficult with sequence diagrams.

UML has another way to capture how objects change over time: state diagrams. These diagrams show the different states an object can be in, and how the object moves from one state to another in response to an event.

State diagrams for the state changes associated with a Cell are shown in figure 3.4, for Auxin hormone in figure 3.5 and for PIN proteins (a kind of EffluxProtein) in figure 3.6. The states of these objects are linked, and a change in the state of one object is linked to that of the others. The state of a cell is defined by the what is happening in that cell, which is determined by what the proteins and hormones are doing.

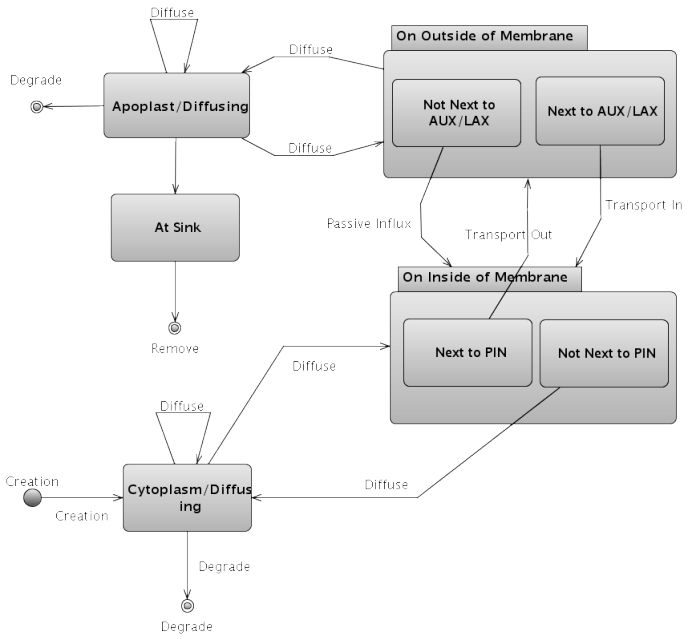


Figure 3.5: The state diagram for Auxin. Defining and expressing this type of complicated behaviour is where state diagrams prove to be useful. The figure clearly shows the different states auxin can get into and the events that put it in those states and the events that can move it into another state.

Figure 3.4 shows that a Cell can be in one of two main states: synthesising, and non-synthesising. When synthesising, the cell can be producing PIN protein (an EffluxProtein), AUX/LAX protein (an InfluxProtein) and Auxin. All cells are capable of making auxin but only source cells do, but they are all capable of producing PIN. PIN is produced in response to the amount of auxin the cell has, so PIN production might be on even if there is no auxin production.

Figure 3.5 shows the state diagram for Auxin. Table 3.1 describes some of the evidence supporting the modelled behaviour of auxin in the simulations. Auxin can be in four main states:

- in the cytoplasm (the inside part of the cell that is not vacuole). It is created here, and may degrade (be destroyed) here. It is in its deprotonated form. It is diffusing around, which either leaves it in the cytoplasm, or moves it to be:

- on the inside of the cell membrane, where it is in one of two sub-states, next to PIN, or not next to PIN. If it is not next to PIN, it diffuses back in the cytoplasm. If it is next to PIN, it is transported out of the cell to be:

- on the outside of the cell membrane, where it is in one of two sub-states,

Auxin: Important Model Assumptions		
Assumption	Description	Evidence
Auxin is created in the cytoplasm of the cell.	In the simulations Auxin is created in a random location of a cell where it is then free to diffuse around the cell. This is a reasonable assumption to make.	Domain experts
Auxin source cells have a fixed concentration of auxin.	Auxin producing cells feedback regulation to keep auxin at a stable concentration.	[65]
Auxin is degraded in the cytoplasm and the apoplast	In the simulations auxin is degraded with a fixed probability from the cytoplasm and apoplast.	Domain experts and [65]

Table 3.1: Important assumptions concerning the behaviour of auxin in the simulations.

next to AUX/LAX, or not next to AUX/LAX. If it is next to AUX/LAX, it is transported into the the cell to be on the inside of the cell membrane. If it is not next to AUX/LAX, it can passively influx into the the cell to be on the inside of the cell membrane, or it can start diffusing to be:

- in the apoplast. It is in its protonated form. It is diffusing around, which either leaves it in the apoplast, or moves it to be on the outside of the cell membrane, or moves it to be at auxin sink, where it is removed from the system. Or it may degrade (be destroyed) here.

In reality, an auxin in the cytoplasm may be no different from an auxin adjacent to a cell membrane: it has no 'sense' of where it is. Therefore the auxin may not have a different *biological* state when it is in these different situations. But we can *model* the biology in terms of such states.

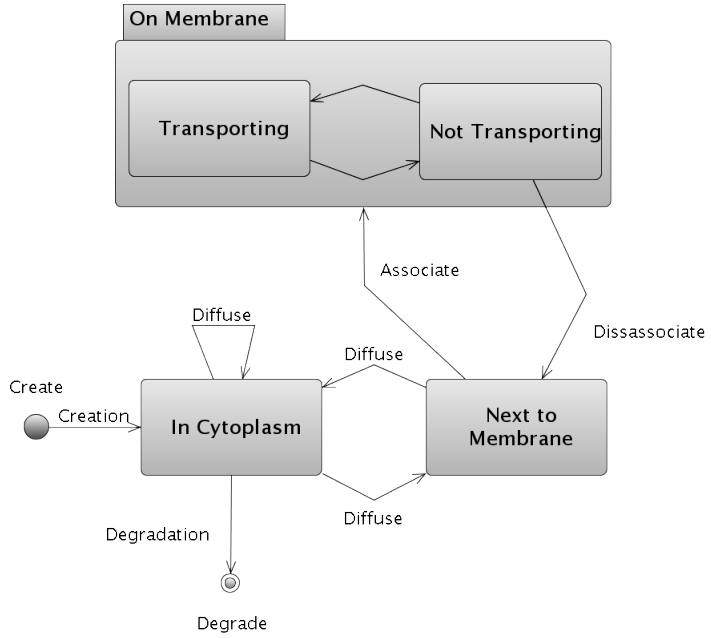


Figure 3.6: State diagram for the PIN Efflux Protein. The PIN is associated with the membrane. It can either be actively transporting auxin or sitting idle. If it is not transporting auxin there is the possibility that it might disassociate and return to the cytoplasm. If it is transporting then it remains attached to the membrane.

State diagrams can be used to model alternative hypothesised behaviours and simulations based on these different hypotheses can be compared. For example, figure 3.6 shows a state diagram for one hypothesis of PIN protein (an EffluxProtein) behaviour, and figure 3.7 shows the state diagram for a slightly different hypothesis for its behaviour. In the latter case, the PIN protein is allowed to move around on the cell membrane if it is not transporting auxin. Therefore the moving state is different from the transporting state when on the membrane of the cell. (Proteins are able to move around on cell membranes [107] and it is theoretically possible that a conformational change in response to actively transporting auxin might stop it from moving.) Table 3.2 states some of the key assumptions in the simulations controlling PIN behaviour. Some evidence come from the the wider auxin transport literature, and some of the domain experts involved in the development of the simulations. Table 3.2 describes some of the evidence supporting the modelled behaviour of PIN proteins in the simulations.

When proteins are being considered, the different states in the Domain Model of the biology correspond more closely to real biological states than in the case of auxin. Proteins are active molecules, and can undergo conforma-

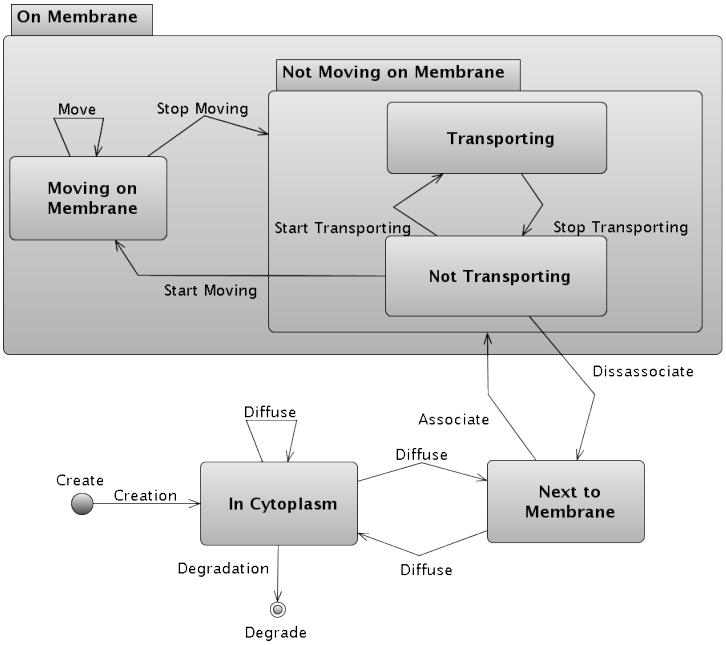


Figure 3.7: State diagram of an alternative hypothesis for the PIN Efflux Protein, with different behaviour on the cell membrane: movement on the membrane is allowed.

PIN: Important Model Assumptions		
Assumption	Description	Evidence
PIN is created in the cytoplasm of the cell.	In the simulations PIN is created in a random location of a cell where it is then free to diffuse around the cell. In a real cell PIN is created in the cytoplasm but its movement inside the cell is highly regulated.	[28, 36, 49]
PIN concentration increases with auxin concentration.	The amount of available PIN in the cytoplasm increases with auxin concentration.	[116]
PIN is degraded in the cytoplasm of the cell.	In the simulations PIN is degraded with a fixed probability from the cytoplasm. Degradation of PIN in real cells is complex.	Domain experts and [48, 68]
PIN responds to auxin concentration	The model assumes that PIN is able in some way to sense its local auxin environment. Either directly or via a signalling molecule.	Domain experts

Table 3.2: Important assumptions concerning the behaviour of PIN in the simulations.

tional and other changes in response to events. Auxin however is a very simple molecule, and more of its behaviour is a passive response to its environment.

One limitation of state diagrams for the depiction of hypotheses of behaviour of objects like PIN in the simulator is that it is difficult to provide information about the circumstances under which a particular event might take place. For example, there might be a number of different possible hypotheses for why PIN should move from the membrane back into the cytoplasm, all of which will look the same in the state diagram. Therefore additional text or other diagrams can be used to provide this type of information.

3.3 Software Model

3.3.1 Software Model Use Cases

The Software Model Use Cases cover the things that our simulation must do to be able to run. These include the set-up procedures required to get the simulation to a starting point, including things like making cells and detecting the internal environment in the cell in order to make things like proteins and hormones. The use cases can be considered from three points of view: that of the plant, which is biological (Domain Model use cases); the simulation of the biology that is required to be there but is not simulated in a particularly biological way; and the things that need to be there to produce a successful simulation but are not part of the biological Domain Model.

As the UML and simulations have developed, the simulated biology has come to represent the real biology as currently understood more closely. However, there are still a few areas where this has not been possible to achieve. For example, when a cell is created in the simulation it is necessary to make the cell and then make its membranes. In reality membranes partly define a cell: a cell cannot exist without a membrane. Therefore the use cases are showing us that our simulation is not doing cell creation, or growth, in a particularly realistic way biologically. In the simulation an abstract object called a cell is created, and is then given a membrane, a vacuole and a starting amount of proteins. It would be more realistic for a cell be the outcome of a particular arrangement of cell membrane, vacuole and other cell elements. The increased flexibility of this could in the future allow for simulation of growth, the lack of which is currently a limitation.

Use cases of the Software Model could also show things like what the graphical user interface (GUI) does, and what I/O the simulation needs to do to pro-

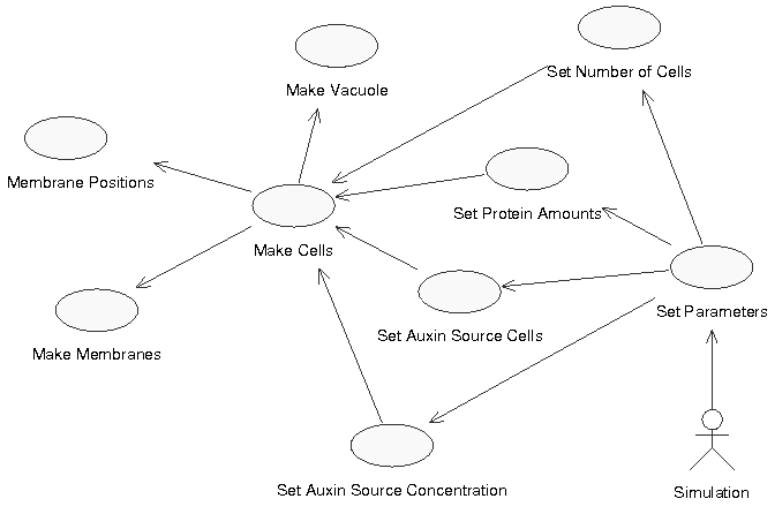


Figure 3.8: The Software Model Use Case diagram, with the Simulation user actor.

vide the user or external systems with results. Such use cases could be a way of modelling high level interaction of different simulations. Figure 3.8 shows the Software Model Use Case diagram that augments the Domain Model use case diagram with use cases for the implementation of the simulation.

3.3.2 Software Model Class diagram

Figure 3.9 shows the class diagram with only the biologically relevant parts of the Software Model shown: to improve its readability, it is missing things like the data and visual output objects.

This Software Model class diagram is produced from the Domain Model class diagram (figure 3.3).

Certain classes are removed: in this generation of the simulator the the apoplast is not modelled explicitly in the simulation. It appears as the gap between the cells in the visualisation. Also removed is the explicit Auxin Canal: this is the emergent property that we desire the simulation to exhibit.

Certain classes are added: inheritance is made use of. Inheritance is used to increase the specificity of objects. Proteins and Hormones share some common features, and are modelled as subclasses of the RegElement (regulatory element) class.

Certain relationships are removed: Auxin is no longer related to Cells. This highlights a difference between the Software Model and the biology. In the

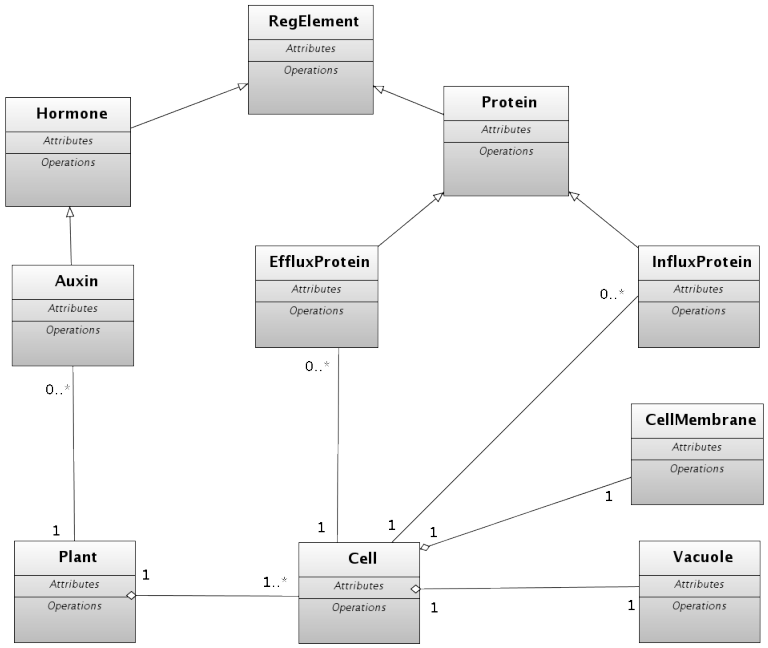


Figure 3.9: The Software Model class diagram, showing only the biological part of the model.

Software Model, the Auxin is related only to the Plant, as are the Cells. For ease of implementation with regard to the diffusion of auxin inside and outside cells, the model records the position of the auxin in the simulation space (part of the Plant, and the Cells can query the Plant to discover how much of that auxin is internal to them. This is a suitable implementation strategy, even though it is not a good model of reality; it shows how the 'same' objects in the Software Model can be quite different in structure from the Domain Model.

The second generation of the simulation does this differently, and explicitly models the apoplast (the space between cells that the auxin is in when not in a cell).

3.3.3 Software Model State diagrams

The Software Model state diagrams follow the Domain Model, except that the production of AUX/LAX is left out, and expressed at a fixed amount and not (currently) regulated.

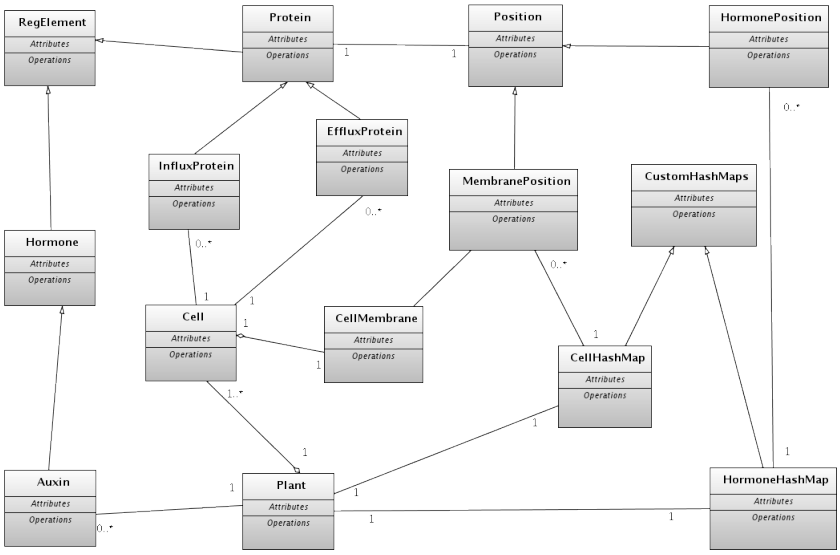


Figure 3.10: A simplified implementation class diagram showing more detail of the underlying implementation of some of the biology in the model. It gives more detail on how the positions of different objects are controlled in the model. The full implementation class diagram shows the classes controlling diffusion and how the threading of the program is controlled. To reduce the complexity of the diagram some classes have been omitted.

3.4 Implementation Model

3.4.1 Implementation Model class diagram

The Implementation Model class diagram (figure 3.10) is a refinement of the Software Model class diagram (figure 3.9). It has all the methods and attributes of the objects (not shown here) and it is use to generate the code skeleton.

The Implementation Model class diagram includes further details of how some of the biological processes are implemented. For example, the positions of components are held by position objects and the hash-map. It may be advantageous to split things in even more detail if the diagrams become overly complex, as certain parts of the implementation are more important than others. Things like the implementation of diffusion and how positions of cells and hormones are stored are of greater interest than how the visual output is achieved.

This separation will be even more worthwhile when things like growth are implemented, as they are likely to be complex and require detailed diagrams. Also, the implementation of such things is much more difficult than conceptualising them, and therefore should be open to more detailed scrutiny to ensure that it is done in a valid way.

The full model has all the methods and attributes of the classes added, and it is this that is used to generate the code skeleton. Once the biology has been produced in both the class diagram and the state diagrams this is often enough to produce a code skeleton from the UML for the model. It might be necessary to define more clearly the interactions between the objects using sequence and collaboration diagrams if the model is large and complex.

3.4.2 Implementation Model state diagrams

The Implementation Model links the Software Model state diagrams of different objects, particularly the overlapping parts of the auxin, PIN and cell state diagrams. This shows how a cell producing auxin influences its own state as it responds to the change in auxin concentration by making more PIN protein. A cell that does not make auxin, but which detects that there is auxin in its cytoplasm, responds by starting to increase production of PIN.

State diagrams are linked by shared events. The Auxin event of entering a cell, either passively or via a protein, is linked to the Cell event of detecting a change in auxin concentration, which causes the cell to enter into a PIN production state. At the moment, we are performing this linking by textual annotations.

This linking of states allows the interactions of the biological objects to be modelled at a higher level than sequence or collaboration diagrams, which are more useful for giving details of how the simulations are going to run. Linking could also be useful to include a notion of space in UML diagrams like state diagrams; for example, states of different objects may influence each other in different ways inside or outside a cell.

3.5 Auxin canalisation Simulator

The code skeletons produced by the UML are then used to develop the simulator. Some of the relevant implementation detail is covered in the following sections.

3.5.1 Molecule diffusion

Two versions of diffusion for the auxin were tested. The method that was eventually used follows an agent based paradigm, modelling (collections of) auxin molecules as moving agents.

It is difficult to estimate the amount of auxin in a cell. The biological data for the amount of auxin in a cell is often determined by crushing up a section of a plant and measuring how much auxin there is in total, then extrapolating that value to individual cells[65]. Although some attempts at using non-invasive methods have also been tried[69]. This assumes an even distribution among the cells, which we believe not to be true. Neither are the cells the same size. This method of estimation gives values for the auxin concentration in one cell that can vary by about two orders of magnitude.

This estimated amount is divided up to give an *auxin unit*, the amount that would occupy 1 unit of space in the model (which is 1 square micron for the 2D model; one cubic micron for the 3D model). From the biological data, this corresponds to 20–2000 auxin molecules in each auxin unit in the simulation.

The auxin units move around the model space in a diffusive manner; they take a random walk around the space. Large numbers of auxin molecules (all those corresponding to an auxin unit) are moving about together as one agent. As agents are allowed to overlap each other in space the only real limit on the number of agents is computational power. It is theoretically possible to make one agent for each molecule but this probably would not be computationally tractable. It is also hard to do from the current data. As we are working at a high level of abstraction throughout the model it is more important to ensure that there are enough agents to avoid the simulation behaviour being adversely affected by low agent numbers.

Some testing was carried out to test if this 'clumping' was a problem, and whether it might be affecting the results of the simulation. A more continuous style model was developed: every unit of space in the simulation has an amount of auxin associated with it, modelled as a number. This could represent the number of individual molecules of auxin, or a fixed number of molecules. This representation allows a portion of the auxin to move into a neighbouring space, and also allows different areas of the simulation space to have different rules, allowing the rate of diffusion to be altered in different parts of the simulation. This is more flexible, but is more costly in computing power and made the simulation code more complicated. Due to being able to scale the number of auxin agents in the first system much more than originally expected, and performance being acceptable this second system was dropped fairly early on in development.

Protein diffusion is implemented in the same way; proteins are coded as objects that describe agents which diffuse in space. Proteins are much larger

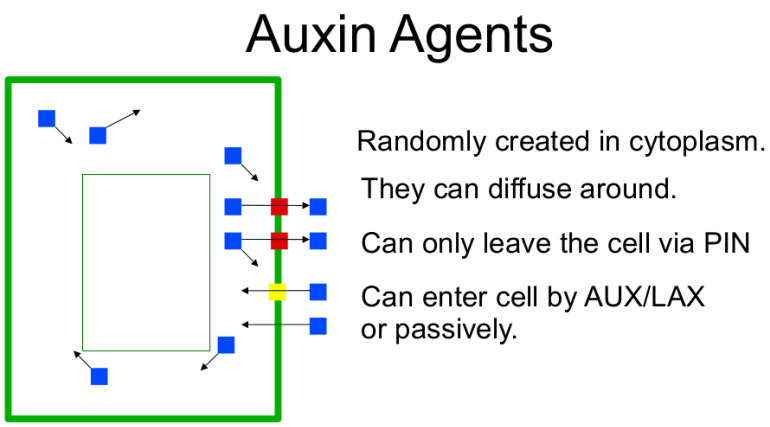


Figure 3.11: Production and movement of auxin agents in the simulations.

than small molecules like auxin, and therefore each protein unit contains fewer protein molecules, and thus the approximation is less problematic.

Agent Behaviours in Space

The behaviours of the auxin agents and the AUX/LAX protein agents do not change in the simulations. The auxin agents are created in the cytoplasm of cells. They are able to diffuse in the cytoplasm. They can only leave the cell by being exported by PIN molecules. Once exported they can diffuse in the apoplast (space between cells). They can degrade anywhere in the space. Figure 3.11 summarises the movement of auxin agents and can be compared with the auxin state diagram, figure 3.5.

The AUX/LAX protein agents are created in the cytoplasm where they are free to diffuse around. If they make contact with a membrane they can become membrane associated. Once on the membrane there is a certain probability that they will drop off the membrane and continue diffusing in the cytoplasm. They are not able to leave the cell. AUX/LAX agents are only degraded when they are diffusing in the cytoplasm. Figure 3.12 summarises the movement of AUX/LAX protein agents.

The space itself normally consists of a grid of cells, normally between five and ten cells in a row. A simulated tissue might have up to fifteen rows. The boundaries on the tissue are hard and not toroidal. Some plant tissues will be very small in size and it is important to understand how the simulations deal with hard boundaries.

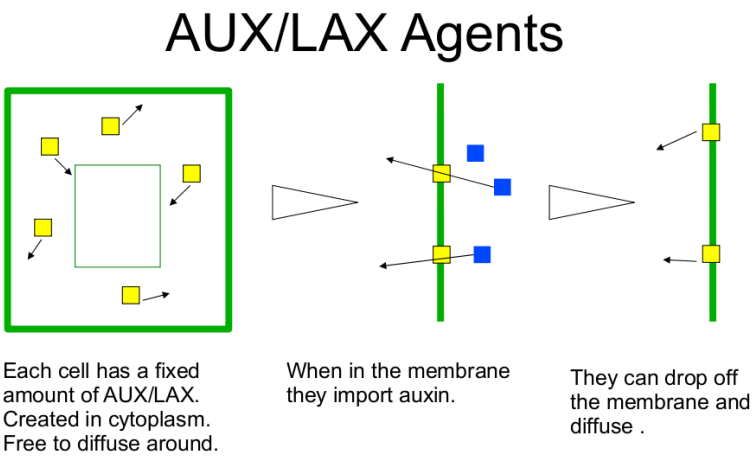


Figure 3.12: Production and movement of AUX/LAX agents in the simulations.

3.5.2 User interface

Biologists interact with the simulator through its graphical user interface (GUI). The GUI provides control over a number of different aspects of the simulator:

- Stopping and Starting
- Saving and viewing of various different still images, showing PIN, auxin, and cell layouts.
- Recording of movie files.
- Altering of all parameters, or the reading of parameters from a text file.
- Displaying output from the model.

Cell layouts cannot be altered using the GUI. Initially a Little Language[8] was developed with the intention of providing an easier interface to setting up the models via an interpreted language. This language was then simplified into more straightforward text configuration files. The parameters.txt configuration file can be used to set starting parameters for things such as the relationship between auxin and the expression of PIN protein. CellFile.txt sets the size of cells and vacuoles, and which cells are to be auxin sources and sinks. It also allows the user to define what the layout of the cells in the model is, by choosing either the number of cells as rows and columns or selecting from a number of more complicated predefined layouts built into the simulator program.

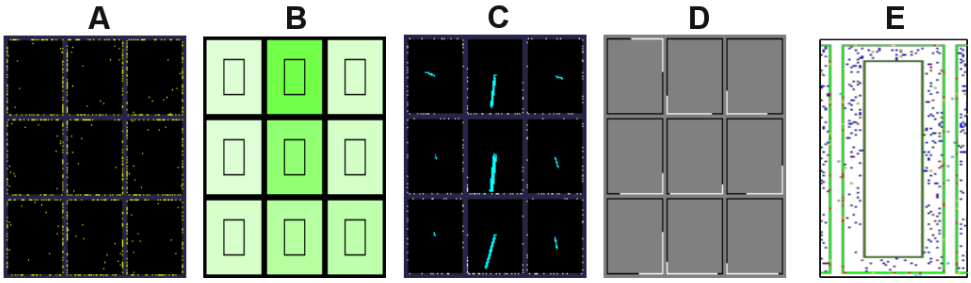


Figure 3.13: Visual output from the simulations. A - shows AUX/LAX proteins in yellow. B - shows auxin concentration in cells, darker green means higher concentration. C - shows polarised PIN. PINs are coloured from grey to white depending on the number of PINs at the particular membrane location. The stick gets wider the more PIN there is in the cell, and it points in the direction of polarisation. The direction is calculated by averaging all the positions of the membrane bound PINs and drawing a line from this point to the centre of the cell. D - another visualisation for PIN polarisation. A sliding window of $\frac{1}{5}$ of the length of the membrane moves round the membrane from a random starting point. The $\frac{1}{5}$ of the membrane with the highest PIN density is then coloured light grey on a black membrane.

The simulator can also be started in "terminal" mode. This uses the configuration defined in CellFile.txt and parameters.txt to run the simulator with no graphical output. This is useful for running on remote computers or clusters. An accessory program can change which cells are sources and which are sinks while the program is running (more detail in appendix B.4).

3.5.3 Visual Output of Simulations

Example visual output from the simulation can be seen in figure 3.13. For the convenience of reproduction only a few rows of cells are shown. Different parts of the simulation output can be visualised separately. When looking only at AUX/LAX proteins they are shown yellow with the cytoplasm and membrane of the cells coloured black and the apoplast dark blue (figure 3.13 A). When looking at auxin concentration in isolation the cytoplasm of a cells is coloured a darker shade of green, with the vacuole and apoplast coloured black (figure 3.13 B).

Visualisation of PIN is more difficult as a method for not only showing PIN, but also its polarisation, is required. Two systems are used in this work. The PINS can be shown on the same background of black cytoplasm and dark blue apoplast as the AUX/LAX. Coloured from grey to white depending on the number of PINs at the particular membrane location. A stick is also added

to the centre of the cell. This stick gets wider the more PIN there is in the cell, and it points in the direction of polarisation (figure 3.13 C). The direction is calculated by averaging all the positions of the membrane bound PINs and drawing a line from this point to the centre of the cell. This method works fairly well as it is visually simple to understand, but it can be fooled. Firstly using this visualisation system a cell with one membrane associated PIN counts as highly polarised, but will have a thin stick indicating a low concentration of PIN. Second, making a stick get fatter and thinner with PIN concentration is a low resolution measure as it is rounded up or down to the nearest whole number of pixels, and there is a limit to the number of pixels that can be used.

The second visualisation system for PIN only provides polarisation information not PIN quantity. A sliding window of $\frac{1}{5}$ of the length of the membrane moves round the membrane from a random starting point. The $\frac{1}{5}$ of the membrane with the highest PIN density is then coloured (figure 3.13 D). These are the automatic visualisation systems for PIN that can be supplemented or altered by hand.

Figure 3.13 E shows a actual cell from a running simulation. The cell membrane is shown as a lighter green box, and the vacuole is the darker inner box. The space between the cell membrane and the vacuole is the cytoplasm where the auxin (blue dots) and proteins (red/yellow dots) are synthesised. The cell structures are on a 1 micron lattice, and the gap between the cells, the apoplast, in the picture is $3um$. In this picture the cells are $60um$ in height and $50um$ across. To avoid questions of sites of synthesis for now there is simply a certain probability that any position in the lattice produces a protein or auxin molecule when required.

3.5.4 2D and 3D simulations

Originally this simulation was developed with the ability to run in 3D in mind. The simulation was designed to use much of the same code to model either 2D or 3D space, with the ultimate goal of having 2D and 3D simulations where the underlying algorithms and code are essentially the same. This is important because if we see significant differences in the behaviour of the different simulations for a given hypothesis or set of parameters we can be fairly sure that difference it due to the extra dimension, as opposed to differences in code.

The aim was to have a 2D simulator for the bulk of hypothesis generation and testing, and only using the 3D model for verification. The early 3D simu-

lation uses two orders of magnitude more memory, and is comparatively very slow. As development went on more of the underlying code in the 2D model became more difficult to easily alter to produce a 3D model. Therefore the 3D version was abandoned to be revisited in the next revision of the simulator.

3.6 Diligent Worker Hypothesis

The most successful of the hypotheses tested has been named for this work as the "Diligent Worker Hypothesis". So named because it describes a mechanism where PINs remain on the cell membrane if they are being productive at exporting auxin into an environment where the auxin is not then building up. The mechanism is described below and shown in figure 3.14. Section 4.1.2 describes the process of finding the parameters that produced the best results. The results from the Diligent Worker Hypothesis PIN polarisation mechanism are presented in section 4.2

The Diligent Worker Hypothesis states that the following mechanism is sufficient to produce auxin transport canals of high flux and high concentration through a simulated plant tissue. The mechanism can be described as follows:

- PINs are randomly created in the cytoplasm.
- The amount of PIN created is linked to the auxin concentration of the cell.
- When in the cytoplasm they are able to diffuse around.
- If they come in contact with the membrane then they may attach (with a certain probability).
- Once attached to the membrane they are able to export auxin out of the cytoplasm and into the apoplast.
- If the concentration of auxin in the apoplast in the space immediately neighbouring the PIN becomes *too high* the PINs drop off the membrane and return to the cytoplasm (an investigation into this behaviour follows in section 4.1.2).
- If the PIN did not export auxin during the last time step then it is able to drop off the membrane and return to the cytoplasm (with a certain probability).

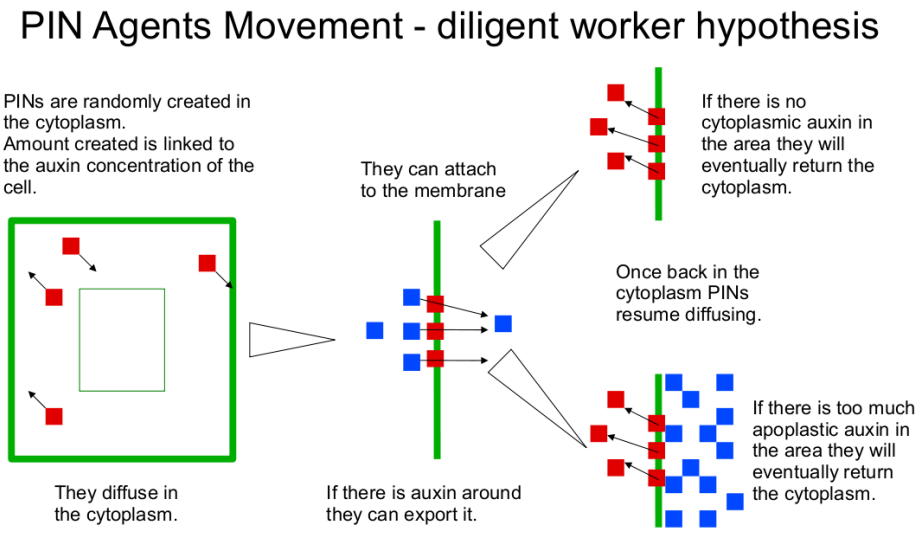


Figure 3.14: Diligent Worker Hypothesis: this mechanism allows PINs to remain in a position on the cell membrane where their export of auxin out of the cell is most productive. PINs will remain in position if there is a supply of auxin which can be exported into an environment where the auxin is not building up.

- PINs degrade only in the cytoplasm.

3.7 Internal Auxin hypothesis: Measuring Internal Auxin Gradients

Kramer reported some success from using internal auxin gradients within a cell to produce auxin transport canals through simulated tissues[52]. As gradients of auxin can been seen in the cytoplasm of cells with highly polarised PINs (see Chapter 4 figure 4.22 panel D for an example of internal auxin gradients) it seemed logical to test if internal gradients could be used to produce auxin transport canals in our models.

A revised form of the Diligent Worker Hypothesis was developed. The only difference is that high internal auxin concentrations, instead of high external auxin concentrations, cause the PINs to drop off the membrane. The hope was that polarisation would occur on membranes due to it being advantageous for PINs to group together on membranes and deplete the auxin in their internal environment. These polarised areas would then hopefully form adjacent to sinks in the tissue as the sinks would reduce the amount of auxin crossing into the cell from the apoplast and disrupting the PINs. The results from the Internal Auxin Gradients PIN polarisation mechanism are presented in section 4.2

The Internal Auxin Hypothesis states that the following mechanism is sufficient to produce auxin transport canals of high flux and high concentration through a simulated plant tissue. The mechanism closely follows that of the Diligent Worker hypothesis and can be described as follows:

- PINs are randomly created in the cytoplasm.
- The amount of PIN created is linked to the auxin concentration of the cell.
- When in the cytoplasm they are able to diffuse around.
- If they come in contact with the membrane then they may attach (with a certain probability).
- Once attached to the membrane they are able to export auxin out of the cytoplasm and into the apoplast.
- If the concentration of auxin in the **cytoplasm** in the space immediately neighbouring the PIN becomes *too high* the PINs drop off the membrane and return to the cytoplasm).

- If the PIN did not export auxin during the last time step then it is able to drop off the membrane and return to the cytoplasm (with a certain probability).
- PINs degrade only in the cytoplasm.

3.8 General Simulation Parameters

Table 3.3 shows the general parameters. These parameters values remain the same for all tested hypotheses, and the values can be compared with other models in the literature. Table 3.4 shows the parameter values for the simulation agents that remain the same in both hypothesis tested. They were found when testing the Diligent Worker Hypothesis using the method described in section 4.1. Parameters specific to the Diligent Worker Hypothesis are discussed in section 4.1.3, and shown in table 4.1. Parameters specific to the Internal Auxin Gradient Hypothesis are discussed in section 4.2, and shown in table 4.2.

Table 3.3 shows parameters general to all the simulations. Auxin diffusion, and degradation rates are broadly in line with the literature[54, 52, 39]. Auxin agents are an abstraction of auxin molecules, and represent more than one auxin molecule. As limits on computational power mean we are unable to model them individually. Auxin production normally only occurs in source cells. Source cells are modelled to maintain their auxin concentration at a set values, consistent with auxin feedback inhibition on auxin biosynthesis[122, 16, 66]. The movement of proteins in the cell cytoplasm is simplified to diffusion at a rate half that of auxin. There is evidence that PIN proteins are trafficked in real cells [49]. However as the process is very complicated for purpose of this work is it better to assume simple diffusion rather than implement a mechanism of trafficking that could be inaccurate and that might affect the results of the simulations in a way that is difficult to understand.

Table 3.4 shows agent parameters that are general to the simulations expressed as probabilities of events happening per time step. Values were attained by a combination of initial results from the GA, and then either checking with the literature or further testing. All cells maintain a fixed concentration of AUX/LAX agents. These are initially created in the cytoplasm where they are able to diffuse around. There are parameters governing their association and dissociation from membranes, and the influx of auxin when attached to membranes. There is no biological data for the association and dissociation of

General Model Parameters			
Parameter Name	Units	Value	Description
time step	s (seconds)	$10e-4$	Model time step
Spacestep	µm	1	Model space step
Auxin Conc.	arbitrary units (a.u.)	0.6	Concentration of auxin agents maintained in source cells.
Auxin Production Rate	$\frac{a.u}{(s)(\mu m)^2}$	5	Auxin agents production rate per second
Auxin Cyto. Diff. Rate	$\mu m^2/s$	600	Diffusion rate of auxin agents in the cytoplasm.
Auxin Apo. Diff. Rate	$\mu m^2/s$	60	Diffusion rate of auxin agents in the apoplast.
Auxin Degradation	s^-1	1×10^{-10}	Rate of auxin degradation.
Protein Diffusion Rate	$\mu m^2/s$	300	Protein agents diffusion rate.

Table 3.3: General Model Parameters. These parameters remain the same for all hypotheses tested.

General Agent Parameters		
Parameter Name	Value	Description
AUX/LAX Conc. (a.u)	0.03	Concentration of AUX/LAX agents. Figures 4.20, 4.21 & 4.22 show the effect of changing the concentration of AUX/LAX agents on canal formation.
AUX/LAX Diss.	1×10^{-4}	Probability of AUX/LAX agents dropping off the membrane (per time step).
AUX/LAX Insertion	0.5	Probability of AUX/LAX attaching to a membrane on contact (per time step).
PIN Insertion	0.7	Probability of PIN attaching to a membrane on contact per time step.
AUX/LAX Degrade	1×10^{-6}	Probability of AUX/LAX agent degrading per time step when in the cytoplasm (per time step).
Membrane Permeability	0.05	Probability of auxin agents crossing the membrane into a cell passively (per time step).
Auxin Efflux	0.5	Probability of auxin agents being effluxed from cells by PIN agents with which they are in contact (per time step).
Auxin Influx	0.5	Probability of auxin agents being influxed into cells by AUX/LAX agents they are in contact with (per time step).

Table 3.4: General Agent Parameters. These parameters remain the same for all hypotheses tested.

AUX/LAX and therefore parameters were chosen that produces a roughly uniform random distribution of around the membrane. As data for the efflux and influx rates of auxin by the proteins is difficult to gather, and the simulation is at a significant level of abstraction, values that allow the proteins to behave as effective auxin pumps were chosen. The auxin transport modelling field has also settled on the idea that efflux by AUX/LAX and PIN proteins can be considered as roughly equivalent, and that it should be approximately one order of magnitude higher than the passive influx of auxin across membranes into cells[39, 52]. Parameters for influx and efflux used in the simulations reflect this.

3.9 Conclusion

The combination of UML and the CoSMoS process has assisted in the development of a flexible simulation framework that simulations biological processes in an intuitive way. Biology maps to UML objects in a straightforward way that can be understood by developers and biologists alike.

The simulations produced are flexible. By concentrating on building the biological components and their interactions into the simulations it is possible to test multiple hypothesis for the regulation of auxin transport. The results should be reflections of truly emergent behaviours, rather than those behaviours being hard coded into the simulation itself.

Using the different level of models that are part of the CoSMoS process to capture domain, simulation, and implementation details has helped produce conceptually cleaner models. The Domain Model looks purely at the biology, and therefore includes emergent properties of interest so they are identified in the model. The Simulation Model does not explicitly include the emergent properties of the System Model: these should emerge from the interactions of the lower level simulated components. The Implementation Model adds the non-biological features to the model, including method for data extraction and visualisation.

The simulations produced by this development process have been used to test the two hypotheses described above and have produced interesting results. The development of this first generation model was also a valuable exercises in using the CoSMoS process with UML. The experiences gained have been taken forward to produce an improved simulator that is described in Chapter 5.

Chapter 4

Results From First Generation Model

A number of different hypotheses for the mechanism of PIN polarisation have been tested using the first generation model. Presented are two hypotheses: the Diligent Worker Hypothesis and the Internal Auxin Hypothesis. The Diligent Worker Hypothesis proved to be very successful; the Internal Auxin Hypothesis is capable of polarising PINs within cells but fails to reliably produce auxin transport canals. Finding biologically plausible mechanisms of PIN polarisation to test in the models is not easy. The hypotheses have to fulfil a number of requirements. Only local information from the agents' environment can be used to determine their behaviour. This is to rule out mechanisms that use long distance information transfer without the implementation of the agents that transfer the information. The behaviours must also be at least biologically plausible. The methods used to parametrise the simulations are introduced in section 4.1. Section 4.1.3 discusses in detail the parameters used to test the mechanism described by the Diligent Worker hypothesis. Section 4.2 discusses the parameters used to test the Internal Auxin hypothesis. The tools used to generate the data from the output of the simulations are discussed in appendix B.2.

4.1 Parameter Search Methods

Parametrisation of models is a difficult problem as data for many of the parameters to allow an agent based model to work do not exist. These simulations are also a significant abstraction of a real cell. This work focuses more on the sensitivity of the simulations to changes in parameters and the relative values

of parameter with each other. In order to explore the parameter space for the agent parameters two techniques were used. Initially the parameter space was explored automatically using a Genetic Algorithm (GA). This was done in two stages, a very rough search to get an idea of values for all the parameters (data not shown). Then a slightly more focused search was carried out on some of the more important parameters. During all phases of parameter testing with the GA manual testing was also carried out. Manual testing involved experiments on a range of different cell layouts with different parameters and carefully observing what was happening. The results from both processes fed-back into each other to produce the final set of parameters used in the experiments.

4.1.1 Genetic Algorithm

Genetic Algorithms are designed to mimic the process of natural evolution in a very simple way. A population of individuals each contain a chromosome that encodes a candidate solution (sometimes referred to as the phenotype) to an optimisable problem. The starting population is normally a number of randomly generated individuals. The phenotype of each individual has its fitness measured against a predefined fitness function. Individuals from the population are then selected based on their fitness and modified (or mutated) and the modified individuals are then used to form a new population. This is the next generation of the population. The population of individuals is evolved towards a solution. The process either terminates when a defined number of generations have been produced, or a satisfactory fitness has been reached. There are many different ways to alter this general scheme. The exact details used for parameterising the two hypotheses are presented later[72].

4.1.2 Diligent Worker Hypothesis Parametrisation

Described in section 3.6 the Diligent Worker Hypothesis proved to be the most successful of the hypotheses tested. Work was carried out to discover parameters for all the agents in the simulation that would allow it to produce canals that form between auxin sources and sinks in a variety of arrangements. A lot of the work done investigating values for the parameters of the simulation was done in the context of how they effect the ability of this hypothesis to form auxin transport canals in tissues. The data shown is the more detailed analysis of the important parameters. Some of the data was collected using the simple GA described in section 4.1.2. This was supplemented with manual testing.

Genetic Algorithm

For testing the Diligent Worker Hypothesis a population of 100 chromosomes was created, each encoding the values of all the parameters to be tested. For an initial very rough search the values for the parameters in the model were randomly generated (data not shown). When looking in more detail at a limited number of parameters the starting parameters (shown in table A.1) were used. Each chromosome was tested by running a simulation using the parameter values it encodes. The success of the parameters – and therefore the entire chromosome – is determined using a fitness function that produces a numerical value indicating how good those parameters were. Then the 10 most successful chromosomes are collected and these 10 chromosomes are multiplied up to make a new starting population of 100. The values encoded in all the chromosomes are then all altered slightly (mutated, the mutation rate used was 0.01 with no crossover). This is the next generation of chromosomes, which are all then tested. This process would normally be repeated for a large number of generations until the fitness stops improving.

The fitness function used in this study is a measure of the ability of the parameters to produce a canal. A number of different automatic methods for calculating this were tried. Many proved unsuccessful as the formation of canals is a self organising process and therefore it is difficult to develop a system that works well for all situations. Frequently the GA found ways to cheat. For example, one method tried was selecting for tissues with a high standard deviation in the auxin concentration of the cells in the tissue. The GA cheated by reducing PIN production and then increasing the amount of auxin the source cell would make. Eventually a simple system was used where the fitness is measured by comparing the average concentration of auxin in the cells that *should* form the canal between the source and the sink. With the average concentration of auxin in all the cells of the tissue, equation 4.1. Where 'C' is the expected canal cells, 'c' is all cells in the space, and 'a' is the concentration of auxin in the cell. The larger the difference between the two average values the better the canal is considered to be.

$$Fitness = \frac{1}{C}\sum_{C=1}^{Canal} a_C - \frac{1}{c}\sum_{c=1}^{All} a_c \tag{4.1}$$

This measure has a few advantages but also some disadvantages. Importantly it is very quick to calculate, which is useful as it is being calculated a large number of times. It is often a reasonable measure of the quality of the

canal. It is hard for the GA to cheat the system into a state with a high fitness according to the measure, but a state that is actually biologically meaningless. The problems with this measure are caused by having to know which cells *should* be in the canal. The severity of this problem was minimised by the tests being carried out with the GA being done on cellular grids with straight files of cells between the source and the sink. These cells are the most likely route that the canal will take. However, as the simulations are self-organising systems sometimes a canal would form slightly out of position and therefore that chromosome of parameters would score a lower fitness than it should. The other problem with testing parameters in complex biological systems is that it is hard to know when to take the measure of fitness. How long do you have to wait for the system to have settled down into something approaching a stable state. Generally simulations were run for twice the number of time steps as it took a known good set of parameters to produce a canal on a cellular grid of the same configuration. The initial set of working parameters were found by manual testing and the very rough GA tests.

Each generation takes about 8 hours to run on a Intel Core 2 Duo 2.4GHz PC. This time combined with the difficulty of developing a good fitness function limited the use of the GA to getting a more general idea of some reasonable parameter values. Once the GA had found some rough parameters that worked well they were then tuned a little by hand. The accessory programs used to carry out this process are described in Appendix B.3. Later a slightly more sophisticated version of the software was developed using a different algorithm (described in [40]). The updated version of the software was only used to test the Internal Auxin Hypothesis.

All parameter testing carried out by the GA was done on a computer cluster with a 5 by 5 grid of cells for 1.5 million time steps. This was a reasonable amount of computer time for a GA, and did not over extended the memory available on the cluster computers. The manual tests to look into particular parameters in more detail were run on 7 by 6 grids of cells for 3 million time steps. Data produced by the GA is presented by plotting the fitness of all individual chromosomes tested on a scatter plot. Appendix B.3 describes the software tools developed to run the simulations on the cluster automatically.

4.1.3 Successful Parameters

The parameters in table 4.1 are specific to producing the behaviour of the Diligent Worker Hypothesis. Detailed results on the testing of the important

Diligent Worker Parameters		
Parameter Name	Value	Description
PIN Production Coefficient (PPC)	15	Controls speed of PIN agent concentration increase with respect to increasing agent auxin (figure 4.10).
PIN Half Concentration (PHC)	0.28 (a.u)	Concentration of auxin agents at which half the possible concentration of PIN agents is attained (figure 4.11).
μ	0.02 (a.u)	Maximum Concentration of PIN agents (figure 4.12).
External Auxin PIN Half Dissociation (EPHD)	9.5	Controls the number of auxin agents in the neighbouring areas that produce a 50% chance of PIN dissociation (figure 4.5).
External Auxin PIN Coefficient (EPC)	20	Controls steepness in probability of PIN dropping off the membrane in response to increasing apoplastic auxin (figure 4.6).
PIN Dis. Const.	1×10^{-10}	Probability of PIN dropping off the membrane if it did not export auxin in the last time step (figure 4.1).
PIN Degrade	1×10^{-4}	Probability of cytoplasmic PIN degrading per time step when in the cytoplasm (figure 4.14).

Table 4.1: Diligent Worker Parameters. These parameters are specific to the Diligent Worker Hypothesis.

parameter values where in sections 4.1.3 to 4.1.3.

Table 4.1 shows the parameters that control the behaviour of the PIN agents in the Diligent Work Hypothesis. Values for these parameters were found using a combination of the GA described above and manual tuning. The effect of these parameters on the behaviour of the hypothesis are described below. These parameters are used for the experimental runs presented in Chapter 4.3.

PIN Dissociation

Key to the functioning of the hypothesis are the parameters governing how the PIN can drop off the membrane of the cells back into the cytoplasm. There are two aspects to this process; dissociation due to not transporting auxin and dissociation due to high concentrations of auxin in the apoplast. Dissociation from the cell membrane due to not transporting auxin is fairly simple and is governed by a fixed probability of dissociation at each time step (PIN Dis. Const.). This is only applied if PIN did not transport any auxin agents in the previous time step. The effect of changing the probability of dissociation can be seen in figure 4.1. Shown is the fitnesses over three independent runs achieved for different values of PIN dissociation. Changing the probability affects the development of auxin transport canals. Figures 4.2, 4.3 & 4.4 show canals produced with different dissociation probabilities. Canalisation improves as probability decreases, with a value of approximately 1×10^{-10} producing the best results.

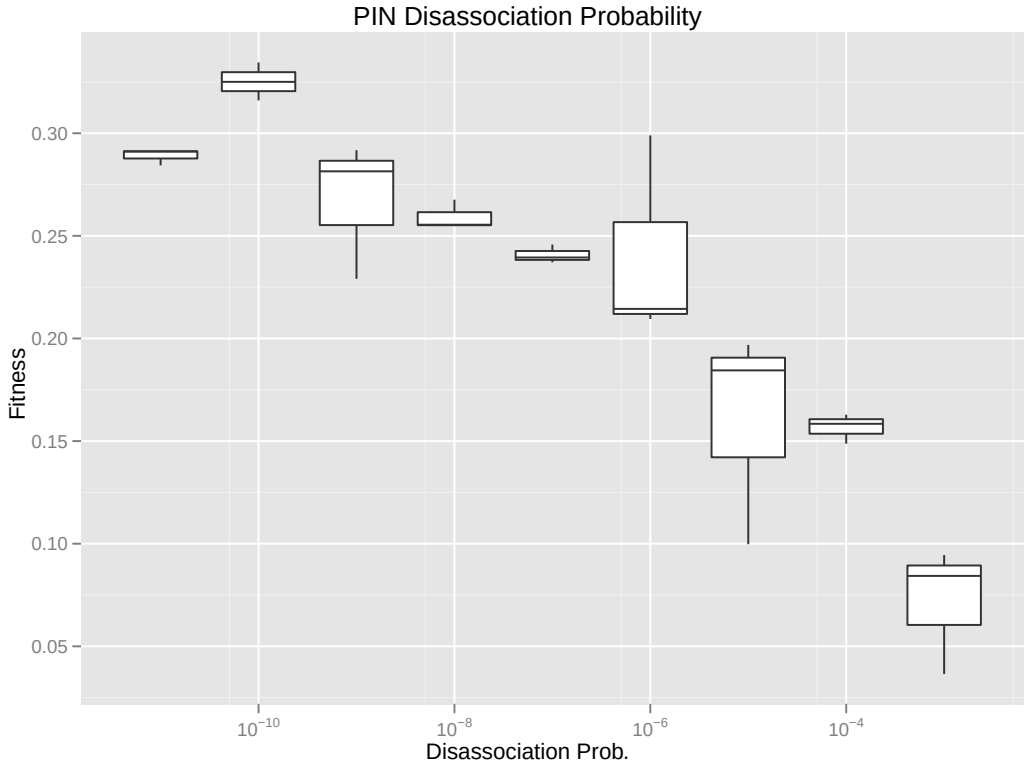


Figure 4.1: The effect on fitness of changing the probability of PIN dissociating from the membrane if it did not transport auxin during the last time step. Shown is a box and whisker plot of the fitnesses over three independent runs achieved for different values of PIN dissociation. For high probabilities the fitness decreases as the PINs are dropping off too easily. This prevents them from becoming polarly localised on the membrane. Too low and PINs are slow to respond to changes in the environment, and PINs remain on the membranes of cells even as the auxin concentration within the cell starts to fall.

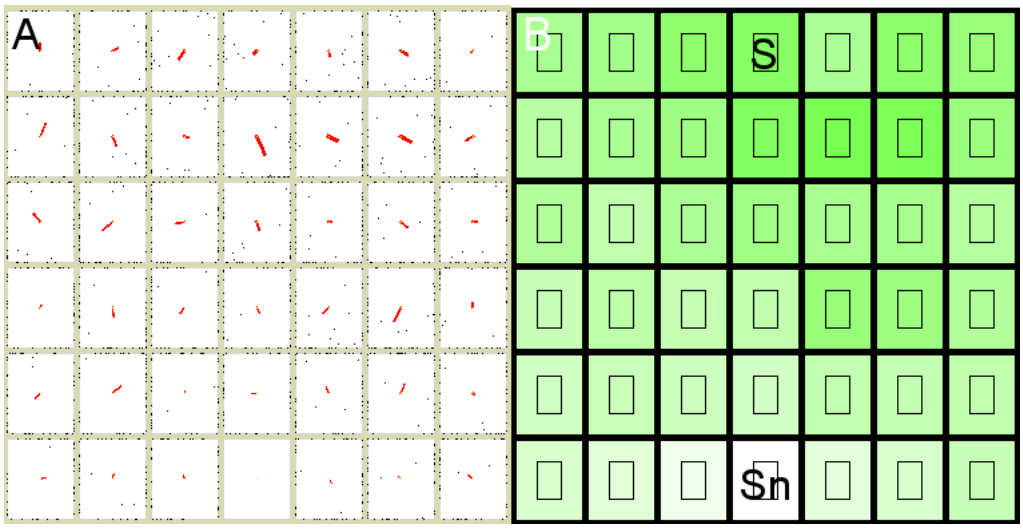


Figure 4.2: The effect on fitness of changing the probability of PIN dissociating from the membrane if it did not transport auxin during the last time step. With a probability of 10^{-3} the PINs do not become polarly localised in the cells. Panel A shows most cells are either not polarised or only weakly polarised. Panel B shows the concentration of auxin in the simulated tissue, the distribution of auxin in the cells shows that no canal is present in the tissue. 'S' indicates the auxin source cell. 'Sn' indicates the auxin sink cell.

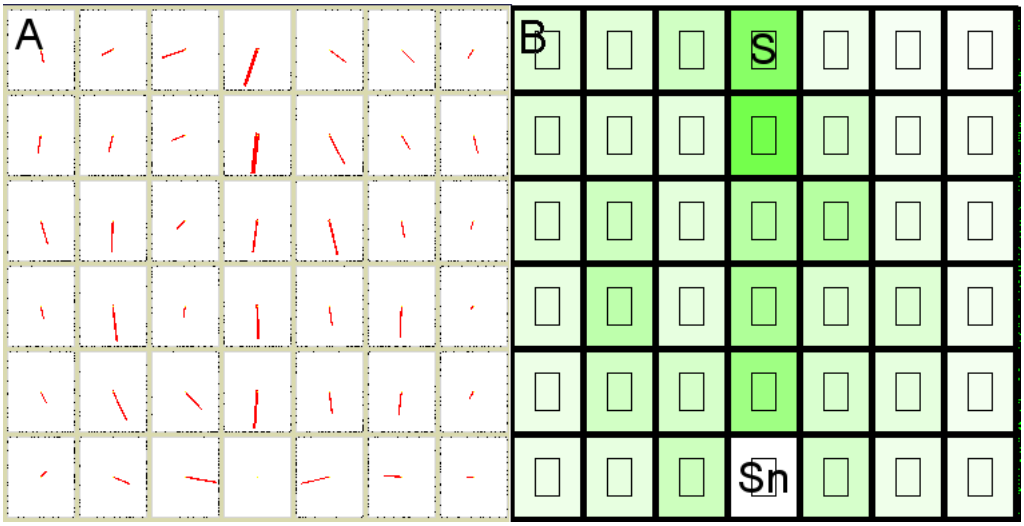


Figure 4.3: The effect on fitness of changing the probability of PIN dissociating from the membrane if it did not transport auxin during the last time step. With a probability of 10^{-7} the PINs do become polarly localised in the cells. Panel A shows cells have become polarised in the direction of the sink. Panel B shows the concentration of auxin in the simulated tissue. The distribution of auxin in the cells clearly shows a canal is forming between the source and the sink. 'S' indicates the auxin source cell. 'Sn' indicates the auxin sink cell.

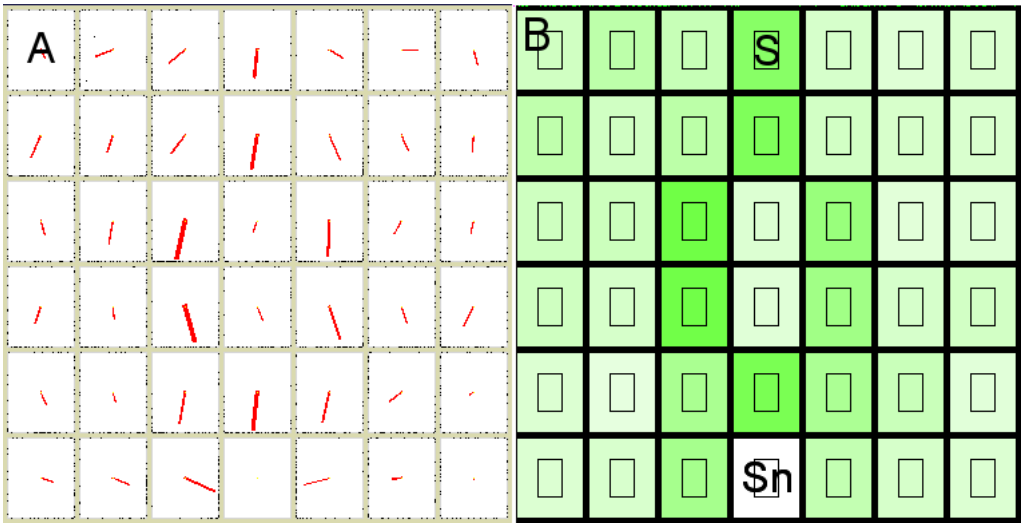


Figure 4.4: The effect on fitness of changing the probability of PIN dissociating from the membrane if it did not transport auxin during the last time step. With a probability of 10^{-11} the PINs do become polarly localised in the cells. Panel A shows cells have become polarised in the direction of the sink. Panel B shows the concentration of auxin in the simulated tissue. The distribution of auxin in the cells clearly shows a canal is forming between the source and the sink. Panel A demonstrates that for very low PIN dissociation probabilities the canals require longer to stabilise as cells remain polarised even if they have low auxin concentrations, and are not carrying out significant auxin transport. 'S' indicates the auxin source cell. 'Sn' indicates the auxin sink cell.

The second aspect of the dissociation process is how the PINs respond to high concentrations of auxin in their local apoplast environment. Although the PIN proteins export auxin out of the cytoplasm into the apoplastic space, their dissociation from the membrane is also partly dependant on the apoplastic auxin concentration. As the concentration of auxin in the space directly neighbouring the position in which the PIN is localised increases, the probability of the PIN dropping off the membrane also increases. In the simulations there are a number of parameters which govern this relationship. A Hill function (equation 4.2) is used that varies the probability of dissociation between 0 and 1. There are two parameters for the function: External Auxin PIN Coefficient (EPC) controls steepness of the response to increasing auxin and External Auxin PIN Half Dissociation (EPHD) controls the number of auxin agents in the neighbouring areas that produce a 50% chance of PIN dissociation. The GA was used to find values roughly suitable for both parameters and these values were then checked by hand.

Figure 4.5 shows the data from the GA for changing values of EPHD and the resulting effect on fitness. The data is fairly noisy, with a fairly wide range of fitness values for the same parameter values being reported. A value of 9.5 was selected for use in further testing of canalisation behaviour, shown in Chapter 4.3. Figure 4.6 shows the data from the GA for changing values of EPC and the resulting effect on fitness. The fitness of the canals is not sensitive to this parameter once the value gets higher than 18. A value of 20 was selected for the canalisation experiments shown in Chapter 4.3. Figure 4.7 shows the behaviour of equation 4.2 for increasing concentrations of auxin using a value of 20 for the EPC parameter and 9.5 for EPHD.

$$(D.Prob.) = \frac{[a]^{EPC}}{(EPHD)^{EPC} + [a]^{EPC}} \tag{4.2}$$

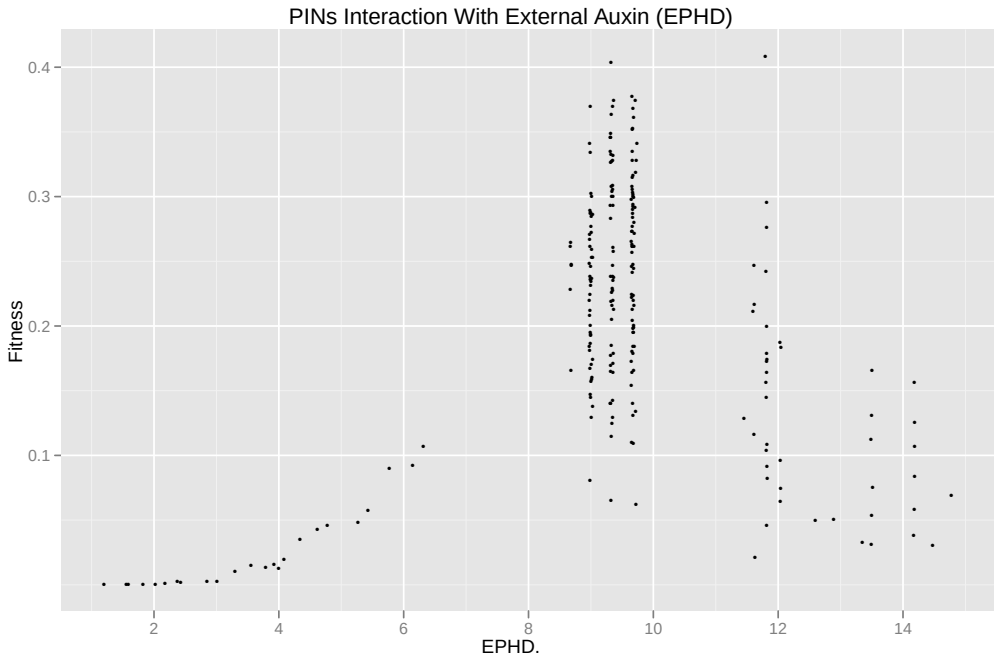


Figure 4.5: Scatter plot showing the fitness of individual runs testing different values of EPHD. EPHD is a parameter in the Hill equation that controls the interaction of PIN with the auxin in the Apoplast: This parameter is the number of apoplast auxin agents in the neighbourhood of a membrane bound PIN agent that gives it a 50% probability of dropping off the membrane back into the cytoplasm. A limited rerun of the GA on a reduced part of the space was carried out to investigate this parameter in more detail. Even with the stochastic nature of the simulations fooling the fitness function a value of approximately 9.5 produces consistently good values of fitness.

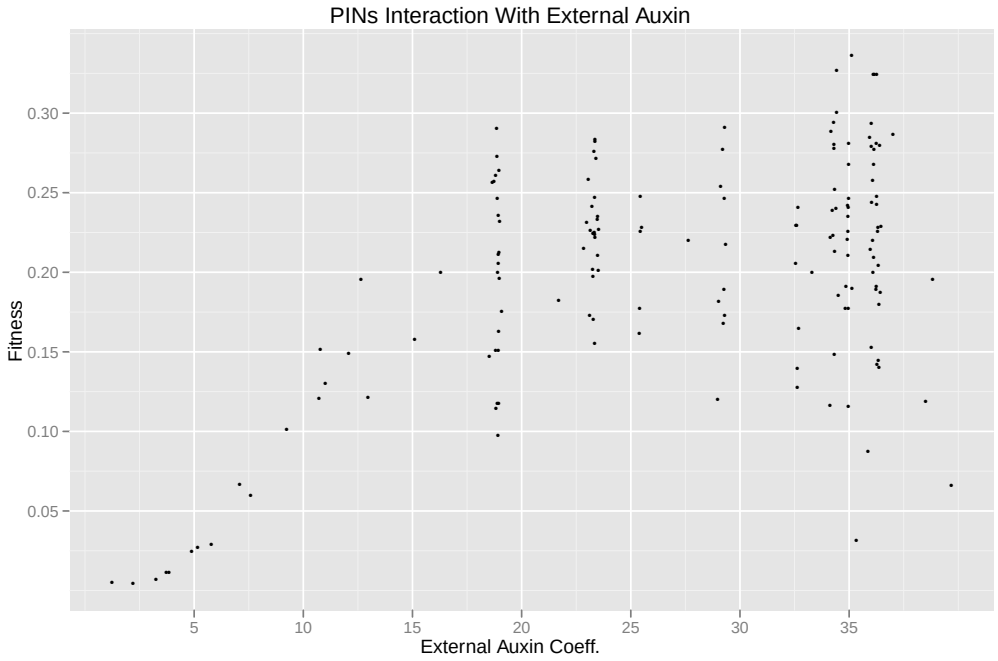


Figure 4.6: Scatter plot showing the fitness of individual runs testing different values of EPC. EPC is a parameter in the Hill equation that controls the interaction of PIN with the auxin of the Apoplast: This parameter alters the steepness with which the probability of PIN dropping off the membrane increases towards 1 with increasing apoplastic auxin. A limited rerun of the GA on this parameter shows that the simulations are not particularly sensitive to the value of this parameter. However steeper increases in probability produced by values higher than approximately 20 seem to produce the best results.

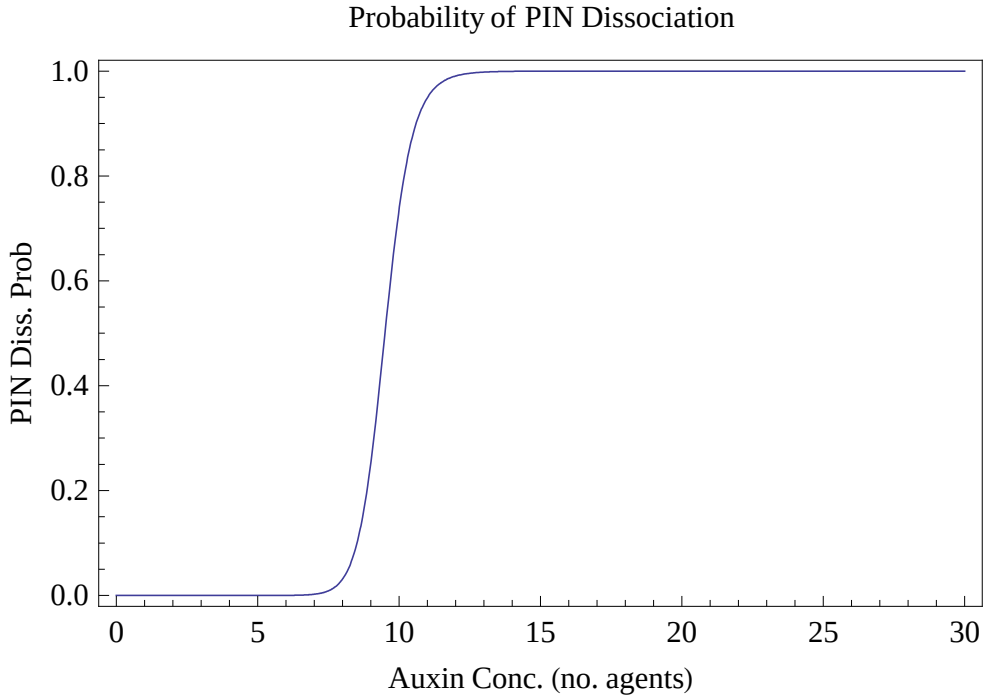


Figure 4.7: Plot showing the probability of PIN dissociating from the membrane due to apoplastic auxin concentration (shown as number of agents in the neighbour space). This is under the control of equation 4.2 using the EPHD parameter at a value of 9.5 and EPC parameter at a value of 20.

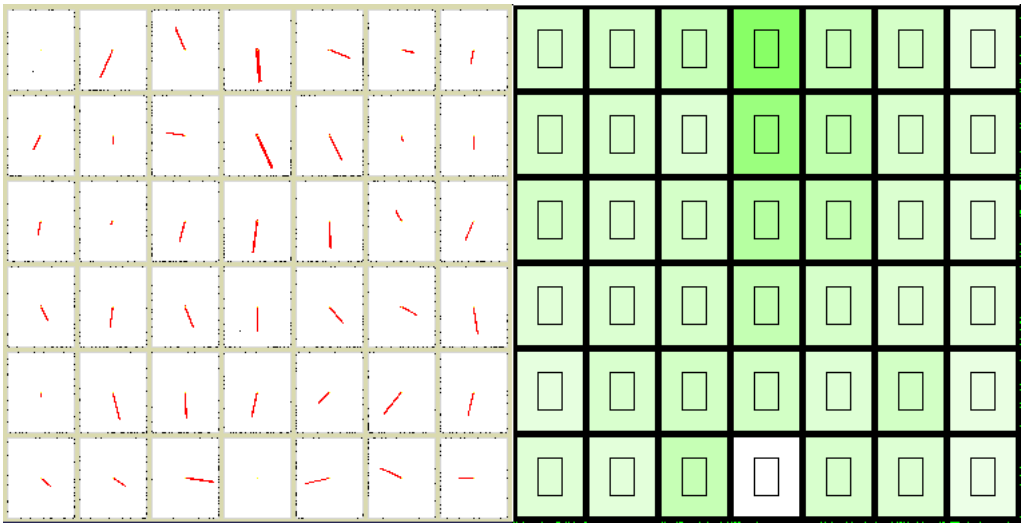


Figure 4.8: Simulation run with only half of the normal regulation of PIN dissociation normally in the Diligent Worker Hypothesis present. The PIN molecules will only drop off the membrane if they are exporting auxin into a local environment of high auxin concentration. Auxin transport canals are still able to form under these conditions, however canals formed under the full hypothesis accumulate more auxin(data not shown).

Figure 4.8 shows that using only the influence of the apoplastic auxin concentration alone on PIN localisation can produce auxin transport canals. Qualitatively these canals do not accumulate auxin as well as those developed by the full hypothesis. Figure 4.9 shows that using only the dissociation of PIN due to not transporting auxin to influence the localisation of PIN does not produce auxin transport canals or polarised cells.

Three other regimes were tested for controlling this relationship. A linear relationship where the probability of PIN dissociation increased in line with increasing apoplastic auxin concentration. An exponential relationship where the probability of dissociation increased exponentially with auxin. Finally, a simple threshold where PIN disassociates when auxin concentration in the apoplast rises over a set value. Although all three can be made to work, they were all more sensitive to the parameter values making them more difficult to work with (data not shown).

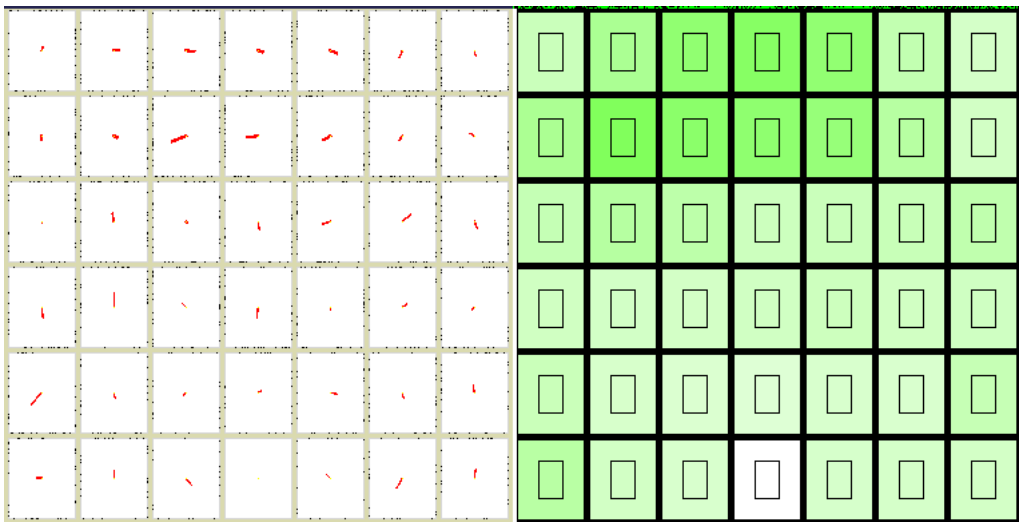


Figure 4.9: Simulation run with only half of the normal regulation of PIN dissociation normally in the Diligent Worker Hypothesis present. The PIN molecules will only drop of the membrane if they are not transporting auxin out of the cell. Auxin transport canals are not formed under these conditions.

PIN Production Parameters

To investigate their effect on canalisation other key parameters were also tested. PIN production does have an effect on the formation of auxin transport canals. PIN concentration increases with auxin concentration regulated by a Hill function (equation 4.3). If cells produce a large amount of PIN for small amount of auxin, the auxin in the cell is rapidly exported out. This causes canals to take longer to accumulate auxin. Figure 4.10 shows the effect of changing the speed that PIN production is increased in line with increasing auxin concentration, PPC. The data is noisy but the highest fitnesses are attained with values from about 5 to 18. A value of 15 was chosen. The second parameter in the Hill function controls the amount of auxin in the cell that will produce half the possible amount of PIN, PHC (shown in figure 4.11). The simulations do not seem to be particularly sensitive to this parameter, and good canals form with values that are slightly lower than half that of the auxin concentration maintained in the source cells. This is logical as it means that cells that accumulate auxin are likely to produce the maximum amount of PIN allowable. It also stops cells with very low auxin concentration maintaining high PIN concentrations. As a value of 0.6 (a.u.) is used for the concentration of auxin maintained in the sources cells, a value of 0.28 is used for the PHC parameter. The final part of the regulation of PIN production is the maximum concentration PIN can reach in cells, μ. Figure 4.12 shows the effect of changing μ on the development of canals. Figure 4.13 shows how the PIN concentration increases with auxin concentration.

$$[P] = \mu \frac{[a]^{PPC}}{(PHC)^{PPC} + [a]^{PPC}} \tag{4.3}$$

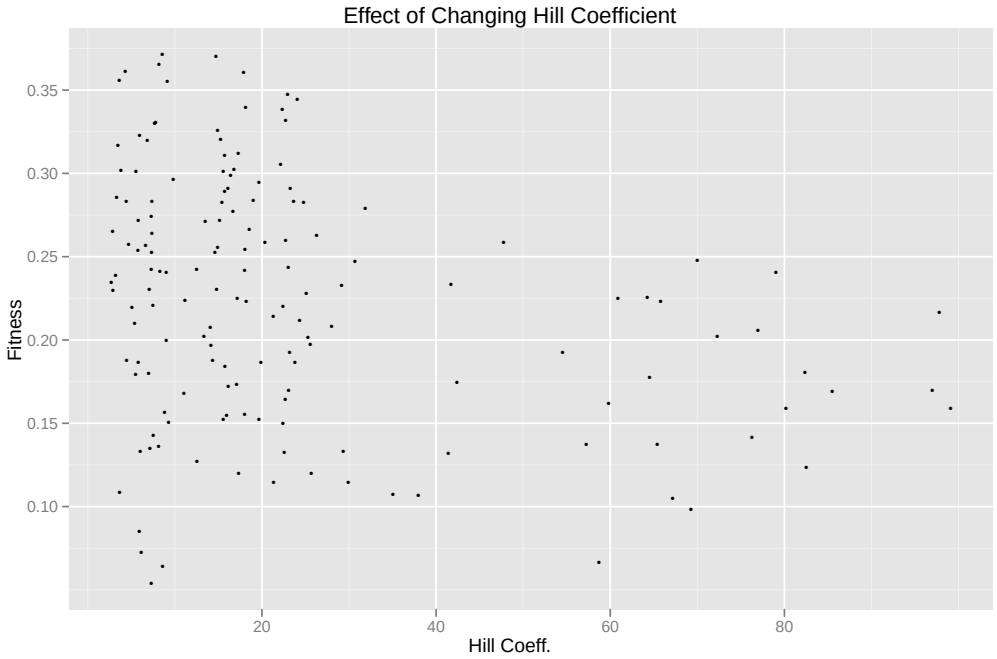


Figure 4.10: Scatter plot showing the fitness of individual runs testing different values of PPC. PPC is a parameter in the Hill equation that controls the production of PIN in response to increasing auxin concentration. This parameter alters the steepness with which PIN concentration increases with auxin concentration. The results from the GA run are fairly noisy. However, these and some hand tuning have lead to a value of 15 being chosen for this parameter.

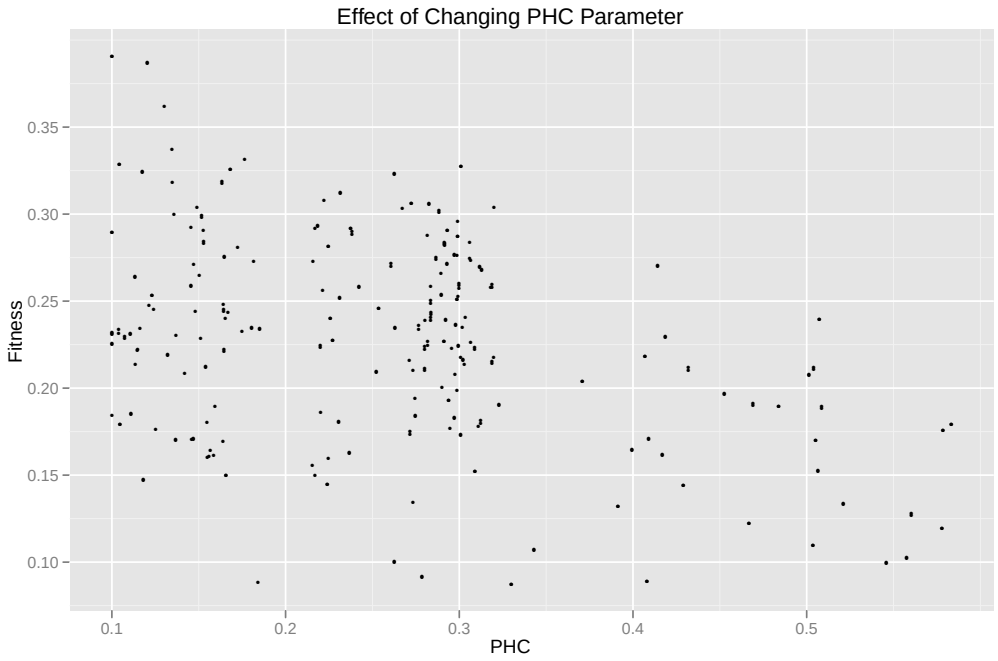


Figure 4.11: Scatter plot showing the fitness of individual runs testing different values of PHC. PHC is a parameter in the Hill equation that controls the production of PIN in response to increasing auxin concentration: This parameter alters the concentration of auxin for which half the total possible concentration of PIN is produced. Model sensitivity to this parameter is fairly low. Qualitatively the results suggest that a value slightly lower than half the concentration of the auxin maintained in the source cells is good enough. This is logical as it would allow cells of high auxin concentration that are a canal to produce the maximum possible amount of PIN protein. A value of 0.28 has been chosen.

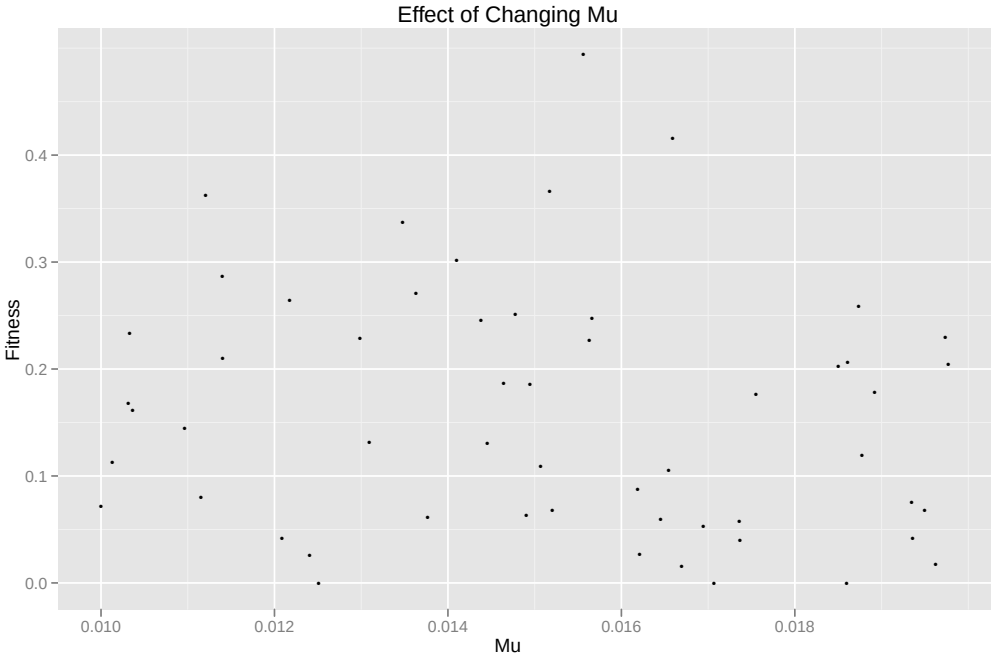


Figure 4.12: Scatter plot showing the fitness of individual runs testing different values of μ. μ is the parameter that determines the highest possible concentration of PIN in the cells. A value of approximately 0.02 has been chosen.

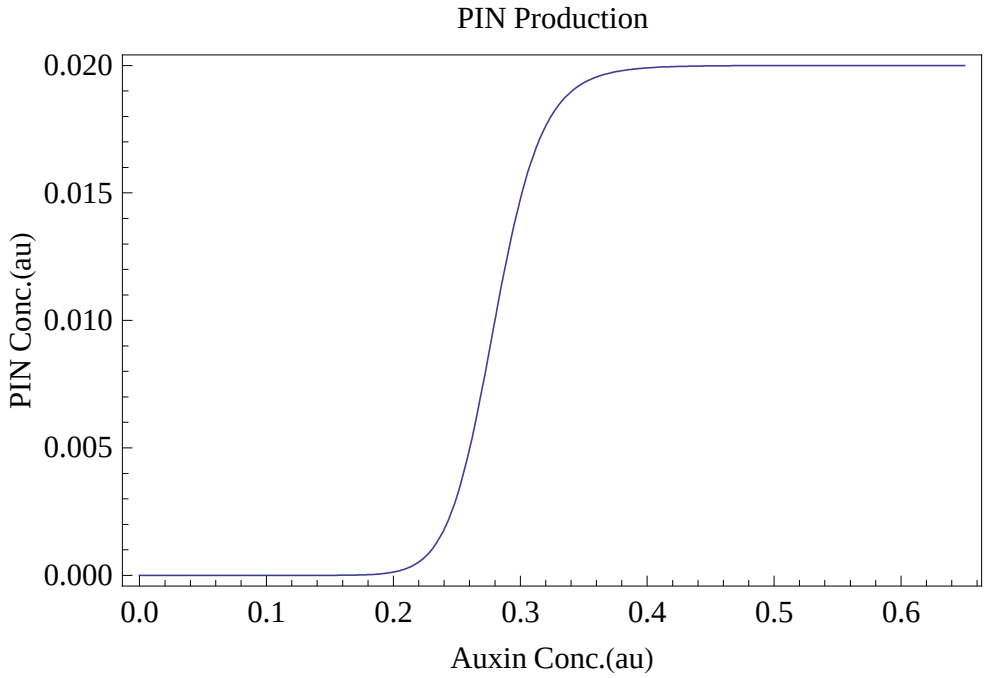


Figure 4.13: Plot showing the changing concentration of PIN protein between 0 and the maximum value of 0.02 (determined by the μ parameter) as the concentration of auxin increases (under the control of equation 4.3). The steepness of the curve is controlled by the PPC (set at a value of 15), the PHC (set at a value of 0.28) parameter alters the concentration of auxin for which half the total possibly concentration of PIN is produced. PIN concentration is ramped up fairly quickly once the cell starts to accumulate auxin.

PIN Degradation Parameters

PIN degradation has also been tested. In a real plant cell the regulation of PIN levels in the cell is very complicated[48]. In order to simplify it and avoid making assumptions that could strongly effect the behaviour of the simulations PIN is degraded from the cytoplasm only, and at a fixed rate. Figure 4.14 shows the effect of PIN degradation on the formation of auxin transport canals. The simulations are not very sensitive to this parameter (with the exception of very high probabilities), canals form over time as long as it is not set at values lower than 10^{-5} or higher than 10^{-4}. If the probability of degradation is set at a very high value like 10^{-2}, PINs are removed from the cytoplasm and then replaced by production too frequently. This interferes with the working of the hypothesis as PIN is essentially removed and then placed in a new random location, slowing down polarisation. Figure 4.15 shows the effect of a high PIN degradation rate: canals can form but seem to take longer to stabilise. Figure 4.16 shows the effect of a very slow degradation rate. Due to its slow degradation PIN remains in the cells even when their auxin concentration has dropped. This produces wider canals, which is linked to the cell PIN concentration being slower to react to the environment. Figure 4.17 shows an intermediate PIN degradation probability which seems to produce the best results.

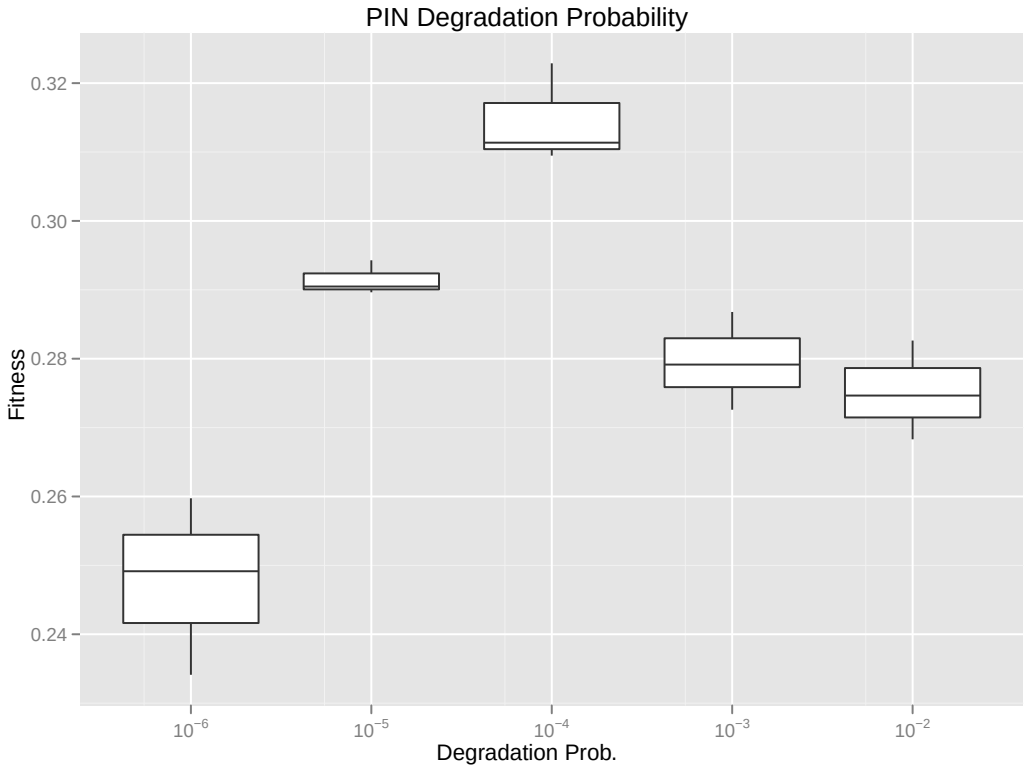


Figure 4.14: Box and whisker plot of the effect on fitness of changing the probability of PIN degrading in the cytoplasm of the cell. If the degradation rate is too high PINs do not have long enough in the cells to move around and come into contact with an area of cell membrane that if they associated with it would put them in a favourable position with respect to the hypothesis. Turning PINs over too quickly in this way increases the time for a good canal to form as PINs are essentially interrupted in their progress to a *good* membrane location and randomly put somewhere else. If the degradation rate is too low then as the auxin concentration of a cell decreases and the excess PIN seems to have a detrimental effect on canals by allowing cells to maintain a concentration of PIN that is higher than the auxin concentration in the cytoplasm would produce. Data from three independent runs.

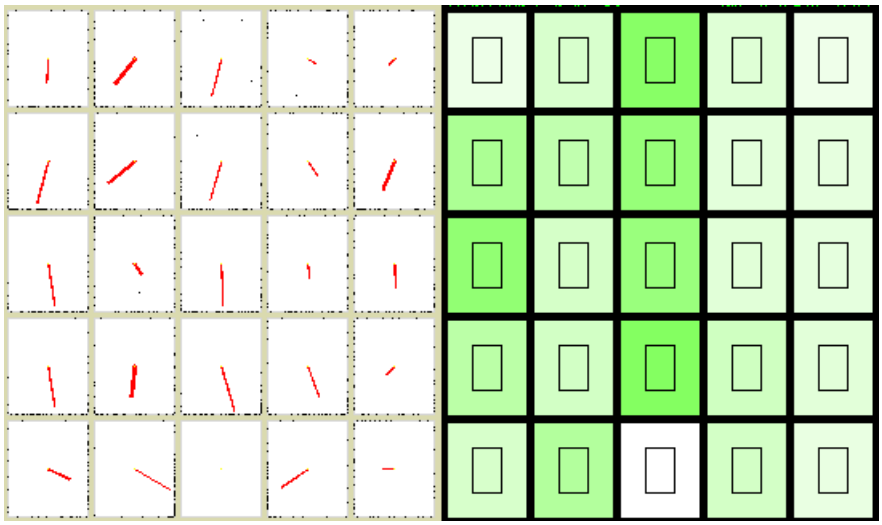


Figure 4.15: With a probability of PIN degradation of 10^{-2} canals are still able to form but qualitatively not as good as those formed with a probability of 10^{-4} (figure 4.17).

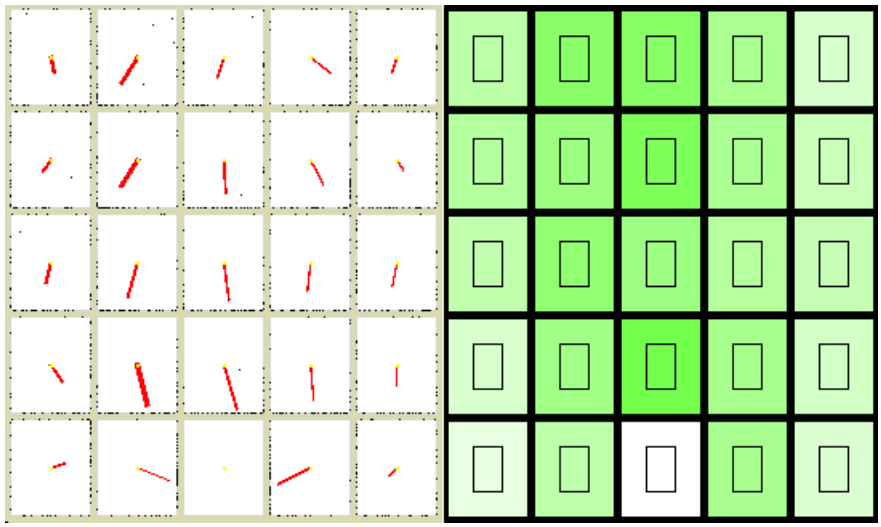


Figure 4.16: Probability of PIN degradation of 10^{-6} canals form between the auxin source and the sink but they seem to be wider. This could be due to the PIN concentration dropping slowly even if the cell has lost its auxin. Therefore cells remain as good sinks for longer.

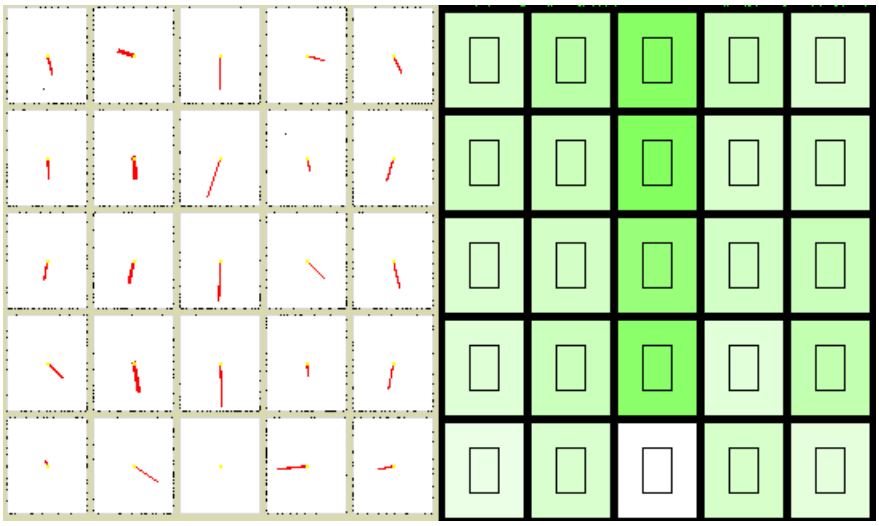


Figure 4.17: With a probability of PIN degradation of 10^{-4} narrow canals form between the auxin source and the sink.

PIN Concentration Over Time

The behaviour of PIN under the control of the parameters shown in table 4.1 can be seen in figures 4.18 and 4.19. Figure 4.18 shows PIN concentration in a canal cell over time. After some initial variation as the canal develops the auxin concentration stabilises and remains high. The cell therefore produces the maximum amount of PIN and keeps it throughout the experiment. Figure 4.19 shows the concentration of PIN in a cell at the edge of the tissue and not part of a canal. Initially the auxin concentration in this cell increases as the auxin produced at the source is transported into all the cells in the tissue. During this period the cell's PIN concentration increases. The auxin – and therefore the PIN concentration – both vary while the canal is becoming established elsewhere in the tissue. Once the canal has been established away from this cell most of the auxin produced by the source is transported through the canal to the sink and does not reach this cell. Therefore the auxin concentration in this cell falls, and the PIN concentration falls with it. Time series data was generated using the program described in Appendix B.6.

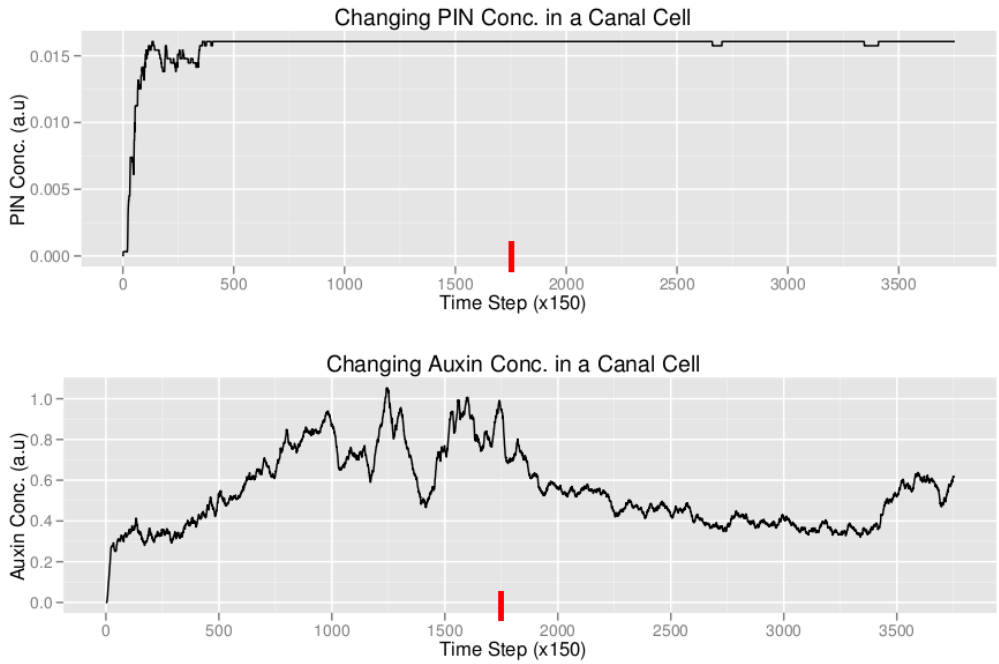


Figure 4.18: Time series data showing the concentration of PIN in a cell that is part of a canal. The top plot shows that very quickly the PIN concentration reaches the maximum allowed by the parameters and remains there. The lower plot shows auxin concentration at same time. Auxin concentration goes through a period of instability at the beginning of the simulation run but then settles down. Data from a single simulation run. Mark on x-axis indicates time of canal formation.

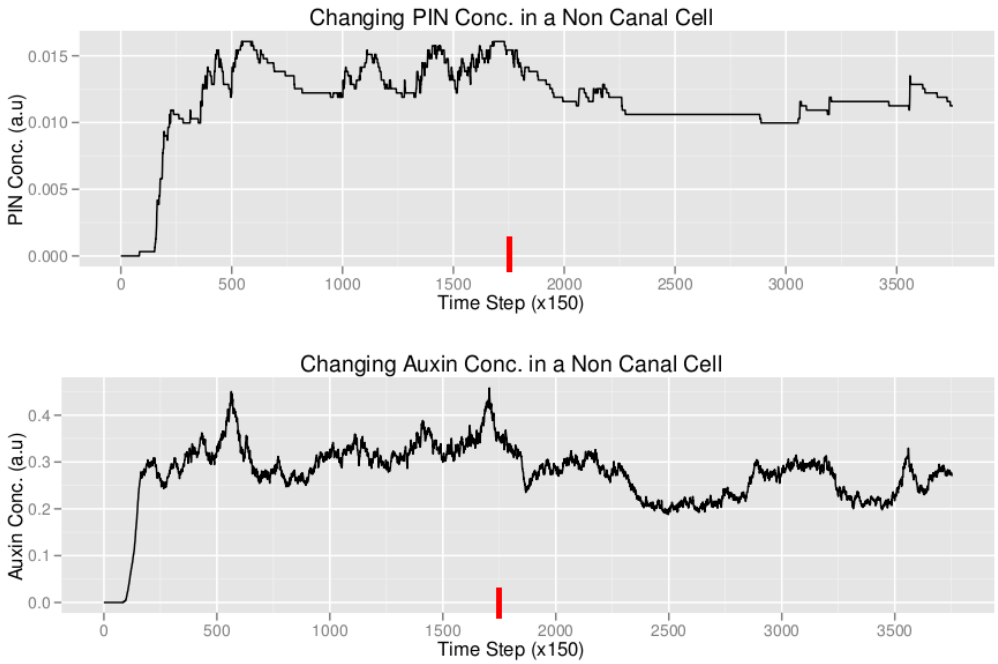


Figure 4.19: Time series data showing the concentration of PIN in a cell that is on the edge of a simulated tissue and not part of a canal. The top plot shows how the PIN concentration goes up and down as the canal is developing in the tissue and then starts to fall once the canal is formed and the cell gradually loses more auxin. The lower plot shows the auxin concentration over the same time. Data from a single simulation run. Mark on x-axis indicates time of canal formation.

The Effect of AUX/LAX Concentration

The amount of AUX/LAX influx protein has a strong effect on the ability of auxin transport canals. When in the simulations AUX/LAX is created in the cytoplasm where it is able to diffuse around the cell. It can associate and dissociate from the membrane of the cells. The dissociation probability of AUX/LAX for each time step was set at 1×10^{-4} so that it will spend a significant amount of time attached to membranes, but diffusion will allow it to become randomly positioned around the entire membrane. Figure 4.20 shows the effect of altering the concentration of PIN on the development of canals (three independent runs). No AUX/LAX protein and the PIN in the cells does polarised towards the sink (qualitatively not as strongly as when there is AUX/LAX). However a canal of high auxin concentration does not form between the source 'S' and the sink 'Sn' (figure 4.21). Figure 4.21 A shows that there is no AUX/LAX in the system. Figure 4.21 B shows that an auxin transport canal of high auxin concentration as not formed between the source and the sink. Figure 4.21 C shows that PIN has been able to polarise towards the sink, however the polarisation is not as strong as when AUX/LAX is present.

When there is significant amount of AUX/LAX present, auxin transport canals of high auxin concentration form between the source and the sink (figure 4.22). Figure 4.22 A shows large amount of AUX/LAX agents (light dots) present in the simulation. Figure 4.22 B clearly shows an auxin transport canal between the source 'S' and the sink 'Sn'. Figure 4.22 C shows that the PIN agents have become polarised towards the sink.

Validation of simulations is difficult. There are a number of important pieces of information that are currently unknown. Experimental estimates exist for how much auxin is in a cells of auxin transport canals, or the speed at which auxin is imported and exported from cells (either directly across membranes or by proteins) [53]. These values can be used as starting points for simulations parameters but the models need to be robust to varying them. The time for PIN protein to become polarly localised and the auxin transport canal to develop can also be compared between different models. Most models report polarisation of a tissue in minutes to hours, rather than seconds [39, 117]. Time for polarisation of tissues in the simulations presented here are broadly in line with these time. For example, the diagonal canal shown in figure 4.26 was established after 100,000 timesteps or approximately 10 minutes.

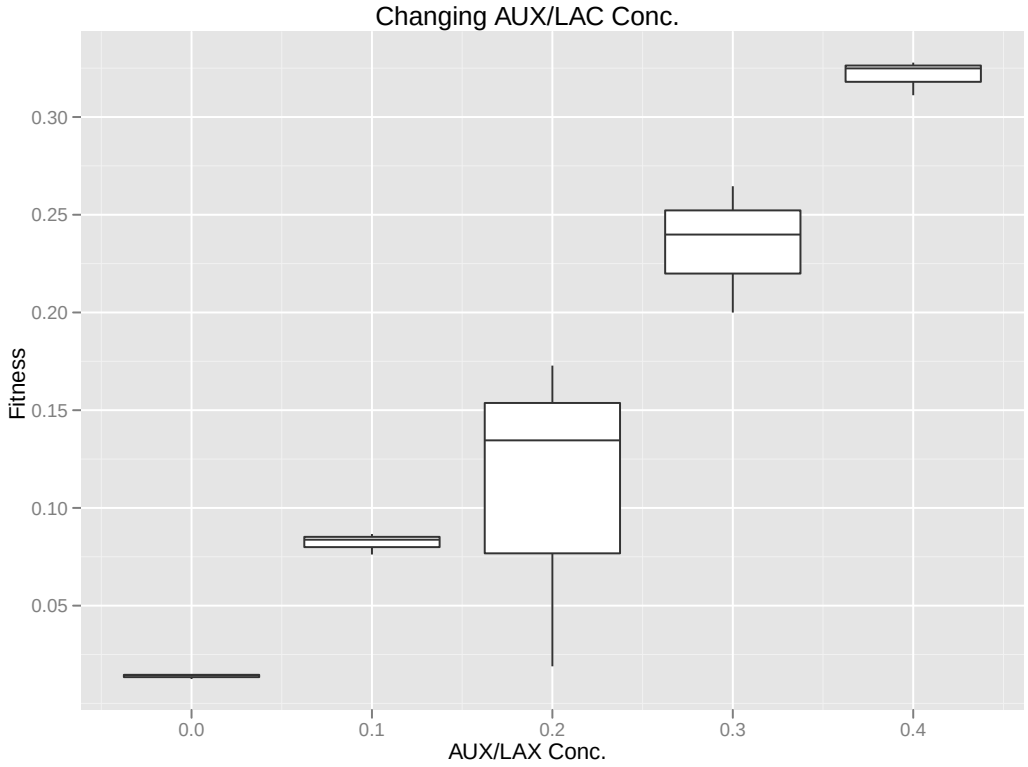


Figure 4.20: Plot showing the effect of increasing AUX/LAX concentration on the fitness. Shown is a box and whisker plot of three independent runs per AUX/LAX concentration value tested. Fitness increases as the amount of AUX/LAX increases. As the fitness measure is based on auxin accumulation of auxin, cells that are able to suck auxin out of their environment efficiently, and are in the auxin transport stream, accumulate large concentrations of auxin. Higher concentrations of AUX/LAX where not tested as much higher would be pushing the limit of what is biologically plausible.

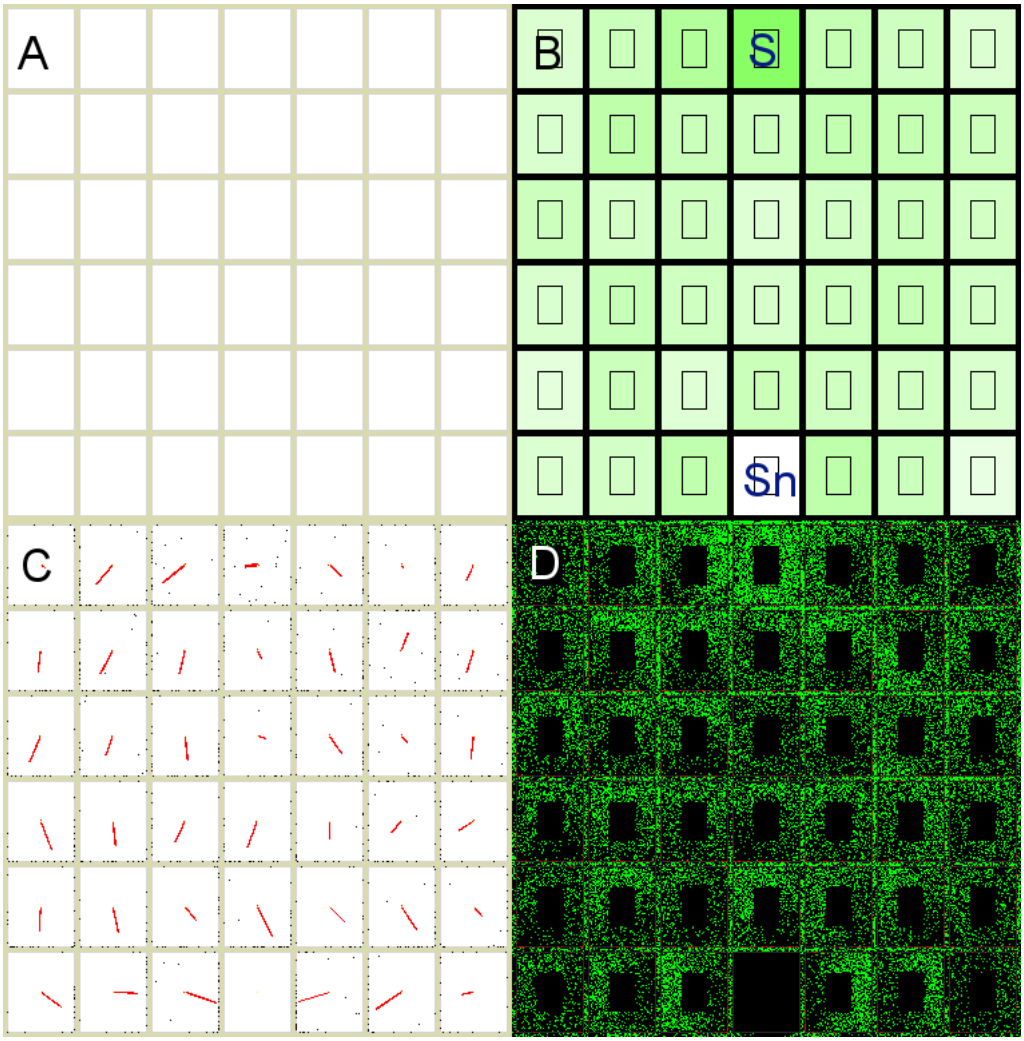


Figure 4.21: Changing AUX/LAX concentration. Panel A shows that there is no AUX/LAX in this simulation. Panels B and D show that auxin is not accumulating in an auxin transport canal between the source and the sink. Panel C shows the PIN in the cells is becoming polarised but as the cells are not accumulating auxin the amount of PIN in the cells remains low. Panel B shows the position of the auxin source cell 'S' and the auxin sink cell 'Sn'.

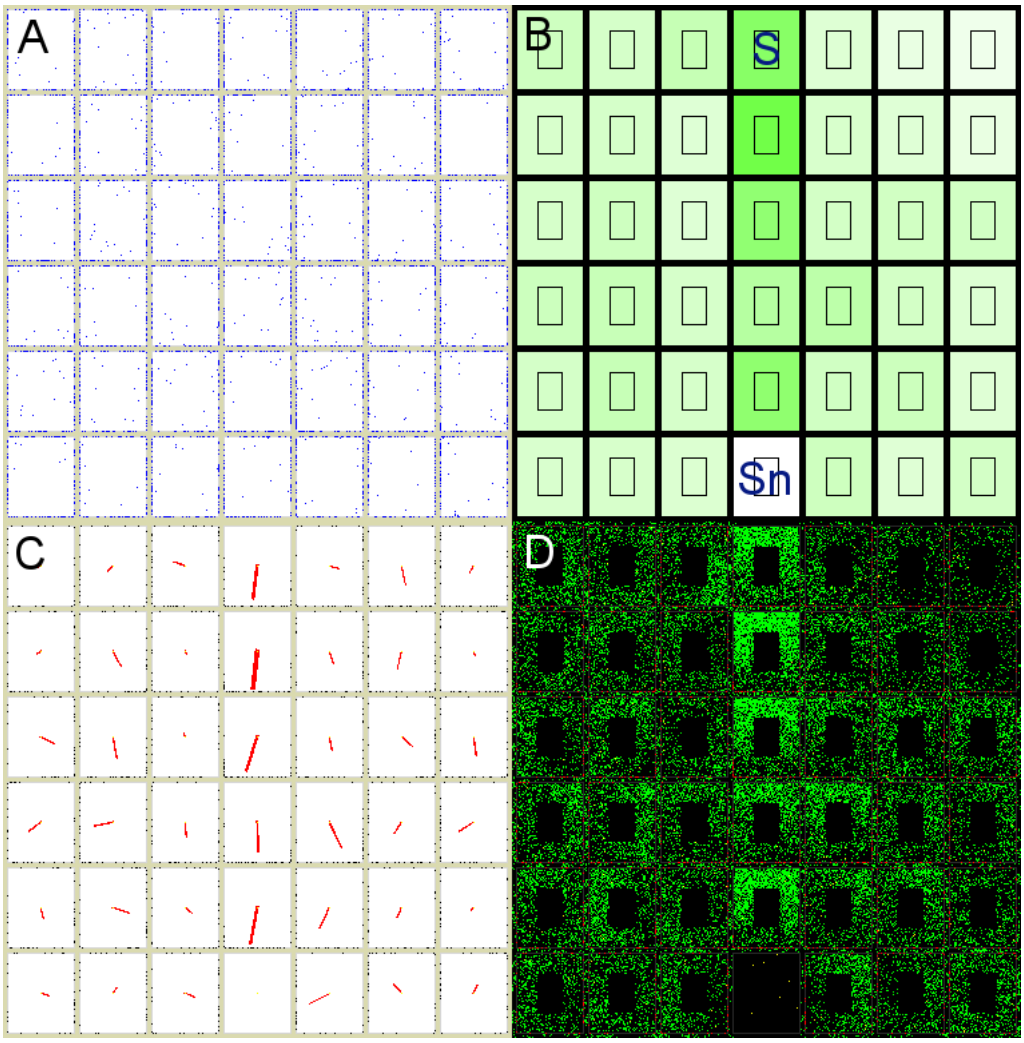


Figure 4.22: Changing AUX/LAX concentration. With an AUX/LAX concentration of 0.04 panel A shows significant amounts of AUX/LAX (blue dots) on the membranes of the cells. Panels B and D show that auxin is accumulating into an auxin transport canal between the source and the sink. Panel C shows the PIN in the cells is becoming polarised and the cells with high auxin concentration have increased PIN. Panel B shows the position of the auxin source cell 'S' and the auxin sink cell 'Sn'.

Internal Auxin Parameters		
Parameter Name	Value	Description
PIN Production Coefficient (PPC)	15	Controls speed of PIN agent concentration increase with respect to increasing agent auxin.
PIN Half Concentration (PHC)	0.28	Concentration of auxin agents at which half the possible concentration of PIN agents is attained.
μ	0.02	Maximum Concentration of PIN agents.
External Auxin PIN Half Dissociation (EPHD)	9.5	Controls the number of auxin agents in the neighbouring areas that produce a 50% chance of PIN dissociation.
External Auxin PIN Coefficient (EPC)	20	Controls steepness in probability of PIN dropping of the membrane in response to increasing cytoplasmic auxin.
PIN Dis. Const.	1×10^{-10}	Probability of PIN dropping off the membrane if it did not export auxin in the last time step.
PIN Degrade	1×10^{-4}	Probability of cytoplasmic PIN degrading per time step when in the cytoplasm.
PIN Insertion	0.7	Probability of PIN attaching to a membrane on contact.

Table 4.2: Most successful Internal Auxin Hypothesis parameters.

4.2 Internal Auxin hypothesis: Measuring Internal Auxin Gradients

Described in section 3.7 the Internal Auxin Gradients Hypothesis is the second system tested that produces polarisation of PIN proteins in cells uses internal auxin gradients. Table 4.2 shows the parameters which produced the best results for this hypothesis.

Under this mechanism we observe strong polarisation of PIN in cells. However, we have be unable to produce good auxin transport canals. The polarised regions do not seem to be able to reliably find the localisation of the sink, or polarised cells that should behave as sinks. This inability to find sinks stops the reliable development of good auxin transport canals through the tissue. There also seem to be some strong edge effects in some of those experiments that did manage to produce canals.

Figure 4.23 shows that PIN proteins in the cells clearly become polarised on the membranes. In these examples canalisation has also occurred. However, this is not reliable and seems to frequently involve canals following the edge

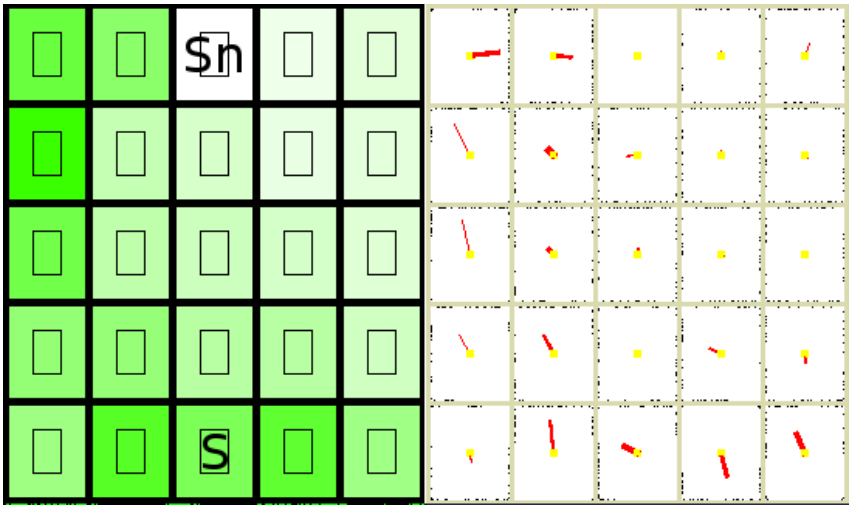
of the tissue. Figure 4.24 shows some more examples of where canalisation has occurred. Figure 4.24(b) is a good representation of the type of results seen under this hypothesis that do not follow the edge of the tissue. Although the cells in the centre of the space are transporting more auxin, it has not developed into the narrow file of cells that are seen in experiments using the Diligent Worker Hypothesis.

It has also been observed that if one particular cell gets a very high concentration of auxin it develops into what is essentially a high concentration sink. Even though it has lots of PIN they cannot stay on the membrane due to the rules of the mechanism. Shown in figure 4.25, the very high cytoplasmic auxin concentration makes it impossible for PIN to stay on the membrane. Although this is intriguing from our current results the mechanism is otherwise so unstable that it is not possible to suggest it as a mechanism for either canalisation or phyllotactic patterning. An updated version of the Genetic Algorithm software was used to do a limited search for parameters, the starting parameters used are shown in table A.2. A description of the updated software is in Appendix B.3.2.

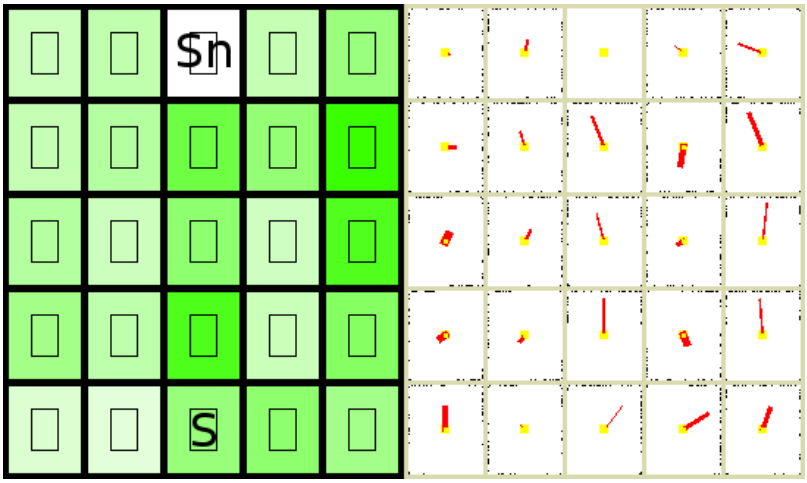
4.3 Diligent Worker Experimental Results

Presented in this section are examples of simulation runs showing that the mechanism described in the Diligent Worker Hypothesis (section 3.6) using the parameter values identified in section 4.1.3 can produce canals between a source and a sink in a variety of different situations. The aim was to test the mechanism is a wider variety of simulated 2D tissues. Figure 4.22 shows a straightforward run in a straight line between a source (marked S) and sink marked (Sn). Figure 4.26 shows a diagonal run, again between a source (S) and sink (Sn). Figure 4.27 shows a run where the canal goes around a block (blue line) and figure 4.28 shows a canal going around two blocks (blue lines). The blocks do not allow auxin to cross and therefore the canals are forced to go around the blocks. These example runs show that the hypothesis is able to produce canals under a variety of different situations similar to the situations of real canals going through plant tissues. Supplementary videos (OneBlock, TwoBlockA, TwoBlockB, and Diagonal) on the DVD show the process of canals developing.

These example runs were inspired by the experiments of Sachs[103]. Figure 4.29 reproduced from Sachs' paper shows canals going round corners. Al-

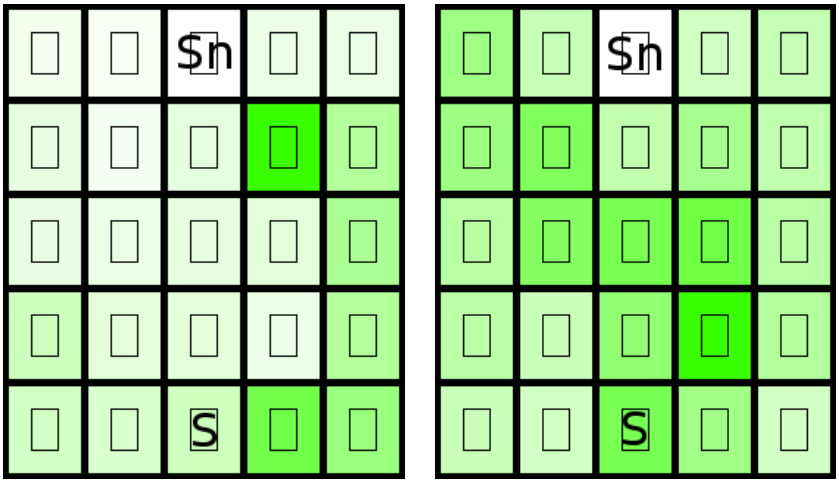


(a) Polar localisation of PIN has occurred and a canal formed between the source (S) and the sink (Sn).



(b) Polar localisation of PIN has occurred and a canal formed between the source (S) and the sink (Sn).

Figure 4.23: Internal Auxin hypothesis tests. Both part (a), and part (b) show that polar localisation does occur with the hypothesis, and canals have formed between the source (S) and the sink (Sn) in the tissue. However a common feature of this hypothesis is that when canalisation occurs it is frequently around the edge of the simulation space.



(a) A canal formed between the source (S) and the sink (Sn) following the edge of simulated space.

(b) A canal has almost formed in a straight line between the source (S) and the sink (Sn).

Figure 4.24: Internal Auxin hypothesis. (a) is an example of the type of canal most commonly formed using this hypothesis. The canal forms around the edge of the simulation space. (b) shows an example where an canal has almost manage to form directly between the source and the sink. The canal is not very well developed and is not typical of the results.

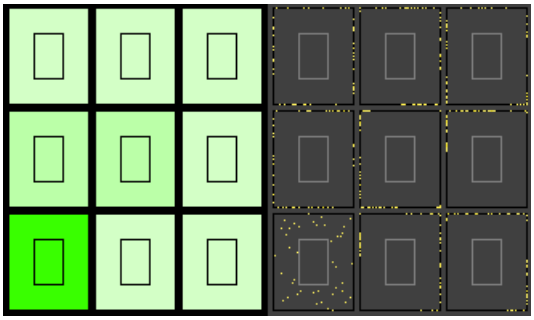


Figure 4.25: Shown here is a section of a simulated tissue where one cell has become a high concentration auxin sink. This cell has so much auxin in its cytoplasm that under the Internal Auxin hypothesis the PINs are constantly dropping off the membrane. This effectively means that this cell is no longer exporting auxin and looks like a sink.

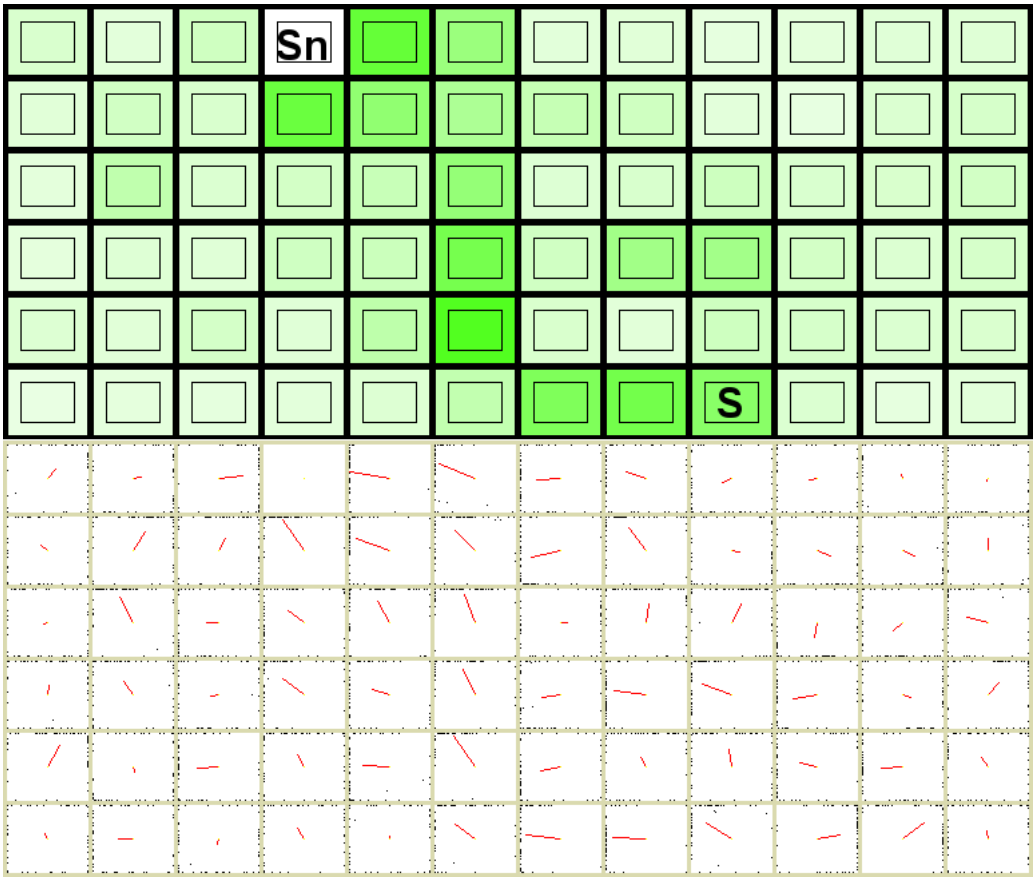


Figure 4.26: A diagonal test, an auxin transport canal is formed between the source cell 'S' and the sink cell 'Sn'. Note: this older simulation run has a datafile that does not support generating PIN polarisation sticks of varying thickness.

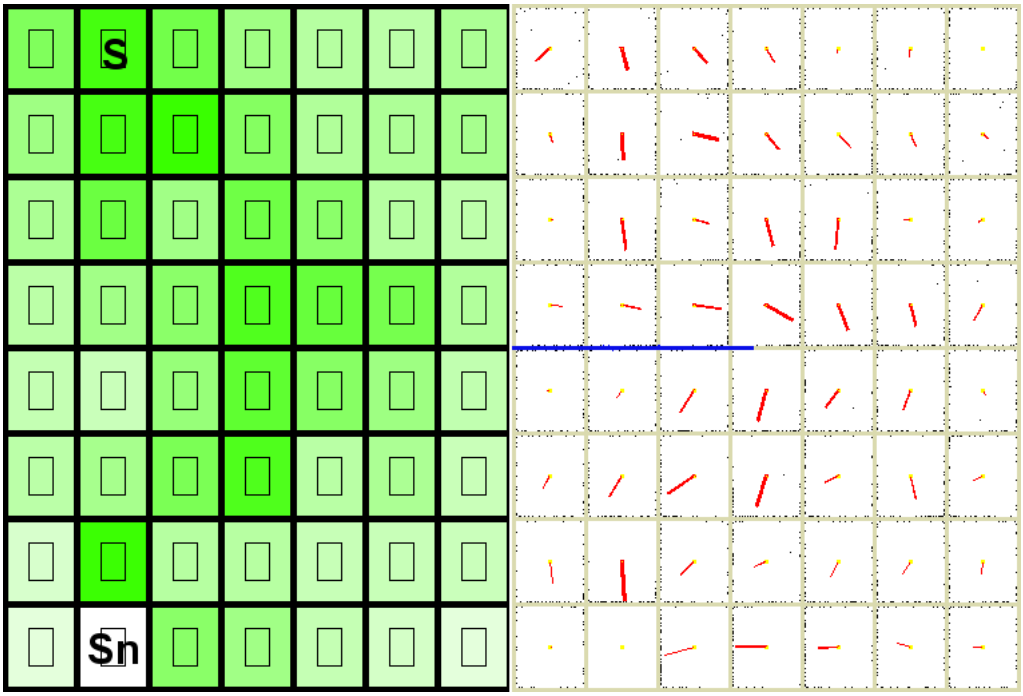


Figure 4.27: One block test. An auxin transport canal is formed between the source cell 'S' and the sink cell 'Sn'. Auxin is not able to cross the block (shown in blue) and therefore the canal is forced to go round the block to join the source and the sink together.

though figures 4.26, 4.27 and 4.28 show experiments running on smaller numbers of cells it is still encouraging to see this behaviour from the simulator.

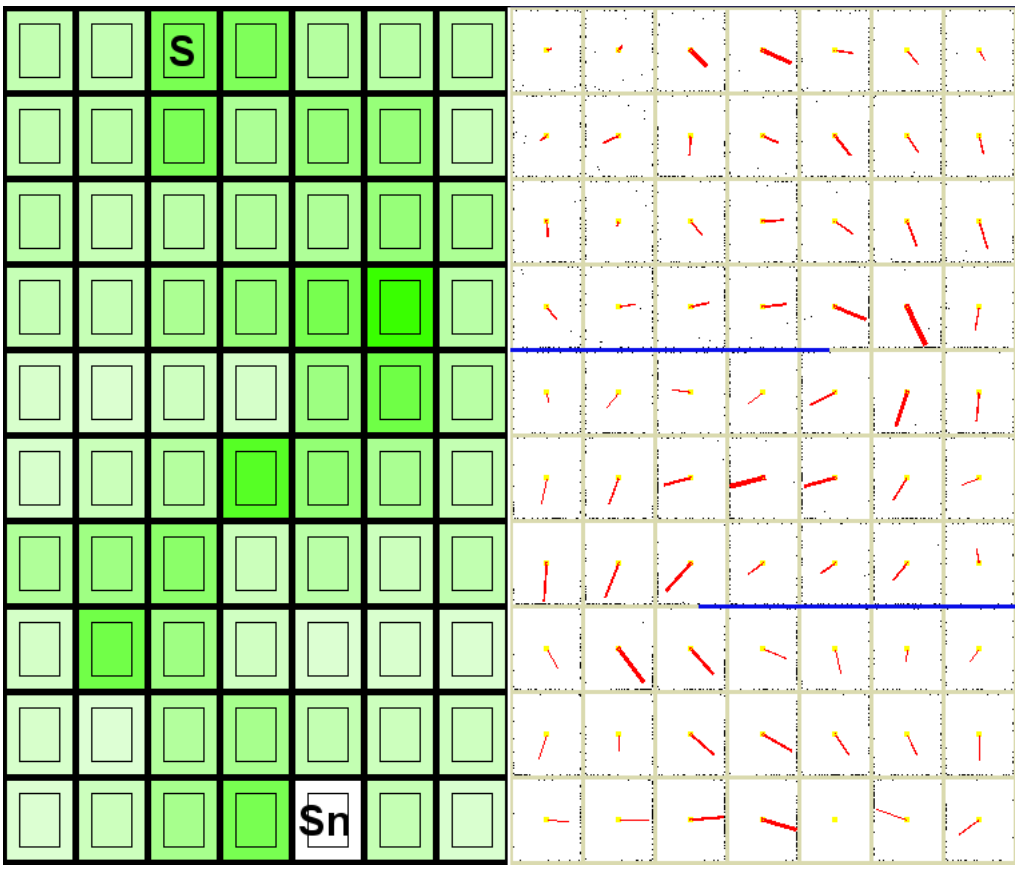


Figure 4.28: Two block test. An auxin transport canal is formed between the source cell 'S' and the sink cell 'Sn'. Auxin is not able to cross the two blocks (shown in blue) and therefore the canal is forced to thread its way around the blocks to join the source and the sink together.

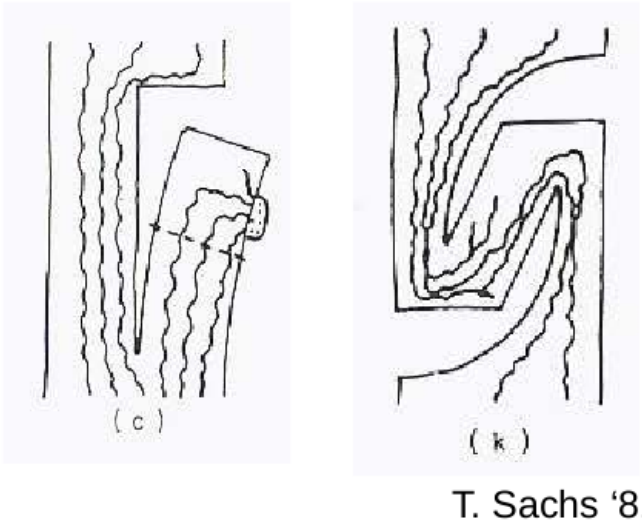
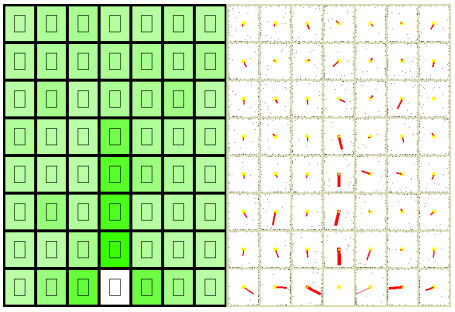
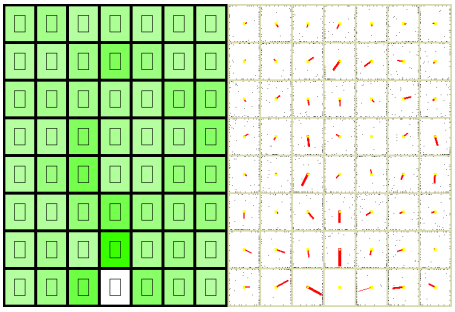


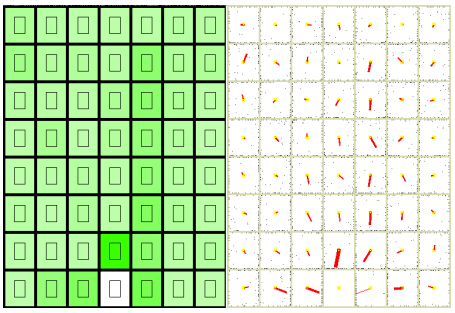
Figure 4.29: Figures from Sachs'[103] paper showing the formation of vascular tissue in response to applied auxin. The canals which form are able to go around corners introduced into the tissue.



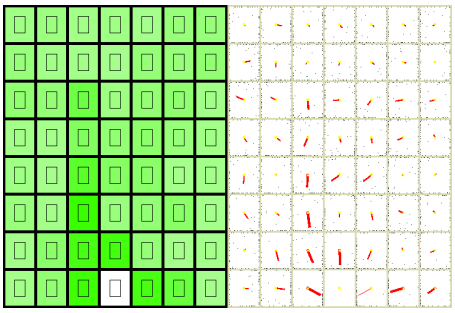
(a) An auxin canal has formed from the sink up the centre of the tissue to just over half way.



(b) The auxin canal has moved slightly off to the left. Another canal is almost present off to the right. The two looked like they were going to meet to form a loop.



(c) The auxin canal has switch to the right hand side of the sink and almost reaches the top of the tissue.



(d) Finally is has switched back to the left of the sink.

Figure 4.30: When tested in a tissue where all cells except the sink produce a background level of auxin canals do form but are more unstable. This series of figures show how a canal changed position over time during one particular simulation run.

4.3.1 Canal Emergence from Background Auxin

Also tested were runs where all cells maintain a background concentration of 0.4 (a.u) of auxin (figure 4.30). There is no high concentration auxin source but there is a sink. Canals do form but they are unstable, the cells that are involved tend to change over time. Figures 4.30(a), 4.30(c), and 4.30(d) show canals that have formed roughly up the centre of the tissue. Figure 4.30(b) shows a tissue where a canal formed up the centre slightly to the left. This was interesting because a second canal was developing to the right and the two looked like they were going to meet.

4.4 Conclusion

Using a combination of manual tuning and GA assisted parameter search it has been possible to find parameters that allow the mechanism described by the Diligent Worker Hypothesis to produce auxin transport canals. The results show that the ability of the mechanism to work is not overly sensitive to the values used for the different parameters. Experiments using the Diligent Work Hypothesis are shown in section 4.3 show that auxin transport canals form in different experimental setups. It has not been possible to find parameters for the Internal Auxin Hypothesis that reliably produce auxin transport canals. This might be because the full parameter space has not been searched, or that mechanism itself needs alteration.

Using a GA to conduct the parameter searches proved to technically more difficult than at first through, due to the amount of time to complete one generation and the difficulty of developing a suitable fitness function.

Chapter 5

Development of Second Generation of Model

The second generation simulator was developed to address some of the limitations of the first. It was specifically designed to remove limitations of cell shape, to provide a simplified route to 3D simulations, and improve performance. An important secondary outcome was improvements in the structure of the simulator code. The processes of enhancing the simulation was again supported by the CoSMoS process. CoSMoS describes a lifecycle for developing simulations. There is no defined end point, and therefore it is well suited to improving existing simulation code.

In section 5.1 an overview of the updated CoSMoS process as used for enhancing the simulations is provided (CoSMoS was under continuous development at the same time as the simulations). In section 5.2 the Research Context is briefly introduced, part of the CoSMoS introduced between simulator generations. In section 5.3 the changes to the biological Domain Model are briefly discussed. In section 5.4 the main changes relating to modelling space that we are addressing in this generation are discussed. In section 5.5 we discuss how the Platform Model has been updated using the CoSMoS process.

5.1 CoSMoS Process: The modelling lifecycle

During the development of the simulations the CoSMoS process underwent a few changes (summarised in figure 5.1, technical reports [2] and [1] provide more detail). The main differences (described below) include renaming of some of the models and the introduction of the **Research Context**. The results model has as yet not been completed for the simulations, instead the models

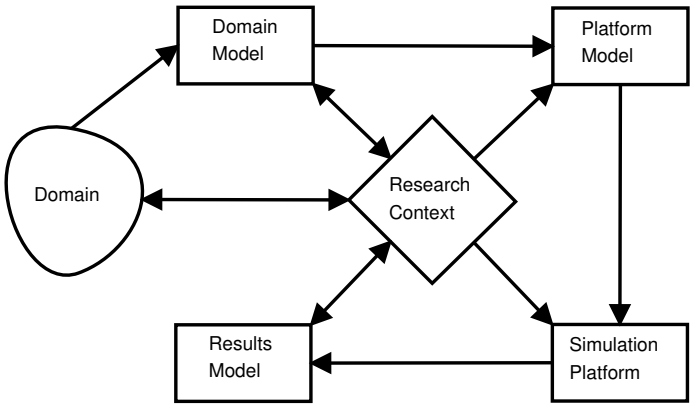


Figure 5.1: The components of the updated CoSMoS process [1, fig.2.1]. Arrows indicate the main information flows during the development of the different components. There is no prescribed route through the process.

are validated against existing auxin transport models and the biology. There is no reason why the Results Models for both generations of simulator could not be completed in the future.

Research Context : the overall scientific research context. This includes the motivation for doing the research, the questions to be addressed, and the requirements for success.

Domain Model : conceptual "top-down" model of the real world system to be simulated. The Domain Model is developed in conjunction with the domain experts, with its scope determined by the Research Context. The model may explicitly include various emergent properties of the system.

Platform Model : (called the Software Model in 3) a "bottom up" model of how the real world system is to be cast into a simulation. This includes: the system boundary, what parts of the the Domain Model are being simulated; simplifying assumptions or abstractions; assumptions made due to lack of information from the domain experts; removal of emergent properties (properties that should be consequences of the simulation, rather than explicitly implemented in it).

Simulation Platform : the executable implementation. The development of the Simulation Platform from the Platform Model is a standard software engineering process.

Results Model : a "top down" conceptual model of the simulated world. This model is compared with the Domain Model in order to test various hypotheses.

5.1.1 Modelling biology and simulations with UML

As with the development of the first generation model UML was used to develop the second. After the initial development of the first model it underwent a phase where attempts were made to improve it. During this time UML diagrams were generated showing the structure of some of the extended simulator. The diagrams assisted in making the decision to redevelop the simulator in a more systematic way, to allow a number of improvements to be made at the same time. The UML was very useful during this period as it showed that ad-hoc changes were having a detrimental effect on the simulation code. Some of the intuitive mappings of UML and OO code to the biology where being lost or not fully exploited. During this time traditional biological 'cartoons' were created to go with some of the UML diagrams to provide extra information.

The most significant modifications can be seen in the Platform Model and associated Simulation Platform Model. These reflect changes to significant assumptions about what should be removed from the Domain Model. The main progress made in the design and implementation of the second generation model has been with the handling of the simulation space, allowing the cells of the tissue modelled to be more naturally shaped. Chapter 7 discusses how the improved 2D simulator has been adapted into 3D.

5.2 The Research Context

The auxin transport community studies many different aspects of auxin transport. These include, but are not limited to: auxin transport canalisation [101, 102]; shoot branching regulation [61, 62, 82]; leaf venation [105]; and phyllotactic patterning [56, 97]. These processes are concerned with the developmental patterning of a plant, at both the tissue level and that of the whole plant.

This research sits within this wider community; it uses background biology derived from the literature, and from wet-lab experiments (for more information see Chapter 2). Again, as with the first generation model, the primarily focus is on modelling the processes of PIN localisation hypotheses, within the context of auxin transport canalisation. There is also the longer term goal of investigating auxin transport canalisation's role in shoot branching regulation.

There are many published mathematical models of auxin transport. This work continues the development of executable models, this modelling technique lends itself to biological systems, and can offer an alternative perspective [24, 31], particularly as we are modelling the PIN protein transporters at

a reasonable level of detail.

Our models focus on the question of PIN cycling and its role in canalisation, and we aim to test different regulatory mechanisms of PIN cycling.

5.3 The Domain Model: auxin transport canalisation

The domain of this model remains auxin transport canalisation. The only difference between the generations of models is that we are taking a fresh look at some of the assumptions in the model. Particularly whether the shape of the cells has a bearing on the success of the hypotheses for the regulation of PIN localisation. This is a change in what has been included from the biological domain into the Domain Model and then fed down through the modelling process.

5.3.1 Domain Model UML

The very high level UML remains the same between the two generations. The high level UML does not capture information about the space in the simulations.

Domain Model use cases. This is used to capture a high level view of what the system does, such as the regulation of proteins and hormones. No changes are required between the two generations.

Domain Model class diagram. This captures the biological entities of interest as objects and classes. Objects map naturally to biological entities such as proteins, hormones, and cells. Cells themselves are composed of a number of objects such as membranes, cytoplasm and vacuoles, which are associated with each other in space. We also need to regulate the production of agents like proteins and hormones, which is done by cells. This again remains the same between the two generations.

Domain Model state diagrams. These remain among the most useful of the Domain Model diagrams for communicating with the domain experts, as they appear to map well to the way these biological processes are understood. State diagrams capture how an object changes through time. They are able

to show the different possible states of the biological objects, and how an object moves from one state to another. Some spatial information can also be captured by state diagrams, as the changes can be associated with a location, within and outside a cell. For example the possible state changes that the auxin object can undergo are different depending on whether it is inside or outside a cell. State diagrams map neatly to the traditional biological 'cartoon' showing processes occurring in cells (such as figure 2.2). The behaviour of auxin can be cross-referenced between the 'cartoon' and the Domain Model auxin state diagram (figure 3.5).

5.4 Modelling Space

Here we discuss an important part of the model that was not explicitly dealt with in the first generation model: space.

The simulation space is a part of the biological domain that cannot easily be captured using UML, and might be based on assumptions that could escape recording. The space in which our biological entities exist is implied in the UML. We can see from the domain class diagram (figure 3.3) that we are representing part of a Plant built from a number of Cells (each with a CellMembrane and Vacuole), surrounded by Apoplast.

However the nature of that space is not captured, nor is any information about how the objects such as CellMembranes or Vacuoles are arranged into Cells, nor how the Cells and Apoplast are arranged into a plant tissue. This information might seem obvious, since it is easy (particularly if you work in the field of plant science) to imagine what a small 2D section of plant tissue might look like. This aspect is easy to capture with a more traditional 'cartoon' and explanatory documentation.

In our initial simulation the assumption is made that a 2D rectangular 'box' is an adequate representation for a plant cell. Therefore the initial simulation is limited to 2D cells of four straight sides. This is a reasonable simplification to make; mature cells in the stem of a plant are often fairly block-like in shape. However, auxin transport canals also form through tissues with cells of varying size and shape, particularly at the interface of a bud and existing vascular tissue. Therefore being able to test the behaviour of our hypothesised regulation of PIN localisation in cells of more natural shapes would be interesting both from a biological and simulation point of view.

Linked to this is the need to try to investigate the effect that 3D cells would

have on the behaviour of the hypotheses. There are a number of differences between real 3D cells and simulated 2D cells that might have an effect on the localisation of PIN. Being able to simulate even a small number of 3D cells could provide interesting insight into the effect of abstracting 3D cells into 2D. Early simulations have been done in 3D, but it is not well implemented in the initial simulation. We also want to allow for more naturally shaped 3D cells.

The first of these issues is linked to the way in which space (the environment of the agents) in the model is handled. This impacts a number of key areas: the interaction between the agents and the space, and how the space is split up into cells and the other structures in the plant tissue.

These modifications are more about changes in the level of abstraction assumed during the development of the Platform Model, about how the simulation is to be constructed from the Domain Model. Sometimes it is possible to change existing simulation code to allow for the change in abstraction. In our case the changes are significant, and the development process of the first simulation highlights a number of areas where improvements could be made.

5.5 Platform Model

The Platform Model includes all the extra components that allow the simulation to run. This includes all the processes required to get the simulation to a point where it is able to start, such as generating the space and populating it with cells.

The Platform Model has three kinds of information: biological processes captured directly from the Domain Model; biological processes required for the proper functioning of the simulation, but not of explicit interest to the Research Context, implemented with regard to efficiency rather than biological fidelity; instrumentation and other such aspects of a simulation that are not part of the Domain, but are needed to observe and document the simulation results.

Throughout the continued developmental process it is the Platform Model that has seen the most change. Not only have we made efforts to make the simulated space more realistic with respect to the real plant, but huge improvements have been made in the data output from the simulations and the organisation of the code.

5.5.1 Platform Model UML

Platform Model use cases: these capture the user requirements for using

the simulator, the traditional use for use cases in software engineering. These are unchanged from the first generation version section 3.2.3 [33].

Platform Model class diagram. This is produced from the Domain Model class diagram, with all emergent properties (such as the Auxin Canal) removed. This high level diagram shows mainly the biologically relevant parts of the model, and is unchanged from the first generation model (figure 5.2).

Platform Model class diagram, implementation level. As we move towards code, implementation level data structures are added to the class diagram. Section 5.5.2 discusses the changes to the implementation level Platform Model class diagram.

Platform Model state diagrams. These follow the Domain Model state diagrams and remain largely unchanged from the first generation (see section 3.3.3 [33]).

As the simulator increases in complexity, keeping the high level and implementation level Platform Models distinct becomes increasingly important. Things that are not biologically relevant, but are needed in a simulator (such as the ability to easily checkpoint to allow restarting) add complexity to the model that biologists do not need to see. We therefore omit such detail from the high level Platform Model diagrams discussed with the biologists, and retain it in implementation level Platform Model diagrams used by the developer.

5.5.2 The Division of Space

The main changes we made in moving from the initial to the enhanced version were to the way the space is handled in the Platform Model and simulation. The initial version treats the space as a largely homogeneous area, a grid of pixels, on which cell membranes and vacuoles are drawn, dividing the space into separate areas. Some areas are associated with objects like Vacuole and CellMembrane; other areas are essentially null.

A CellMembrane is a continuous line enclosing the cell (figure 5.3A). It is straightforward to define a cell membrane if it is built from straight line segments. However it is more difficult to define realistic-shaped cells with curved membranes (figure 5.3B) using this approach. The membranes would need to be drawn correctly somehow, and then read into the simulation. It would be easier to place the cells into the space as continuous areas of cytoplasm, and

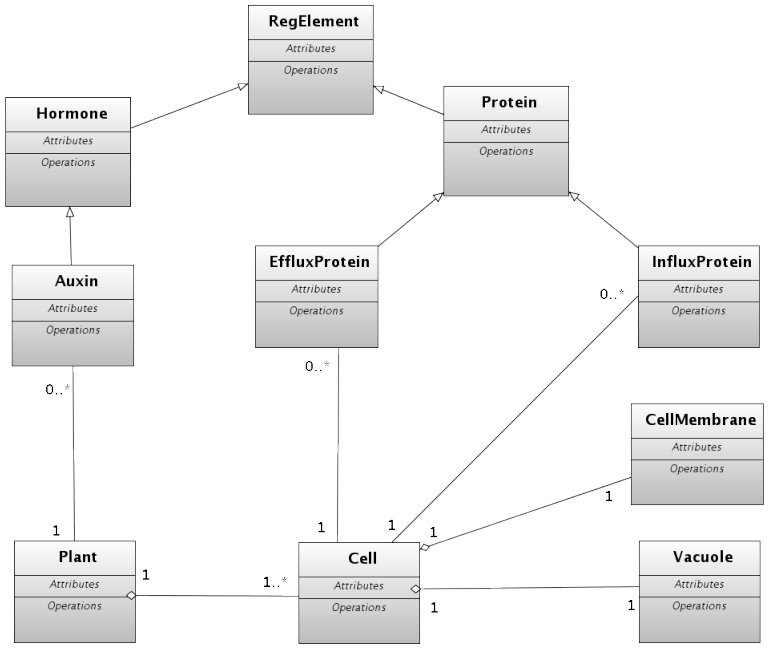


Figure 5.2: Platform Model class diagram, the same as figure 3.9 [33, fig.10]. Note that space is not explicitly dealt with, rather it is generic unless something like a CellMembrane object is put into a position.

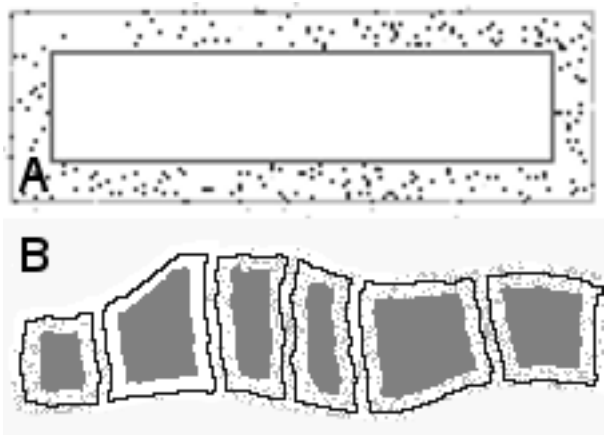


Figure 5.3: (A): Section of visual output from the initial simulator. The thin line of the cell membrane (outer grey line) is drawn into the space to define the cell. The vacuole is defined by drawing another membrane (darker grey line). This is a simple task for boxes, but more difficult for natural shapes. (B): Section of visual output from the enhanced simulator, showing a continuous curved membrane (black line).

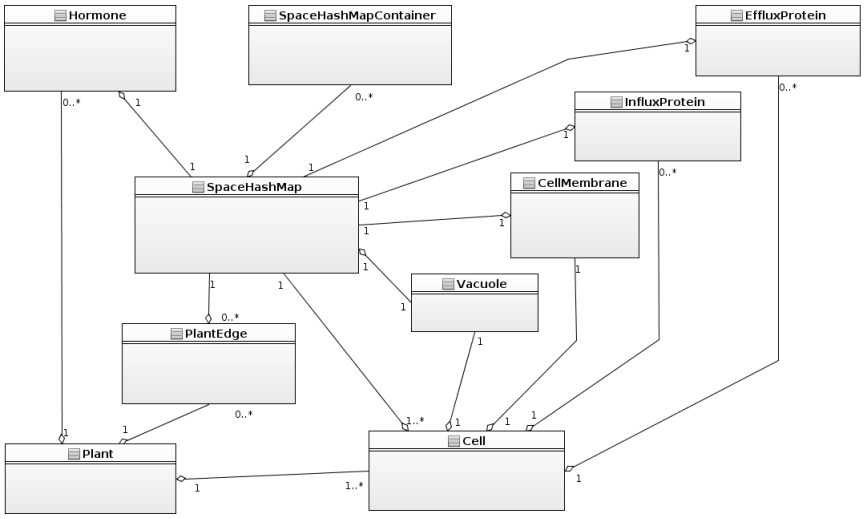


Figure 5.4: Implementation level Platform Model class diagram of the initial simulator. All objects in space require access to a singleton class SpaceHashMap that provides them with information about the space they are in, via the SpaceHashMapContainer class. As more kinds of space are needed, the resulting code becomes inefficient and untidy. (Inheritance has been left off this diagram to improve readability.)

then determine the position of the membranes around the edge (which is how it is implemented in the enhanced version). A new method of handling the space needs to be able to address such issues. The new method of handling space also needed to allow for easier addition of new types of space.

In the initial version of the model, all space is described by a single object. Figure 5.4 shows the relevant part of the implementation level Platform Model class diagram. A single class, SpaceHashMapContainer, has different attributes that allow it to represent all of the different types of space in the simulation, depending on the values the attributes are given. However, the complexity and size of this class increases each time we add a new kind of area of space in the simulation.

Another significant issue with having all the kinds of space specified in a single class is that some of the methods in the class need to behave differently depending on what the kind of space the object made from the class is representing. This increases the complexity of the individual methods in the class as they must contain code for all the different behaviours. The organisation of the first generation simulator code suffered from having a richer space added to the model, something that could have been avoiding if it had been designed with space in mind from the beginning.

For the enhanced version, we refactor the code to handle the space in a

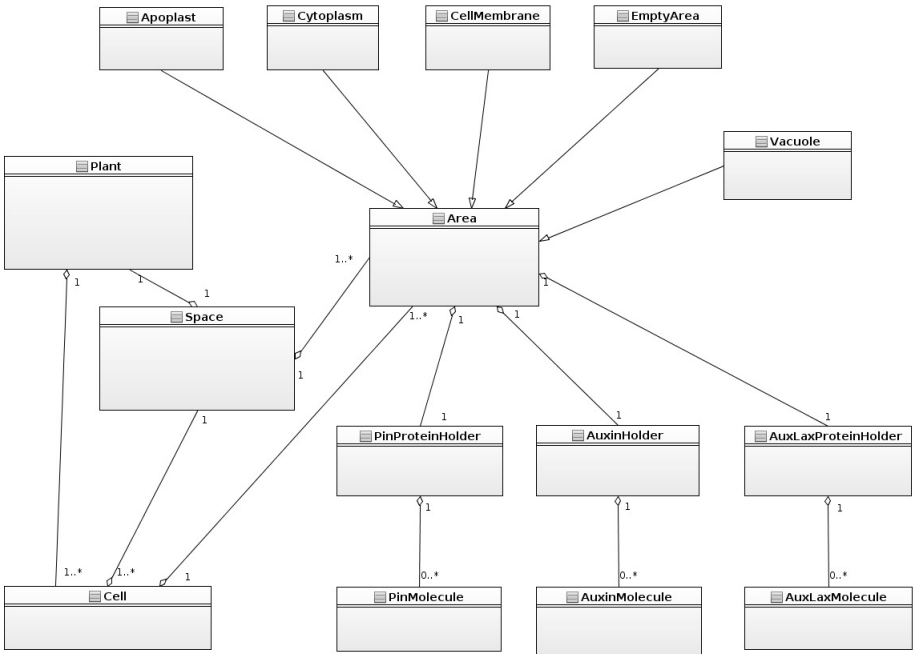


Figure 5.5: Implementation level Platform Model class diagram of the enhanced simulator. The space is now built from different child classes of the Area class, each with a holder looking after the different Molecules. The Space contains many areas which compose a single Plant, the Plant has many Cells. Cell requires access to the Space directly but also contains within it a list of all its associated Areas. A Cell does not directly contain any Molecules. (Inheritance of the different Holder and Molecule classes are not shown, to improve diagram readability.)

more area-specific manner, to improve its structure and extensibility, and to allow more natural-shaped cells.

In the initial model, space is general unless it is given a particular type. In the enhanced model, all the space is given an area type. An abstract class, Area, has attributes common to all the different types of area in the simulation. Sub-classes extend the abstract Area class into more specific kinds of space. Currently there are five types of area. Cells have Cytoplasm, Membrane, and Vacuole. Outside the cells there is Apoplast: the cell walls. Finally there is EmptySpace; this is used to allow more elaborate shapes of space to be used in the models, and is not processed. Apoplast areas separate all the cells from each other, and also separate cells from EmptySpace (see figure 5.5).

The abstract Area class contains many attributes and methods common to all the different types of area. These attributes and methods tend to be the system aspects of the class, such as accessing the colour of the object or its position in the space. The specific area type then adds extra methods that

give that space more biologically specific behaviour, and if necessary overload particular methods. This has many advantages, including simplicity of code maintenance reducing the likelihood of introducing errors. When a new type of space is added to the model much of the code is already in place.

5.5.3 Agents in Space

In the initial version, the code that determines how the agents move around in the simulation space is held in the agents themselves. This results in the classes describing the agents becoming more complicated each time a new kind of space is added to the simulation. The agent requests information about its current environment from the environment directly. It then uses this to make an appropriate decision about what it should do. There is also an inconsistency in where the agents are stored. Figure 5.4 shows that auxin (Hormone objects) are held in the Plant class, but the proteins are in the Cell class. This makes *biological* sense, since the PIN and AUX/LAX proteins do not leave the cell, but auxin does. However it makes better *implementation* sense to think as all three as being held in the Space, and whether or not this is in a Cell is determined by what the space is. This is the case for the enhanced simulator, as shown in figure 5.5.

The movement of agents is also the responsibility of the Space in the enhanced simulator. Each Area sub-class that can have agents contains an AgentHolder with methods for storing the agents that are contained within it. The different AgentHolder sub-classes (such as AuxinHolder) for each agent inherit properties from the parent AgentHolder, but are also given specific behaviours. The AgentHolder classes accept incoming agents to their area. The movement of the agents is controlled by the Area sub-class, which has methods for moving any agents in the relevant AgentHolder. This puts the responsibility for moving agents onto the Area class. Therefore when a new kind of space is added, the areas are updated to allow agents to move into this new kind of space. These changes are reflected in figure 5.5 (to improve the readability of the diagram, the inheritance from the abstract AgentHolder class and the Molecule class are not shown).

5.5.4 Space from Templates

The more natural-shaped cells are defined using templates derived from images of real plants.

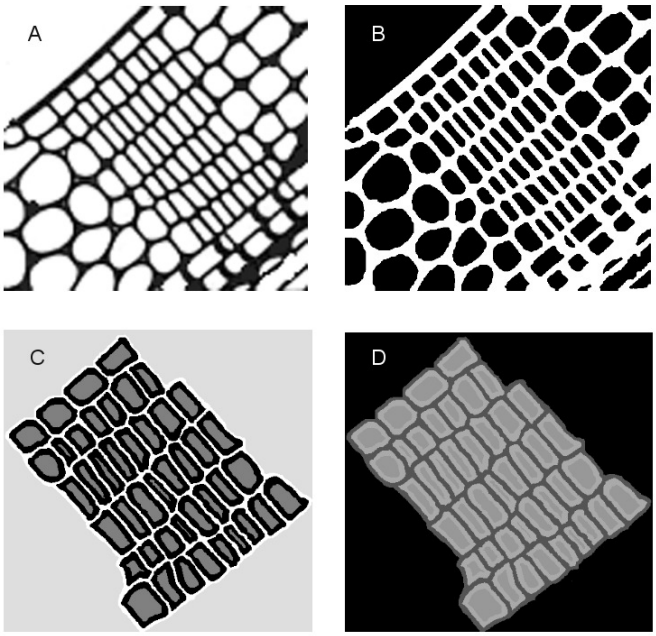


Figure 5.6: Processing sections of plants into the model. If the section photos are of high enough quality the processing can be done automatically. (A) Photographic section from a real plant, tided up to allow it to be processed. (B) Image processed for reading into the model: black areas will become Cytoplasm, white areas Apoplast. (C) Image modified by hand to isolate a patch of cells: light grey areas will become EmptySpace. Vacuole areas are then added automatically (dark grey). (D) Template as it finally appears in the simulation visualisation. CellMembrane areas are added automatically at the interface between Cytoplasm (here light grey) and Apoplast (here dark grey).

Figure 5.6 shows the lifecycle of a template: it starts as an image of a section of a plant, and ends as a representation of the simulation space. Templates can either be generated automatically (normally with a little manual processing), or fully by hand. They need to contain only three pieces of information: the areas of the space that are empty (not active as simulation space but required to be spatially present); the areas that are apoplast; and the areas that are cells. The template is then processed to add vacuoles into the cells. These are not added directly from the image being used because simulated 2D cells need smaller vacuoles than are shown in sections of real 3D cells. Instead they are added automatically by filling the centre of the cell a certain amount (see section 7.0.1 for discussion of this). Cell membranes are also added automatically around the cytoplasm. Once the vacuoles and cell membranes have been added into the space we essentially have areas presenting cell cytoplasm, cell membranes, vacuoles, apoplast and any empty areas. All are displayed as different colours in the image (shown as different shades of grey in the figure).

In the simulation the space is created to match the pixel size of the template, and the entire space starts off as apoplast. Each pixel of the template is then read and its colour determines what it is in the space. The next task is to group areas of continuous cytoplasm and the vacuole inside them into the more abstract notion of a cell. In a plant, a cell is essentially a container of elements that need to be held together. The elements have no concept of togetherness, they are just associated in space. The way the different elements interact is through the common environment. In the simulation a cell is more abstract. It is similar in that it contains lists of all of its spatial contents but it also needs methods to create more proteins or hormones when they are required. Essentially the nucleus of a real cell, which regulates what is expressed, is part of the more abstract Cell class in the simulation. The Cell class provides access to the common environment, to allow cell regulation.

5.6 Conclusion

This Chapter describes how the CoSMoS process was used to take an existing simulation and enhance it to improve its performance, add features, and clean up the implementation. The enhanced simulator has been used run simulations on larger numbers of cells, and over more naturally-shaped cells.

Continuing to use the CoSMoS process assisted by UML as allowed the simulations to be updated in a efficient and systematic way. Using this ap-

proach helped identify which of the assumptions we made when making the transition to the Platform Model from the Domain Model might need to be reassessed. Both the CoSMoS process and UML helped to highlight when progressing down a particular development path was increasing the gap between the biology we were trying simulate and how we were implementing it.

Going through the CoSMoS process identified when it was necessary to return to the Platform Model of the simulator to include more natural cell shapes derived from the biology. Both the CoSMoS process and UML allowed the identification of parts of the simulator code that were becoming over complicated and could be improved. From this it was possible improve how the biology of the Domain Model is captured in the Platform Model, and simultaneously improve the simulator code itself.

Chapter 6

Results From Second Generation Model

The purpose of the new generation model was to test the effects of two abstractions, and to try to improve the performance of the simulator. In this chapter section 6.1 briefly looks at the performance of the two simulators. Section 6.3 looks into the question of whether 2D grids of cells with four straight lines is a good enough environment for testing hypotheses for PIN localisation. Will our most successful localisation hypothesis behave differently when more naturally shaped cells are used. Section 6.4 presents experiments investigating the effect of the size of the vacuole on the localisation of PIN. Carried out with this simulator with both square and irregular cells, this work was partly in preparation for comparison with 3D cells. The enhanced simulator was also developed to be able to test the effect of abstracting 3D plant cells into 2D simulators, and what they would mean for hypotheses of PIN localisation. This work is not complete, but work done so far is presented in chapter 7.

6.1 Improved Simulation architecture allows performance improvements

The improvements in the underlying simulation code and multithreading have provided performance improvements as well as improved code structure and simulation flexibility. Many of these improvements have been back-ported into the first generation model, thus closing the performance gap and making the first generation simulator faster when only one computational thread is available. Figure 6.1 qualitatively compares the performance of the second

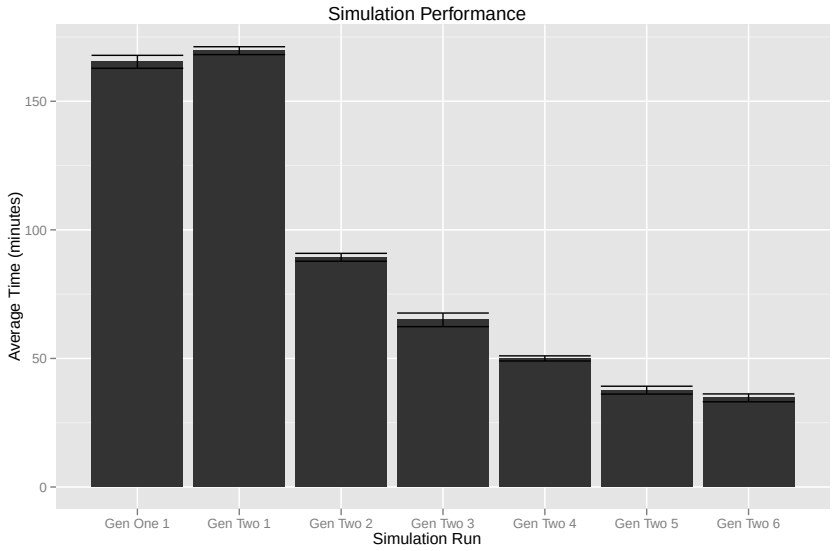


Figure 6.1: A qualitative comparison between the performance of the two generations of simulator. Shown is the time to forty-five thousand time steps on a 5x5 grid of 2D cells. The first generation simulator using one processor thread (Gen One 1) takes 163 minutes. Due to the back-porting of performance improvements this is slightly quicker than the second generation simulator with 1 processing thread (Gen Two 1) which takes 170 minutes. The second generation simulator can take advantage of multiple processors and achieves a significant performance boost from using up to 5 threads. Beyond five and the over head of the threads blocking each other removes the advantage of using more processors. Shown is an average of three independent runs.

generation model with the first, and the improvements from the ability to take advantage of multiple processing cores. The plot shows the time to forty-five thousand time steps on a 5x5 grid of 2D cells for both the (updated) first generation, and second generation simulators (experiments run on a 16 core 2.3GHz AMD Opteron 8356 computer with 32GB of RAM). Much of the cost of improving performance comes in the form of increased memory requirements.

6.2 Comparison with first generation model

The underlying structure of the second generation simulation is different to the first generation simulation, but the high level behaviour of the agents in the two models is exactly the same. There is therefore no point in reproducing all the same experiments and parameter searches with this version. Figure 6.2 shows that when some of the basic simulations carried out with the first generation model are repeated with the second generation model the results are

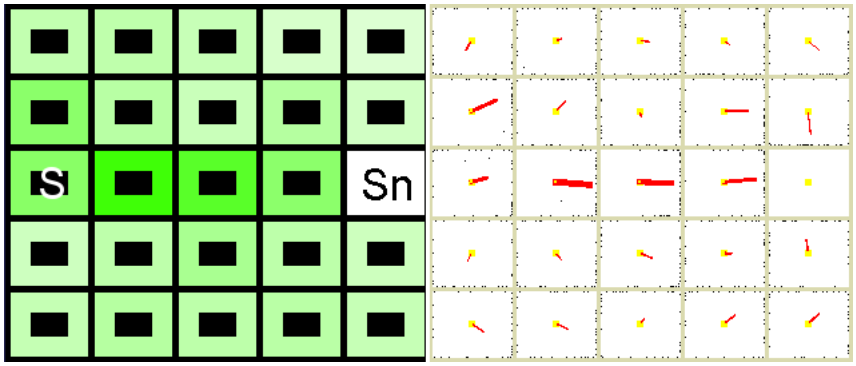


Figure 6.2: Basic test with regular cells. The canal forms between the source 'S' and the Sink 'Sn'.

very similar. Figure 6.3 shows that when testing a 5x5 grid of cells the fitness of the canals formed is comparable between the two generations of simulator (allowing for the stochastic nature of the two simulations). With the first generation simulator scoring approximately 0.46 and the second generation scoring approximately 0.41 with regular cells. The fitness achieved with the second generation simulator in irregular cells is slightly lower at 0.29. The fitness of the simulations was calculated using the method described in section 4.1 and an average of three independant runs is shown. Experiments were also done diagonally through rectangular regular cells (figure 6.4).

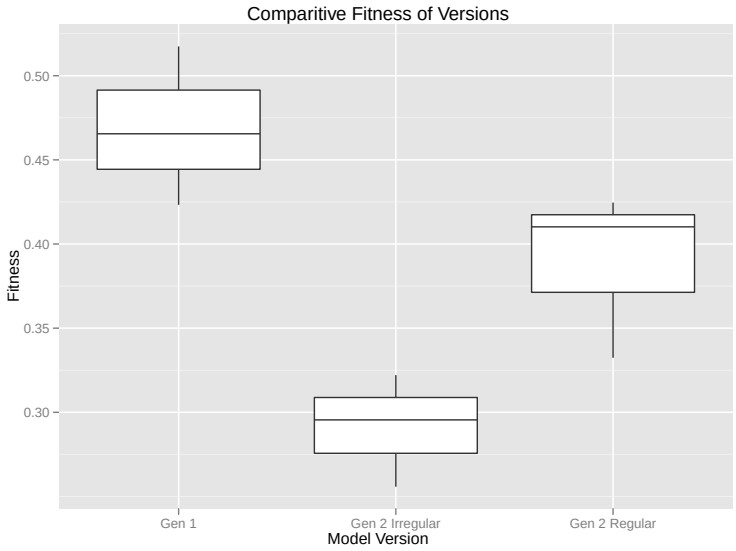


Figure 6.3: Box and whisker plot showing a comparison of the fitness of a canals formed through a 5x5 grid of cells over three independent runs of the simulators. A run with the first generation simulator with regular cells, then a run of the second generation simulator with irregular cells and finally a run of the second generation simulator with regular cells. When using regular cells the different generations of model achieve comparable fitness values. When using irregular cells in the second generation model the fitness is slightly lower due to the fitness calculation being affected by canals generally not being as neat.

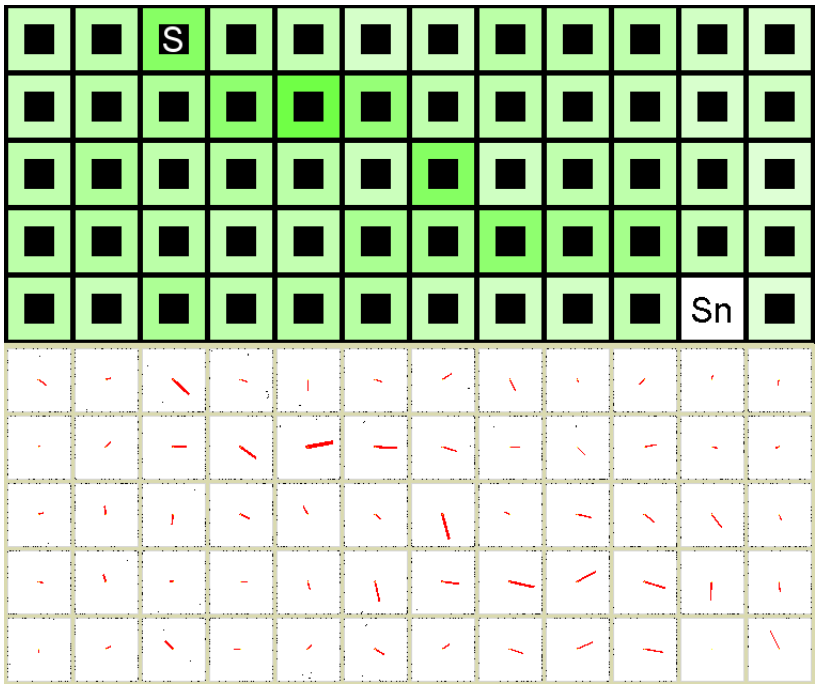
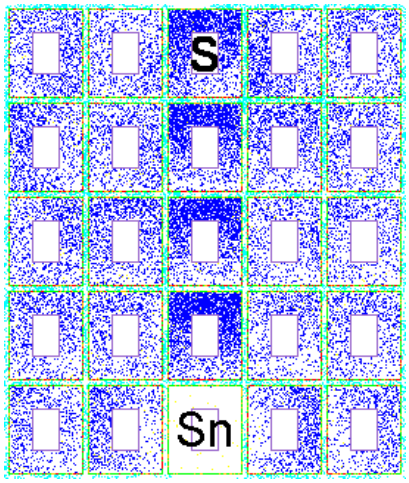


Figure 6.4: Basic diagonal test with regular cells. The canal forms between the source 'S' and the Sink 'Sn'.

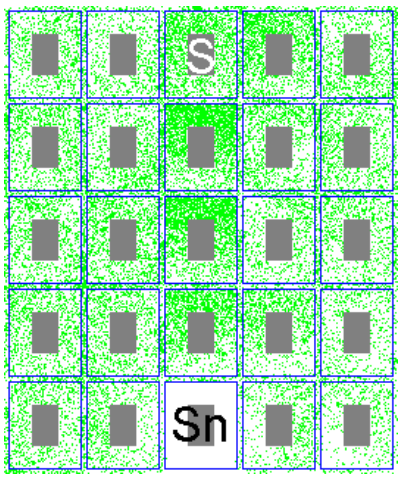
6.3 Irregular Cells

During testing of the second generation model it was found that canals would not form as well through irregular cells as they would through regular cells using the same parameters (figure 6.3). Qualitative assessment of the results suggested that PINs were slower to polarise in the cells. When investigated it was found that the Cell Membrane areas in irregular cells have fewer Apoplastic area neighbours than those in regular cells when averaged over all membrane areas (figure 6.6). Adjusting the EPC (see table 4.1) parameter from 9.5 to 8.6 compensates for this effect and closes the qualitative gap between canals formed in regular and irregular cells.

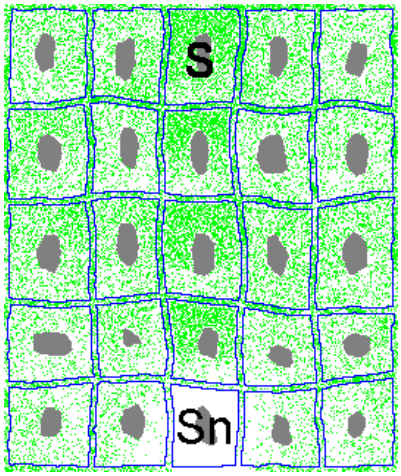
Experimental simulations were carried out that are similar in design to those tested with rectangular cells. These include a couple of tests of canals forming in roughly straight lines through irregular cells (figure 6.5(c), and figure 6.5(d)). Figure 6.7 shows a canal formed in a line through a larger tissue of irregular cells. Finally figure 6.8 shows a canal going around a block. The templates for these experiments were drawn by hand.



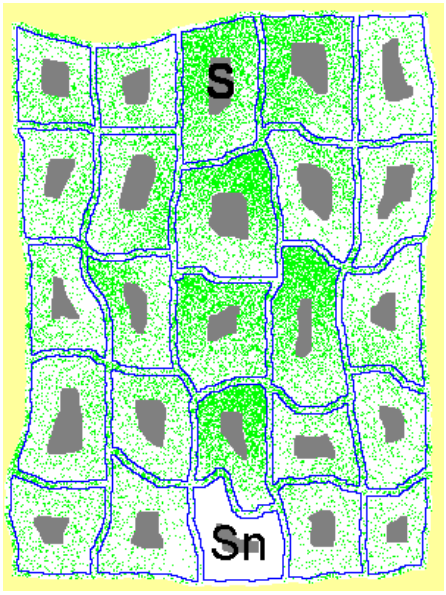
(a) First generation model. A canal formed between the source 'S' and the Sink 'Sn'.



(b) Second generation model. A canal as formed between the 'S' and the Sink 'Sn'. Using the same parameters as the first generation model.



(c) Second generation model using slightly irregular cells. A canal has formed between the auxin source 'S' and the Sink 'Sn'. If the cells are only slightly irregular the same parameter values used in the first generation model can be used.



(d) Second generation model using irregular cells. A canal has formed between the auxin source 'S' and the Sink 'Sn'. When more irregular cells are used the parameters have to be adjusted

Figure 6.5: Comparison between the first generation model and the second. When testing regular cells, or cells that are very nearly regular, the parameters used in the model can remain exactly the same ((a), (b) and (c)). However, it was discovered that the cell membrane areas of very irregular cells (d) have fewer apoplastic are neighbours and therefore the EPC parameter has to be adjusted from 9.5 to 8.6.

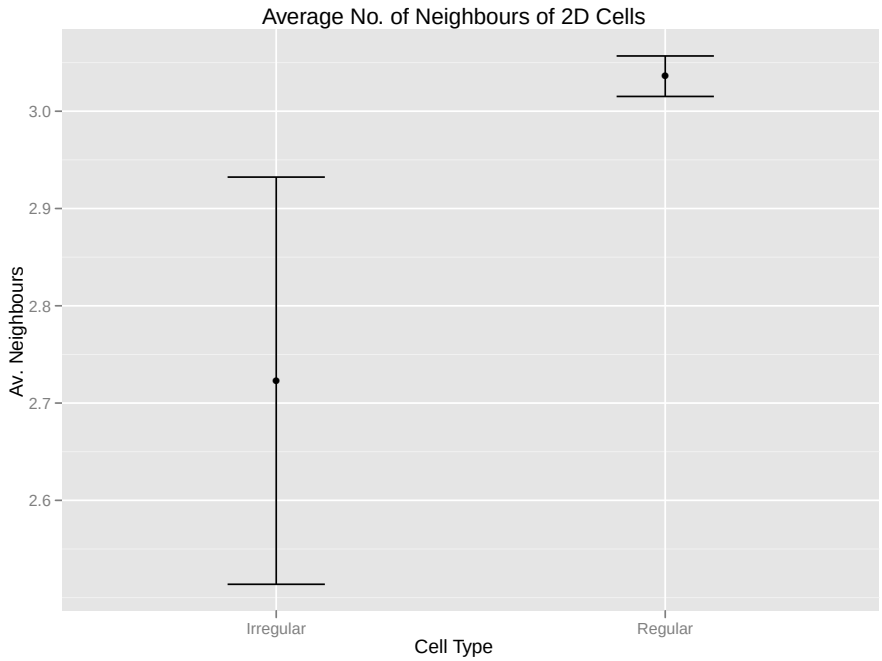


Figure 6.6: Plot showing the average number of Apoplastic neighbours a Cell Membrane area has. Data shows the average from three different tissues, error bars are the standard deviation of the results. The Cell Membrane areas of irregular cells have on average fewer Apoplastic neighbours than cell Cell Membrane areas in regular cells. This effects the formation of canals and requires the EPC parameter to be changed from 9.5 to 8.6.

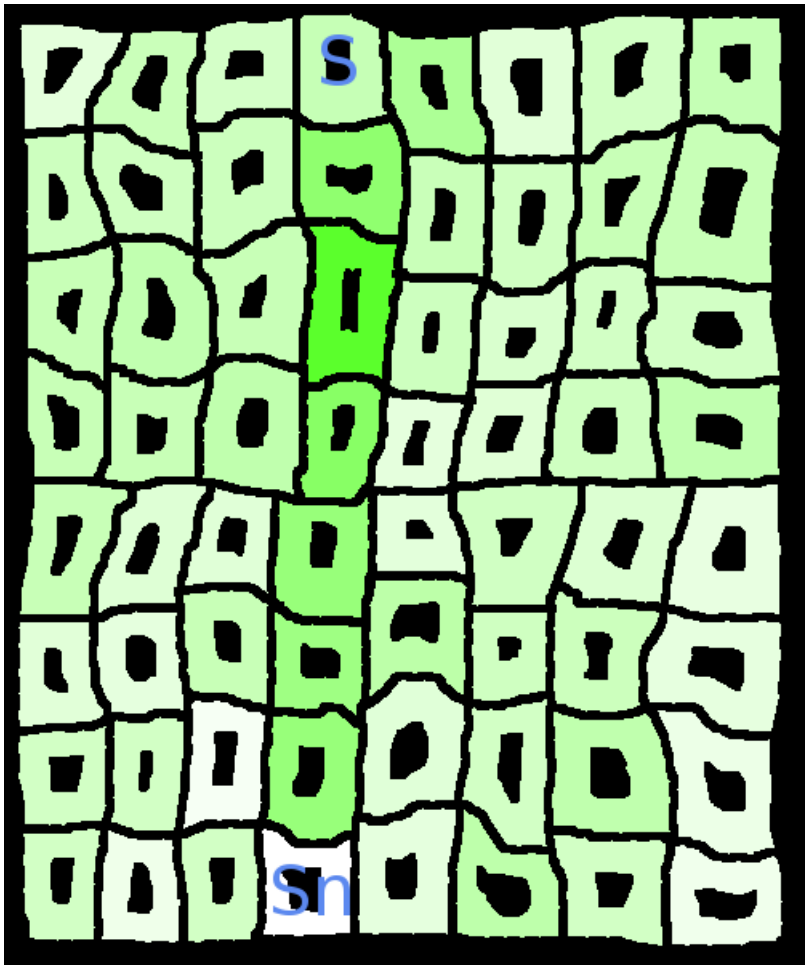


Figure 6.7: An auxin transport canal has formed in a straight line between an auxin source cell 'S' and a sink cell 'Sn' through an 8x8 grid of irregular cells.

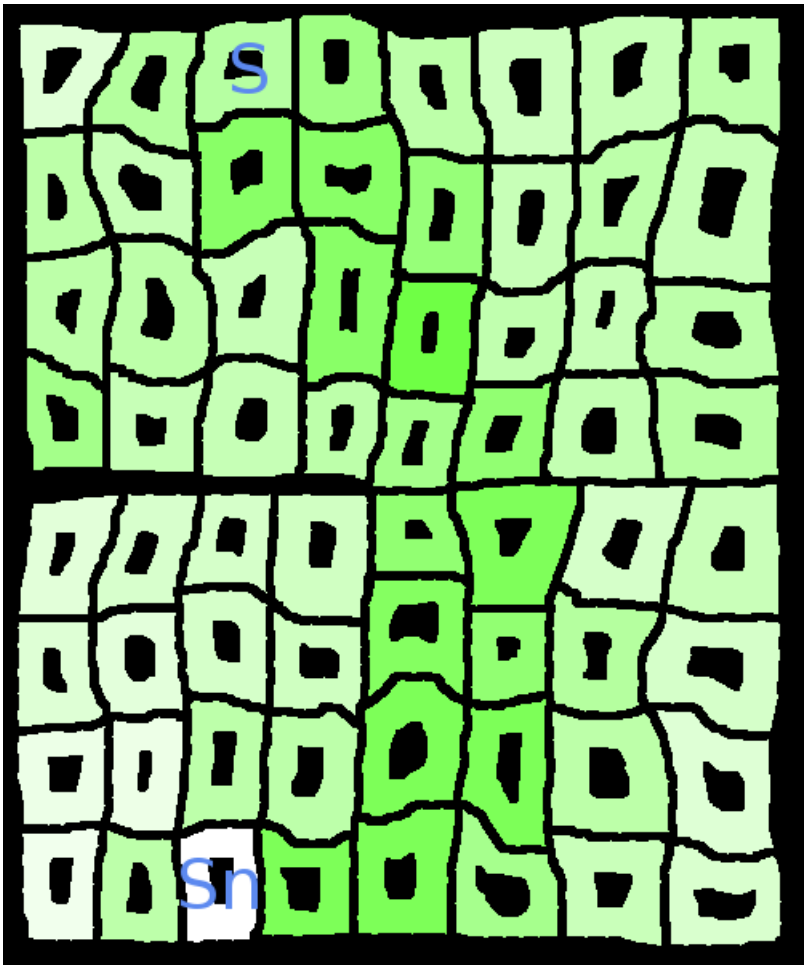
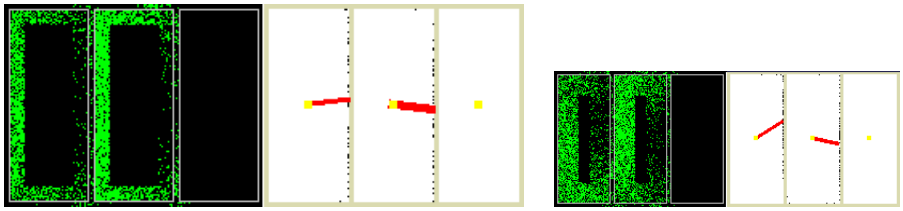


Figure 6.8: An auxin transport canal has formed between an auxin source cell 'S' and a sink cell 'Sn' around a block through an 8x8 gird of irregular cells.



(a) Even if the vacuole is occupying most of the cell area PIN distributes fairly evenly along the membrane.

(b) When the vacuole occupies a smaller amount of the cell area PIN distributes evenly along the membrane.

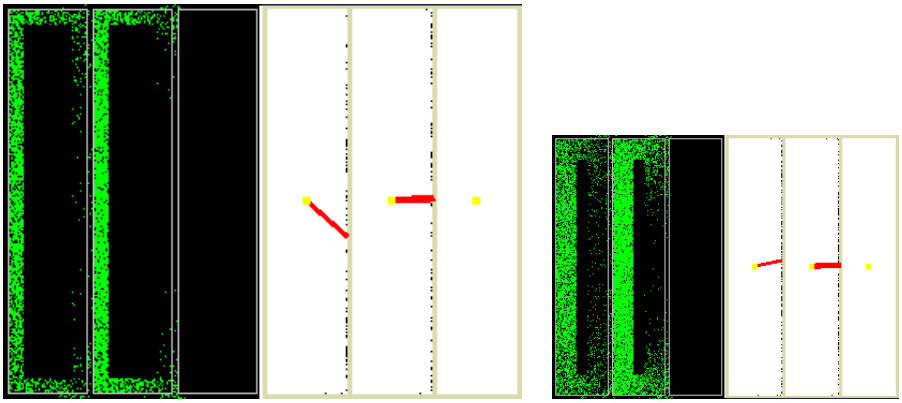
Figure 6.9: In cells of a shape approaching that of a square PIN polarisation is not adversely affected by vacuole size. Even if the vacuole is large when compared to the total cell area the PIN is able to distribute evenly along the membrane.

6.4 Vacuole Size

Simulations were run to test the effect of changing vacuole size on PIN polarisation. These tests were carried out on small numbers of rectangular and irregular cells of different sizes and vacuole sizes.

Rectangular cells that are approaching a square in shape do not strongly demonstrate any effect due to vacuole size, even if the vacuole takes up a very large amount of the cell area. Figure 6.9 shows that PIN is able to distribute evenly along the whole length of the membrane facing towards the sink. Effects can be seen in rectangular cells that are much longer than they are wide. In these cells PINs can causes parts of the cell to be strongly depleted of auxin which eventually will cause the PINs in these depleted areas to dissociate from the membrane and return to the cytoplasm. This tends to cause patches of PIN to form on the membrane, normally towards the corner of the cell (as shown in figure 6.10(a)). The effect disappears if the vacuole area is reduced relative to the cell (figure 6.10(b)). The effect of the vacuole was partly overcome by adjustment of the parameters of the Diligent Worker Hypotheses during the parameter search phase (chapter 4 section 4.1.3) but it can still be seen. Irregular cells show the same effect; those that can be considered as having a shape that is approaching circular behave in roughly the same way as square cells. Very long irregular cells behave like rectangular cells and suffer the same problem of patchy PIN distribution (data not shown).

These experiments also highlighted some interesting observations concerning the PIN localisation hypothesis. The effect of the hypothesis is to encourage the PINs to pump auxin from an internal environment with a ready supply of auxin, into an external environment where the auxin is able to leave. The



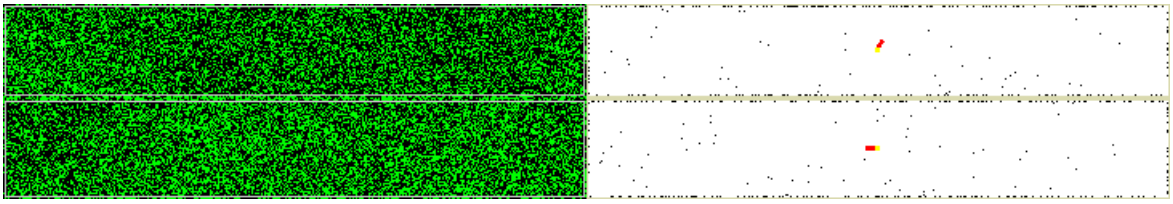
(a) In larger rectangular cells, when the vacuole occupies most of the cell area PIN distribution tends to be patchier, with one or two denser patches along the membrane.

(b) In larger rectangular cells, when the vacuole occupies a smaller amount of the cell area PIN distributes evenly along the membrane.

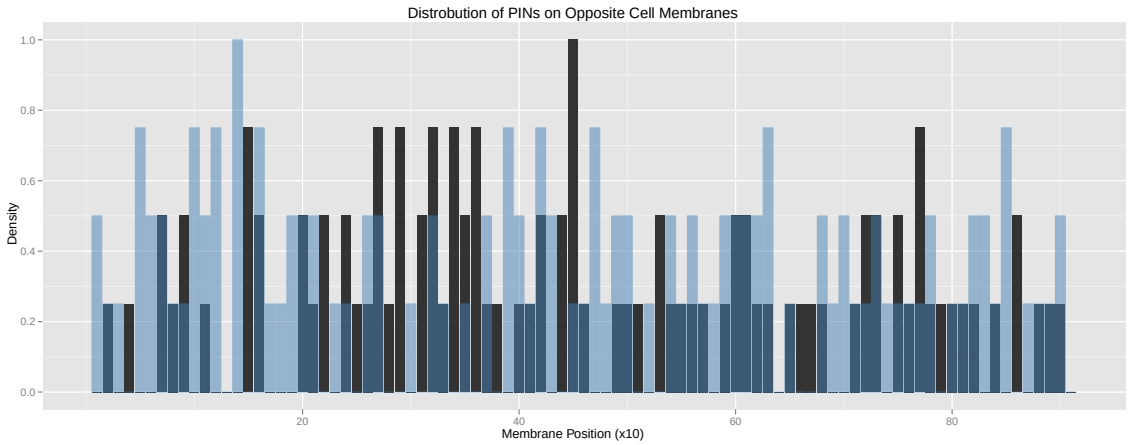
Figure 6.10: In cells of a more rectangular shape PIN polarisation is affected by vacuole size. If the vacuole is large when compared to the total cell area the PIN tends to form higher density patches along the membrane.

effect of high apoplastic auxin causing PINs to disassociate means that PINs do not *like* pumping auxin into an environment that lots of other PINs are also pumping auxin into. The result is that if PINs from two neighbouring cells are pumping auxin into the same area of apoplast the auxin concentration is likely to reach a level high enough to start causing PINs to drop off membranes. This can be seen fairly well when very long thin cells are placed next to each other.

Figures 6.11, 6.12 and 6.13 show an experiment where two long thin cells are placed together. Both cells are made auxin sources. Once the simulation is started both cells reach the set auxin concentration and the amount of PIN in the cells starts to increase. At first the PINs randomly start associating with the membranes (figure 6.11). Figure 6.11(a) shows PINs are starting to randomly associate with the cell membrane. PINs do not remain on the membranes facing the edges of the simulation space as the auxin pumped here has nowhere to go and tends to build up. Instead, the PINs tend to build up on the membranes facing the neighbouring cells. Once on this membrane they are affected by the PINs on the opposite membrane. Therefore patches of PINs start to form along the membrane, and they tend to form out of sync with similar patches on the membrane of the other cell (figure 6.12). Figure 6.12(a) shows the PINs on the membrane. The patches are more clearly represented by the histogram of PIN distribution shown in figure 6.12(b). This distribution is very interesting as it has been hypothesised that similar patching of PIN proteins occurs during the development of leaf pavement cells. If left longer the PINs eventually settle at one end of the cell, and this is the opposite end to the PINs in the neighbouring cell. Figure 6.13 shows the PINs on each cell have become polarly localised. Once in this configuration the two cells are essentially pumping auxin around in a circle. The polar distribution can be seen clearly in 6.13(a), and an auxin gradient can also be seen in the cells. Figure 6.13(b) shows a histogram of PIN distribution on the membranes. The histograms represent the number of PIN agents on the membrane and each bar represents 10µm of membrane.

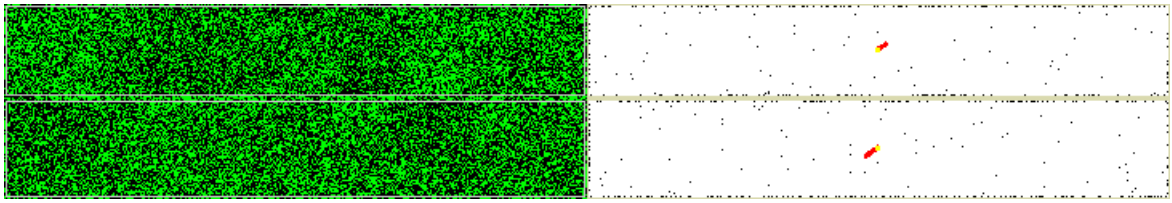


(a) Early PIN distribution.

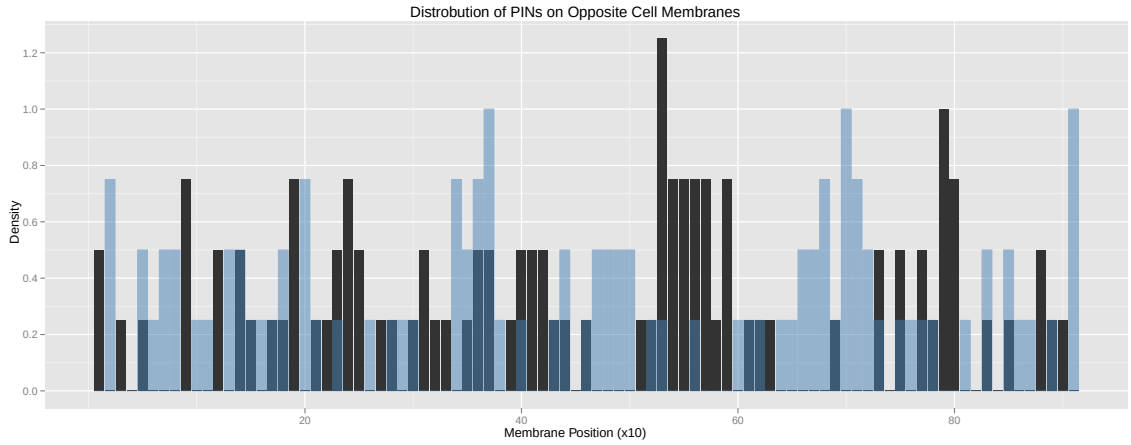


(b) Histogram of initial PIN distribution, each bar is number of PIN agents on a 10µm section membrane. Blue bar is the bottom cell.

Figure 6.11: Looking at the changing distribution of PINs on the membrane between two auxin source cells. Not long after the simulation is started the PIN is distributed fairly randomly around the membrane of the cells.

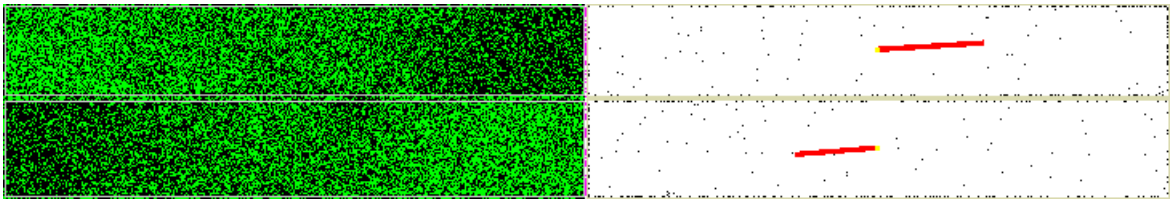


(a) Later PINs go through a temporary phase of forming patches along the opposing membranes.

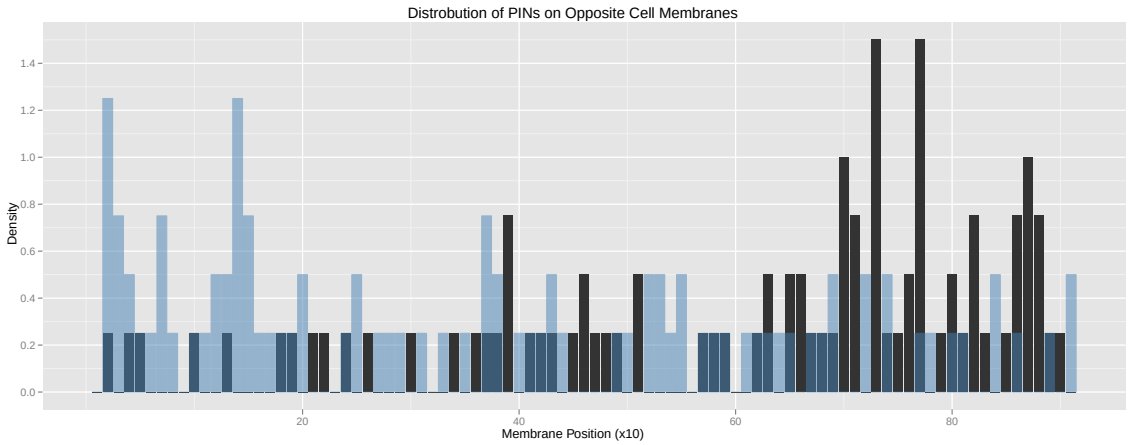


(b) Histogram of patchy PIN distribution, each bar is number of PIN agents on a 10µm section membrane. Blue bar is the bottom cell.

Figure 6.12: Looking at the changing distribution of PINs on the membrane between two auxin source cells. There is an intermediate period during which the PINs distribution becomes patchy along the membranes facing the opposite cell.



(a) Finally PINs settle at opposite ends of cells.



(b) Histogram of final PIN distribution, each bar is number of PIN agents on a 10µm section membrane. Blue bar is the bottom cell.

Figure 6.13: Looking at the changing distribution of PINs on the membrane between two auxin source cells. Eventually the PIN in both cells become polarly localised at opposite ends of the two cells.

6.5 Conclusion

The updated simulator has reproduced the auxin transport canals that the first generation simulator produces using the same parameters and the mechanism described by the Diligent Worker Hypothesis. It has also been able to test the Diligent Worker Hypothesis in more naturally shaped cells and has shown that the same mechanism is sufficient to produce auxin transport canals in irregularly shaped cells.

The updated version of the simulator can take advantage of multiple CPU cores to reduce simulation time and allow simulations to be run over larger numbers of cells. The improvements to the simulator structure and code will also allow for future testing of auxin transport in 3D cells. This version of the simulator will allow for more ambitious experiments to be tried in the future.

Chapter 7

3D Model

Throughout the development of the simulators, 3D versions have always existed. Having a 3D simulator is very useful for general testing of the performance of the hypotheses for PIN localisation in 3D cells. A 3D simulator can also provide a platform for looking at the effect of abstracting from 3D to 2D. One good example is vacuole size. A 3D simulator can be used to consider the difference between taking a slice through a 3D cell and turning that into a 2D simulation, and simulating a 3D cell directly. Improvements made between the two generations of models were intended to produce a 3D simulator that was directly comparable to the 2D simulator, easy to maintain, and that would allow any improvements and developments made in the 2D model to be easily transferred over. It also needed to match the specification of the 2D simulator, including having facilities for testing both regular and irregular cells. The second generation simulator code was designed with this in mind. The changes required to produce the 3D simulator are limited to a much smaller number of classes in the model than in previous generations, the most significant being the Space class, and visualisation/data output classes. Importantly the classes controlling agents like auxin and PIN are the same in both the 2D and 3D versions.

Unfortunately the visualisation and data output for the 3D model is not complete. However, some basic output is possible. Testing of the 3D model is difficult due to very high memory requirements and long simulation times. Obtaining results from the 3D model could be carried out in the future. However, it may prove more useful for testing auxin dynamics in a single cell, but should be able to verify if hypotheses work in a small number of cells.

7.0.1 3D Space

The improvements to simplify the maintenance of the 3D model are linked to the improved systems for implementing space. The enhanced simulator space is implemented by ensuring that all Areas know who their neighbours are, and therefore the move to 3D is simpler as it mainly involves giving the Areas more neighbours. The code for the 2D and 3D versions of the simulator are therefore very similar.

We can either generate block-shaped 3D cells from algorithms or naturally shaped cells by stacking prepared 2D templates together in a careful order to create a 3D space. This requires three kinds of templates containing: only Apoplast; Apoplast and Cytoplasm; Apoplast, Cytoplasm and Vacuole. Programs that automatically generate the slices have been made (see appendix B.5).

We are interested in 3D simulations to investigate how our hypotheses behave in 3D and the effects of using a 2D representation of 3D structures in simulations, particularly on the effects of vacuoles. If the possible paths an auxin molecule can take in a 3D cell with a large vacuole to that of a 2D cell with a large vacuole are compared, the possibility for a significant effect can bee identified. Figure 7.1 shows that in a 3D cell an auxin taking the path through the vertical section is has to travel much further than if it took a path through the horizontal section at the position of the dashed line. In the 2D cell there is only the vertical path. All other diffusing agents in the simulators have the same problem. This could have an effect on auxin transport in a 2D tissue. We can use the 3D simulation to help calibrate the required size of the 2D vacuole, or alternatively allow diffusion at a reduced rate through the vacuole to simulate the horizontal path.

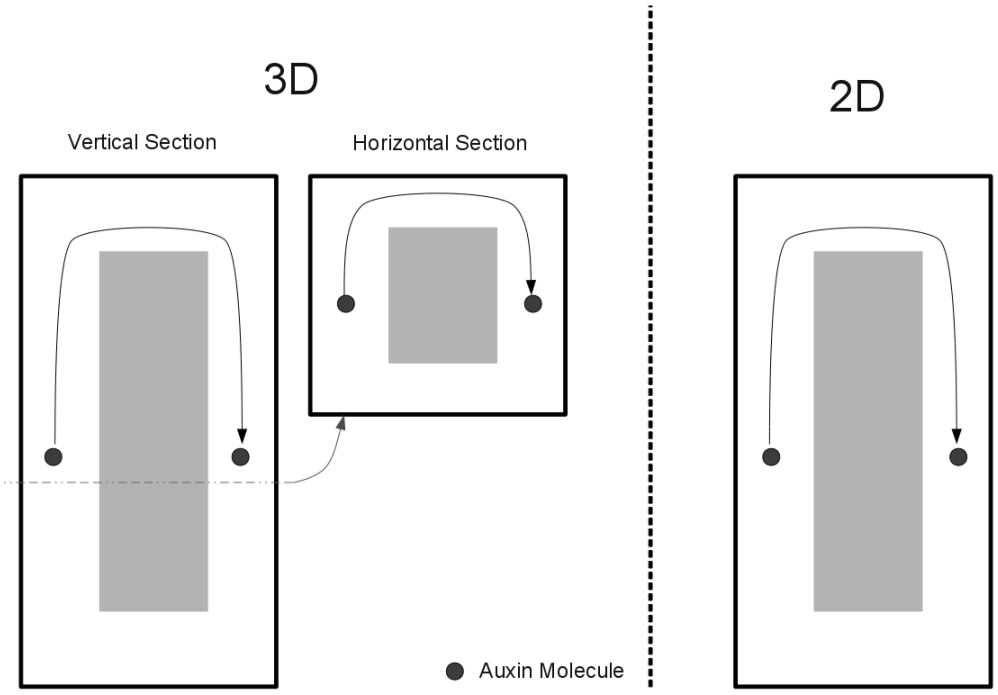


Figure 7.1: Comparison of possible paths of auxin molecules (or other agents) in 2D or 3D cells. In the 3D cell the auxin has the possibility of taking a short path to the same position. This is not possible in a 2D cell with only one path.

7.1 Conclusion

An addition goal of updating the simulator was to improve its ability to test different hypotheses of auxin transport canalisation in 2D and in 3D. Not only that when finished the 3D version of the simulator could prove a useful tool for modelling other cellular processes in 3D. Exploiting the 3D simulator remains a possible major area of future work.

Chapter 8

Discussion

Using the CoSMoS process with UML it has been possible to develop an agent based simulator for auxin transport canalisation that can allow for the testing of different hypotheses for the localisation of PIN. The mechanism described by the Diligent Worker Hypothesis 3.6 can successfully produce auxin transport canals through simulated tissues. The CoSMoS process encourages a systematic approach to the development process, helping to ensure that all decisions made at each point in the development process are understood and known. The first generation simulator has been tested extensively, and a number of limitations were identified. The main ones being an inability to have cells of a more natural shape, the improved 3D version of the simulation, and some more general performance issues. The CoSMoS process was then used to successfully enhance this existing simulator to address these three points. Using the enhanced simulator this work has shown that the mechanism described by the Diligent Worker Hypothesis is sufficient to produce auxin transport canals in irregular cells. There is a lot of scope for future work, including more experiments with the enhance simulator. The 3D version of the enhanced simulator could also be extensively used in the future.

8.1 Model Development

The CoSMoS process ensures that at each stage of modelling and simulator development effort is made to understand and acknowledge what decisions have been made and why. It is also flexible enough to work with software engineering tools like UML. UML is able to produce detailed information about the structure of a biological system. It is then possible to extend these UML descriptions of the biology into code skeletons of a simulator, even though

the final UML and code includes much more than just the underlying biology. That underlying structure should be visible (visibility can be improved by maintaining a separate Platform Model and Refined Platform Model), and areas where it has had to change or has been deliberately changed (such as the removal of emergent properties) can be highlighted and the reasons made clear. UML diagrams, particularly state diagrams, can be compared with more traditional biological 'cartoons' to enhance cross-disciplinary communication of model structure and included biology. This can help increase information flow between modellers and domain experts.

8.1.1 Development of First Model

Used for the development of the first generation model the CoSMoS development process (assisted with UML) has provided a number of advantages to our simulations. The diagrammatic nature of UML as a tool for producing various levels of models, including descriptions of program code, has helped produce simulations that not only work in an intuitive way, but that are also built intuitively. Biology maps to UML objects in a straightforward way that can be understood by developers and biologists alike.

The use of the CoSMoS process, with its different level models to capture system, simulation, and implementation details has helped produce conceptually cleaner models. The Domain Model looks purely at the biology. Here class diagrams and state diagrams are of greatest use. The class diagrams allow us to look at the static structure of the model, and how the different parts are connected together. They can include the emergent properties of interest, so that we have these properties captured rigorously in a model. State diagrams provide detailed information of how the objects change in response to events. They are normally produced by thinking about the known biology of the different biological elements. These map nicely onto traditional biological cartoons which can be used to provide extra more specific information about what is happening in a model.

The Platform Model does not explicitly include the emergent properties of the Domain Model: these should emerge from the interactions of the lower level simulated components, and can be compared against the Simulation Model for plausibility. The Platform Model can also transform biological components to behave in non-biological ways that are more readily simulatable. For example, it is necessary to look at the state diagrams to see if in order to produce a simulator we require more states to capture events than are provided in the

Domain Model. The existence of the two models highlights areas where the simulation is breaking with the biology, and if this break is causing a problem, the models can sometimes provide the solution. Additionally, it allows the simulation to be validated against the Domain Model. At this stage inheritance is added to class diagrams, to indicate classifications and generalisations, and to be used in implementation to reuse code and reduce duplication.

8.1.2 Enhancement of First Model

The CoSMoS process underwent development at the same time as these models were being developed. The CoSMoS process was again used produce an incremental change to the pre-existing model and resultant simulator. The enhanced simulator has improved performance, potentially allowing us to run simulations of canalisation over larger arrays of cells, and over more naturally-shaped cells. Canals still form in the latter case, indicating that the observed process is not an artifact of the rectangular cells. Slight alterations in parameters are required when using naturally-shaped cells. This is due to a slight reduction in the average number of neighbours a Cell Membrane area has in an irregular cell when compared to a regular one.

Continuing to develop our simulations with the CoSMoS process assisted by UML has ensured efficient and systematic progress. Using this approach helped to identify which of the assumptions made when making the transition to the Platform Model from the Domain Model needed to be reassessed. Both the CoSMoS process and UML highlight how progressing down a particular development path was increasing the gap between the biology we were trying simulate and how we were implementing it.

The CoSMoS process has allowed us to see that we needed to return to the Platform Model of our simulator to include more natural cell shapes derived from the biology. Both the CoSMoS process and UML allowed us to identify parts of the simulator code that were becoming over-complicated and could be improved. From this we were able to improve how the biology of the Domain Model is captured in the Platform Model, and simultaneously improve the simulator code itself. This can be clearly seen in the class diagrams produced as the first generation model was developed to an intermediate model, before the decision to follow a more systematic development process was made.

There is no end point to the CoSMoS process: it is continuous and during the development of the simulators it is intended that the developers will continue round the CoSMoS life cycle as required. This is a worthwhile process

as each time round the lifecycle highlights a number of areas that could be improved and as soon as the latest improvements have been integrated into the simulator code, then it is often the case that new idea for improvement or new requirements have been identified. Not only does the CoSMoS process help with identification of such improvements, but it also help to ensure that their inclusion in the simulator is systematic, and not ad-hoc. The developers and domain experts are also encouraged to discuss any changes by returning to the domain model before implementation takes place.

Modelling languages like UML could help people to better understand how biological models work. Therefore, their potential use in this area is becoming more important. Biology papers list details of how the experiments that produced the data were carried out in a lab. It should also be the case that how a simulation works and produces the data presented should be equally well explained, in order to allow independent verification of results and the sharing of methods and techniques among the modelling community. The increased use of modelling and the complexity of the simulations produced make this a more pressing need. UML can provide an effective way of developing and communicating simulations. It also provides scope for the possibility of producing interfaces between simulations, which might allow for simulations at different levels of abstraction to work together.

8.2 Simulator Results

Using the simulators produced we have been able to produce results for two different hypotheses for the localisation of PIN during auxin transport canalisation. One of these hypotheses was tested further in cells of a more natural shape using the enhanced second generation simulator.

8.2.1 PINs: The Diligent Workers of Auxin Transport Canalisation

The first generation simulator has been used extensively to explore two hypotheses for PIN localisation during canalisation. The most successful is the Diligent Worker Hypothesis. This hypothesis describes a PIN localisation mechanism that *encourages* PINs to remain on the cell membrane in a position where they are working productively (as defined in the hypothesis described in section 3.6). When PINs become associated with the membrane of the cell

they are able to export auxin from the cytoplasm into the apoplast. The PIN will remain on the membrane exporting auxin *ad infinitum* unless the local auxin environment of the PIN causes it to dissociate from the membrane and return to the cytoplasm. This can happen in two ways. If the PIN did not export auxin during the last time step there is a certain probability that it will dissociate. Or, if the local apoplastic auxin concentration becomes very high the probability of dissociation increases. The combination of these two mechanisms results in PINs remaining in position on the membrane where they have a ready supply of auxin to export, and an environment to export it into that is not experiencing a build-up in auxin (suggesting that the auxin has somewhere else to go).

The hypothesis mechanism is split into two parts, both under the control of a number of different parameters. The first part of the mechanism is the probability that a PIN agent will dissociate from the membrane if it did not transport auxin in the previous time step. The probability of this event occurring has to be kept very low otherwise is has a negative effect on the formation of auxin transport canals. A high disassociation probability prevents the PINs from polarising in the cell. Disassociation at a low value (1×10^{-10}) is useful however as this mechanism is the only other way that PINs can get off the membrane and return to the cytoplasm. This is important as it allows PINs to move locations in low auxin concentration environments where otherwise they might become stuck in one location. It is also important for PIN degradation in cells of low auxin concentration, as the simulation only degrades PIN from the cytoplasm.

The second important part of the hypothesis is the control of the probability of the PIN dissociation due to increasing apoplastic auxin concentration. This is under the control of two different parameters as the probability is changed by the environment of the PIN agent. A Hill function is used, with a parameter which changes the steepness with which the probability of dissociation increases (EPC set at 20) with apoplastic auxin concentration and a second parameter for the number of auxin agents in the PIN's environment that produce a 50% probability of dissociation (EPHD set at 9.5). Using these parameters PINs do not respond to low numbers of auxin agents in their environment, but once the number increases above 5 auxin agents the probability of dissociation increases rapidly.

The testing on AUX/LAX concentration follows the pattern of behaviour where if no AUX/LAX protein is present auxin transport canals form but do

not accumulate auxin[73]. Once AUX/LAX protein is added the canals are able to accumulate auxin[50]. It is reassuring to see the simulator matching the general behaviours reported in the field[50, 53].

Successful Experimental Runs

The Diligent Worker Hypothesis is able to produce auxin transport canals in a variety of situations. Chapter 4.3 shows results from experiments in regular cells using the first generation simulator. This includes a reduced scale attempt to replicate the results of Tvsi Sachs[103]. The hypothesis is very good at making a canal in a straight line through a small simulated tissue directly from a source to a sink. In this situation any effects due to the presence of the vacuole (discussed in section 8.2.3) are minimised and the canals formed are often narrow and represent the shortest path between the source and the sink.

In rectangular cells the hypothesis is able to produce canals that also go through tissues diagonally. Under this situation the canals do sometimes find the shortest path (figure 6.4) but are sometimes affected by the grid layout and therefore take a slightly longer path (figure 4.26). This is likely to reflect more on the use of rectangular cells rather than the hypothesis itself. It is possible that if allowed to run for more time steps all canals would eventually take the shortest path. However the experiments of Sachs showed that auxin transport canals do not always take the shortest path [100, 101]. Rather that once an auxin transport canal has successfully connected an auxin source with a sink it becomes established due to being an efficient transport stream for auxin even if its path meanders within the tissue. This can be seen from Sachs' experiments where canals formed around wounds might go a short distance and fizzle out, or wander around trying to find a new path to the sink. More experiments with irregular cells and cellular differentiation might help to unpick the significance of cell shape and the mechanism of PIN localisation have on path length.

The mechanism described in the hypothesis is also capable of producing canals that go around blockages in the tissue. In these experiments an impermeable barrier to auxin is put between some of the cells in the tissue. The source and sink cells are then chosen so that the normal route that the canal would be expected to take is blocked. In these cases the canals are seen to form around the block, and are still able to connect the source with the sink.

Finally some experiments were carried out in tissues were there was no single point source of auxin but instead every cell (except the single sink in the tissue) produced a background concentration of auxin (0.4 a.u). This was

to see if canals would form through the tissue from the sink. Canals do emerge from this set-up, but they are not stable and tend to move around in the tissue over time.

The resulting emergent canals are due to some cells randomly ending up with more auxin than their neighbours. The increased auxin means they have more PIN. If they are in contact with the sink the PINs become polarised towards the sink. The polarisation of the PIN makes these cells behave more like sinks to their neighbours, which may cause their PINs to polarise in the direction of this polarised cell. This produces the dynamic canals observed. This is interesting as the other possible outcome of this experiment would be that all cells would become polarised in the general direction of the sink over time. An effect that has been shown in the mathematical models developed by Feugier et al.[22]. It might be possible that adjusting some of the parameters would allow more stable canals to develop from background auxin production.

Cell differentiation could be added to the simulators and could potentially produce some interesting results. Differentiation could fix the PIN on the membrane of highly polarised cells. This cell would then remain in this highly polarised state. The cell would behave as a sink to the neighbouring cells which would eventually become highly polarised towards it. They could then undergo differentiation. Potentially this could cause the dynamic canals in the tissue to become fixed and stable. The difficultly is defining a method of making the decision to differentiate. Data does exist for how long it takes to reverse the direction of polar transport in differentiated cells by moving the auxin source. These are however mature cells and not cells developing from an auxin transport canal, and might re-polarise much slower [119]. Despite the lack of good data a number of mechanisms where considered. A good starting candidate would be to measure how long a cell had been polarised in a particular direction, and then fix the PIN in that cell. Or perhaps, fix them only on the section of membrane that contains the polarised PIN. Even with the slightly arbitrary decisions on the length of time until the PINs become fixed, testing differentiation would make an interesting piece of future work.

8.2.2 Comparison with Other Canalisation Models

The diligent worker hypothesis as tested in the two simulators compares very well with other simulations of auxin transport canalisation. As the simulations are agent-based the hypothesis can be implemented at a fairly low mechanistic level. This is an attractive feature of agent-based models. Most mathematical

models abstract a lot of mechanism into a small number of equations. An important part of the canalisation models is PIN allocation. This is normally done by measuring the net flux of auxin out of the a cell across the membrane which is split into a number of different regions. The membrane region with the highest net flux gets the most PIN. There are a few differences as to whether there is a fixed pool of PINs for which the membrane regions compete, or if there is an unlimited supply with PINs allocated as a function of flux[5, 22, 73]. This works well as the membrane regions near sinks experience the highest net flux and therefore accumulate PIN. To date however, no flux sensor has been discovered, and flux sensors are considered to be biologically less plausible than measuring concentration.

The mechanism described in the Diligent Worker Hypothesis is able to produce auxin transport canals through a tissue without requiring a flux sensor for auxin. Instead the PINs react to the auxin concentration of their immediate environment, and whether there is auxin in their environment for them to transport. The results show that the interaction between PIN and the auxin in the apoplast is very important. Without this part of the hypothesis auxin canals do not form (see figure 4.8 page 66). Having the mechanism implemented helps to make it clear what is happening, and more information can be gathered about what parts of the mechanism are important. These models are still at a high level of abstraction, and the mechanism simulated is not meant to be the full biological mechanism. An attempt was made to find a level of abstraction where more information about a possible mechanism can be sought, without producing a simulation that was difficult to understand.

Bayer et al.[5] present a hybrid model of phyllotaxis and midvein formation. The model switches between an up-the-gradient model of PIN localisation at low auxin concentrations, and switches to a with-the-flux model at high auxin concentrations. The most relevant part of this model to this work is the development of the midvein, formed from a region of high auxin concentration to a sink using the with-the-flux model. In order to get this to work reliably the sink produces a signal that helps the canal find its location. The Diligent Worker hypothesis does not require any additional signals from the sink in order to link the auxin source to the sink. This removes the need for sinks to produce a specific signal for canalisation to work. The model presented by Bayer et al.[5] was designed to work for phyllotaxis and canalisation, and therefore is likely subject to additional constraints. However there have been theories about whether or not auxin sinks produce signals to aid canalisa-

tion, the work presented in this thesis, along with other canalisation models, supports the theory that this is not necessary.

There are aspects to the Diligent Worker Hypothesis that will need biological verification. Firstly it is not know if PINs can directly respond to auxin concentrations, or if a signalling protein will be required for this part of the mechanism. It is feasible that a transporter might respond to being unable to release the molecule it is transporting, or it could detect if it is bringing in as much as it is exporting, through conformational changes to the protein's shape. Good candidates for self-regulating transport proteins that undergo conformational changes due to their substrate have been identified in animals[67]. It is also not know if PIN could respond to not transporting auxin out of the cell. It is again plausible that transporting auxin changes the conformation of PIN making it less likely to dissociate from the membrane. One aspect of the Diligent Worker Hypothesis that needs to be tested is how it responds to large amount of exogenous auxin being added to the tissue. In biological experiments the application of exogenous auxin increases polar localisation of PIN. If large amounts of the applied auxin gets into the apoplast under the Diligent Worker Hypothesis it would cause PIN to dissociate from the membrane, reducing the amount of polarly localised PIN. If however most of the applied auxin ends up in cells under the Diligent Worker hypothesis this would cause more PIN to be produced in the cells and possibly (requires testing) more polar localisation of PIN. Unfortunately it is difficult to measure where applied auxin goes in a tissue but it would be interesting to test the behaviour of the Diligent Worker Hypothesis' response to auxin application experiments.

Resent work as highlighted the protein ABP1 (AUXIN-BINDING PROTEIN 1) as a possible signalling molecule for the localisation of PIN on the membrane. Robert et al. have shown that auxin signalling to the auxin receptor ABP1 inhibits the internalisation of PIN proteins. ABP1 is normally enhances endocytosis of PIN1 from the membrane back into the cytoplasm of the cell. However if auxin is bound to the ABP1 protein this prevents the enhancement of endocytosis and increases the amount of membrane bound PIN. Therefore as the auxin concentration of the cell increases the amount of membrane bound PIN also increases[98]. This mechanism has been modelled and is able to produce auxin transport canals in a variety of situations[117]. This model includes ABP1 in the apoplast of the plant tissue. Once bound with auxin the ABP1 attaches to the membrane of the cell and reduces the endocytosis of PIN1 proteins at that local. Whether or not ABP1 functions

in this way in the apoplast is unknown, but the model does provide interesting new predictions that could be experimentally validated. This mechanism could be implemented in this modelling framework allowing the results from the two different models to be compared.

8.2.3 Vacuole Size

Vacuole size does have a subtle effect on the ability of the Diligent Worker Hypothesis to develop canals. Figure 6.10 shows that in long cells, if the direction of auxin transport is across the long cell the PIN tends to localise closer to the corners of the cell membrane rather than in the middle. This is due to the PINs being affected by depletion of auxin in the cytoplasm close to middle of the membrane. Vacuoles are not always included in other published models of auxin transport canalisation, and there have been no published accounts of mechanisms for PIN localisation being affected by the presence of vacuoles. It is difficult to say if this is because affects have not be looked for, or if they were tested and the results not reported.

8.2.4 Cell Shape

The diligent worker hypothesis has been tested in cells of more natural shape. Figures 6.7 and 6.8 show that it is able to produce canals in tissues with irregular cells. Some minor changes were made to the parameters to take account of cell membrane areas in irregular cells having a reduced number of apoplastic neighbours. Observations suggest that this modification did close the qualitative gap between the speed at which the PINs polarise in irregular cells when compared with regular cells.

The second generation model that can simulate irregular cells has not been tested in as much depth as the first generation model. More experimental runs need to be done with irregular cells, matching more of the examples done with the first generation model shown in chapter 4.3. Testing in more natural cells is important, not only to ensure that the model can reproduce biological observations but also because other auxin transport models implemented using ODE are tested in more natural cells[70, 5, 45]. Models like those produced by Bayer et al.[5] have started to use templates in the simulations and therefore it is helpful to be able to use templates in the simulators to make comparisons.

8.2.5 Localisation by Cytoplasmic Auxin Gradients

The Internal Auxin Hypothesis is very similar to the Diligent Worker Hypothesis. The difference is that instead of the probability of PINs dissociating from the membrane as the apoplastic auxin concentration increases, they respond to increasing cytoplasmic auxin instead. Using internal auxin gradients in cells to regulate the localisation of PIN was not as successful as the Diligent Worker Hypothesis. Figures 4.23, 4.24 and 4.25 show that the hypothesis can successfully polarise the PIN in cells. Looking closely at the PIN in the cells highlights a difference between the two hypotheses. Looking at the source cell in figure 4.23 most of the PIN is on the left side of the cell membrane. However there is some along the top, and some down the side too. There are a number of patches of PIN, but overall the cell is polarised.

The internal auxin hypothesis works in two parts. Groups of PIN that are on the membrane together deplete the cytoplasm in their neighbourhood of auxin and therefore increase the likelihood that they will remain in that position on the membrane. This is not enough to polarise a cell; the PIN needs to become localised to the same area of membrane. The theory is that auxin travelling into the cell from the outside will increase the likelihood of PIN dissociating from the membrane. The patches should eventually develop where the amount of auxin coming back into the cell is low (in the vicinity of a sink). However auxin transport canals do not readily form, suggesting that the patches of PIN are not being influenced by auxin coming back into the cell enough for polarisation to occur towards a sink. The PIN becomes fixed on the membrane of a cell and then does not change. Increasing the sensitivity to cytoplasmic auxin would help the patches find a sink by reacting to auxin crossing into the cell, but simultaneously might stop them forming in the first place.

When canals did form they tended to follow the edge of the tissue (figures 4.23, 4.24 and 4.25). This could be due to auxin building up on the edge of the space and crossing back into the cells. There could be enough crossing back into the cell to cause PIN to build up on other membranes. This extra environmental influence might help the canals to form. This hypothesis needs further testing. It could be possible to find parameters that produce more stable canals and the experiments so far cannot rule it out as a possibility.

There is one other published model of auxin canalisation driven by measuring internal auxin gradients[52]. In this model PIN is allocated to the cell membrane at the low concentration end of a cytoplasmic auxin gradient. The

gradient has to be larger than 1% for PIN to start being recruited to the membrane. The membrane at the lowest part of the cytoplasmic gradient received the most PIN. PIN is constantly removed from all membranes. This is significantly different from the hypothesis presented here, as the PIN is trafficked to the membrane area experiencing the lowest part of the cytoplasmic gradient. The experimental set-up between the models is significantly different. The sink in [52] is down one side of the tissue, much larger than the single cell used here. The auxin source was also much larger. In order to test fully the hypothesis described by Kramer[52] it would have to be the association of PIN with membrane that was affected by cytoplasmic auxin concentration in the area. This could easily be tested in the future.

8.3 Simulation Run Time

The simulations are about on the limit of what is possible with current computer technology and programming techniques. When using a 2.4GHz Intel Processor it takes about 12hrs. to run a canalisation experiment on a grid of 8x6 cells. More complex experiments with more complex space, like blockages, take longer. It is also difficult to predict how much longer, it could be twice or three times. The stochastic nature of the simulations allows them to get off to a bad start, resulting in significant variation between experiments on identical tissue set-ups. It is possible that some of the goals of the work are made impossible by long simulation time. For example, the 3D model is limited to testing very small groups of cells.

The long simulation times is not because the models are badly implemented, but more because modelling hundreds of thousands of small auxin molecules, and tens of thousands of proteins, is very computationally intensive. The present method of multithreading the simulation has also reached its limit. Under the current implementation the threads block each other once more than 6 threads are used, this blocking starts to reduce the overall performance. Future technologies and programming techniques that are emerging, like General-Purpose computation on Graphics Processing Unit (GPGPU), could be used to speed up the computational process. Compared with well optimised mathematical models, the simulations require much more computational time. However the simulations can provide additional insight by implementing mechanisms at a lower level. The insight gained at the lower levels can help with determining what it is safe to abstract away in other models

built with different techniques.

8.4 Simulator Interfaces

Originally the intention was to use a Little Language[8] to provide an interface to the set-up of the models. Currently some set-up can be changed using a text file, but this is limited to some control of the layout of cells in the space, the size of vacuoles and control over sources and sinks. The original idea of using a Little Language could be revisited to allow easier access to some of the deeper parts of the simulations, for example, to allow the addition of new proteins and their behaviours without the need to delve into the Java code of the simulation. Of course, there is always the danger that the Little Language itself grows until it is of the complexity of Java. However, the intention is that it should be created in biological domain specific terms, not generic programming terms. An alternative to this would be to use UML as an interface into models for biologists. Biologists could be able to draw UML diagrams of new proteins (or other objects), associate them to other biological objects, and link them to implementation objects that allow them to function. The links with the biology would confer the biological behaviour, and the links with implementation would handle diffusion, positioning, I/O, etc. The models produced would be more general in their capabilities and allow for more hypotheses to be explored. The users could specify new proteins from a series of diagrams, which in collaboration with some implementation experts can be implemented into the simulator. This might be too ambitious, or perhaps even undesirable. It may prove much more productive, in terms of the longer term goals of the biological modelling, to maintain the collaboration between domain experts and modelling experts. Rather than building interfaces to simulations that could prove to be very difficult to maintain, or even reduce the flexibility of what can be modelled. There is always the risk with generic interfaces that they are used in ways that they were never intended to be, or tested for, risking models being produced that are not well understood.

8.5 Support Tools

Some effort was made to future proof the simulators. Both versions produce a datafile (calDat file) that stores very detailed information about the model at the time step which the file was saved. This file is designed to be persistent

and robust to updates or changes to the simulators. This was to try to allow for newer versions of the simulator to be compared with old ones, and allow any new tests to be run on past data. This works by separating the programs doing the analysis of the data from the simulator. The simulator just produces a file containing lots of raw information about where PINs and auxins are, and what the state of the cells are. The analysis programs then extract this data. If new types of space or new molecules are implemented these can be read by an updated version of the analysis program but the old version can still be read as it just ignores the missing bits of raw data. The raw data stored is rich enough that it is very likely to be possible to run a new analysis on the simulation without rerunning the experiment. This is very useful considering the time that might take. The simulations also periodically save checkpoints which can be used to restart a simulation from a previous state.

The use of Genetic Algorithms for searching parameter space was not as successful as it perhaps could have been. The combination of the difficulties of developing a fitness function that could not "cheat", and the long time required to complete a single generation, stopped an optimal set of parameters from being found. The process was further hampered by results being lost due to the computer cluster introducing some additional (unwanted) stochastic death. Good parameters were found but a systematic search through the space would probably have worked as well, especially as some hand tuning was required anyway.

8.6 Other Areas of Future Work

There are numerous possibilities for future work in addition to those already identified. The models have been successful in producing auxin transport canals but there is scope for additional work in this area. The models are also flexible and could be put to use in other areas.

8.6.1 Experiments with different arrangements of Sources and Sinks

It would now be interesting to test tissues with multiple auxin sources and sinks. Sachs[103] showed that there is interaction between existing auxin transport canals and new sources of auxin. If the auxin transport canal is transporting auxin a canal from the new source will not join the existing canal.

However, if the existing canal is not transporting auxin it behaves like an auxin sink, and a new canal can form from the new source to the existing canal. The ability to have multiple auxin sources could also be used to see if the auxin transport switch described by Prusinkiewicz et al[90] (successfully used to model bud activation) can be tested in the context of auxin transport canalisation and PIN localisation. The paper describes a mechanism where buds activate and are able to grow only when they can export auxin into the main stem of the plant. Using the one of the simulators presented, an auxin canal could be established in a tissue representing a main stem with a bud. Various experiments of source activation could then be tested. An ideal result would be that if there is an auxin canal present in the tissue that is transporting enough auxin so that it does not behave as a strong sink for auxin from the surrounding tissue an auxin transport canal would not form from the source of auxin in the bud. However if the auxin source in the main stem is shut down, the auxin levels would drop allowing the stem to behave as a sink. Hopefully an auxin transport canal would then form from the bud to the canal in the main stem. This would be an auxin transport switch partly driven by PIN localisation.

It would also be interesting to add a source to a tissue maintaining a background level of auxin where emergent canals are developing (like those shown in section 4.3.1). The addition of a source in the tissue might allow for an emergent canal to become stable within the tissue. How a canal develops from the new source would be of interest; would it cause a new canal to form in a similar manner to existing experiments? Or would it join to any existing emergent canal in the tissue at the time the source is activated? Using the application SourceSink (described in appendix B.4) new sources and sinks can be added to a tissue while the simulation is running. This will allow for this type of experiment to be tried in the future.

8.6.2 Experiments with Single Cells

The existing simulators could be used to look at auxin dynamics in a single cell or small group of cells in more detail. This could used to test localisation mechanism for PIN in more detail. They could also be used to investigate auxin homoeostasis in the cell and the subcellular compartments. There is evidence that auxin is transported into the Endoplasmic Reticulum by PIN5, another member of the PIN family of auxin transporters[77]. Coupling the production of different members of the PIN family with auxin concentration

in cells and subcellular compartments would be very interesting as long is it could be tied to a specific biological question. The 3D model might prove useful for modelling single cells.

8.6.3 Simulator Extensions

In the future we may wish to include more aspects of the Domain in the models and simulation. One important example is growth. Introducing growth into the current simulation architecture would be very difficult to do. Introducing growth in an agent-based model of this type would pose some difficult challenges. One difficult problem is what would be put in new areas of space, such as new cytoplasmic volume. Grieneisen et al.[39] describe a model of auxin transport in the growing root. The models are modified cellular Potts models and the new areas of space are filled with content such as auxin based on averaging out the auxin of the surrounding space. This is very difficult to do in a discrete system. The CoSMoS process could be used to make the transition between the current simulator to a new one in a way that allows us to fully understand the differences between the two simulators produced.

8.7 Concluding Remarks

The work presented here has shown that is it possible to model auxin transport canalisation with an executable agent-based simulator. Two mechanisms for PIN localisation have been tested. The Diligent Worker hypothesis has been shown to be successful. Implemented as rules describing the behaviour of PIN in the different environments of the cell in response to auxin concentration, it is able to polarly localise PIN in cells and also produce auxin transport canals from an auxin source to a sink elsewhere in the tissue. It does not require a flux sensor to be implemented or any additional long distance signals. The second hypothesis tested uses internal auxin gradients. This hypothesis was less successful and although it is able to consistently polarise cells, canals do not reliably form between the source of auxin and the skin, and the results display what look like edge effects. Again this hypothesis is implemented as rules effecting the PIN proteins. This hypothesis has not be as extensively tested and it is possible that its behaviour could be improved with more work.

Both generations of the simulation software are flexible and future hypotheses for PIN localisation could be tested. It would also be possible to look into the existing hypotheses in more detail. There remains more work to be done

on the second generation model, and the 3D model. Some of the features of the simulators have not been fully exploited either, being able to switch sources and sinks off could be used to conduct additional experiments. The simulators could also be used for related but different uses. For example auxin homoeostasis and biosyntheses could be investigated in more detail.

The CoSMoS process has proved to be an extremely useful tool in assisting the development of the simulators. It promotes a systematic approach to the development of the simulators, through the production of intermediate models describing the different stages of the development process. Thus increasing the likelihood that what is included, or not included, in the model is known and understood. The intermediate stages also help keep the domain experts on board throughout the entire modelling process. Following the CoSMoS process also highlighted the point where it was time to leave the first generation model and start development of the second generation model. The use of UML in the development process, which was constantly updated as the CoSMoS lifecycle was followed, also highlighted when the existing simulator code was becoming difficult to maintain and not living up to the initial desires of the project, of having well implemented, object-orientated code, that was easy to understand. The CoSMoS process, again with UML, was then used to develop the enhanced model. The CoSMoS process proved very useful for developing well implemented simulations.

Appendix A

Supplementary Data

A.1 Starting Parameters Used for GA

Table A.1 shows the starting parameters used with the GA to investigate some of the Diligent Worker Hypothesis parameter values in detail.

Diligent Worker GA Starting Parameters		
Parameter Name	Value	Description
PIN Production Coefficient (PPC)	30	Controls speed of PIN agent concentration increase with respect to increasing agent auxin.
PIN HC (PHC)	0.29	Concentration of auxin agents at which half the possible concentration of PIN agents is attained.
μ	0.015	Maximum Concentration of PIN agents.
External Auxin PIN HD (EPHD)	9.0	Controls the number of auxin agents in the neighbouring areas that produce a 50% chance of PIN disassociation.
External Auxin PIN Coefficient (EPC)	18	Controls steepness in probability of PIN dropping of the membrane in response to increasing apoplastic auxin.
PIN Dis. Const.	1×10^{-5}	Probability of PIN dropping off the membrane if it did not export auxin in the last time step.
PIN Degrade	1×10^{-5}	Probability of Cytoplasmic PIN degrading per time step when in the cytoplasm.
PIN Insertion	0.7	Probability of PIN attaching to a membrane on contact.

Table A.1: Diligent Worker Parameters used at the start of the search carried out with the Genetic Algorithm

Internal Auxin GA Starting Parameters		
Parameter Name	Value	Description
PIN Production Coefficient (PPC)	15	Controls speed of PIN agent concentration increase with respect to increasing agent auxin.
PIN HC (PHC)	0.15	Concentration of auxin agents at which half the possible concentration of PIN agents is attained.
μ	0.016	Maximum Concentration of PIN agents.
External Auxin PIN HD (EPHD))	9.5	Controls the number of auxin agents in the neighbouring areas that produce a 50% chance of PIN disassociation.
External Auxin PIN Coefficient (EPC)	35	Controls steepness in probability of PIN dropping of the membrane in response to increasing cytoplasmic auxin.
PIN Dis. Const.	1×10^{-9}	Probability of PIN dropping off the membrane if it did not export auxin in the last time step.
PIN Degrade	1×10^{-3}	Probability of Cytoplasmic PIN degrading per time step when in the cytoplasm.
PIN Insertion	0.7	Probability of PIN attaching to a membrane on contact.

Table A.2: Starting parameters used in for Internal Auxin hypothesis parameter search.

A.2 Starting Parameters Used for Internal Auxin GA

Table A.2 shows the starting parameters used with the GA to investigate some of the Internal Auxin Hypothesis parameter values in detail.

Appendix B

Accessory Programs

A fairly large number of accessory programs were developed for data analysis. Varying in complexity from sophisticated programs to simple scripts. Where the programs are either important tools or are at a mature enough point in there development they have been included on the supplementary DVD. Others will be released at a later point.

B.1 Simulation Check pointing

All simulations produce checkpoints at regular intervals while running. These allow the simulators to be restarted at the point at which the file was saved. Details on how to use checkpoints is in the ReadMe.txt files with the simulators.

B.2 Mining Data from the Persistent Data File

DataMining application. A significant amount of effort was put into developing a data file for the storage of results from the simulation runs. This compressed data file, called calDat, was designed to store as much information about the simulation state as possible at regular intervals during the running of the simulation. The file produced is design to be persistent, and therefore allow old simulator runs to be compared with new. It should also allow any new analysis done on new data to be run on old data. This is achieved by saving as much information about the positions of agents and the layout of the tissue into the calDat file as possible. The analysis is then the job of the DataMining program. If a new piece of information is required in the calDat file – for example a new type of agent – the system still works as the missing data in the old calDat is ignored.

All of the images produced of canals, and all the videos, were generated by DataMining.jar. Generating videos particularly requires the processing of a very large number of files. So DataMining is written with the Java Communicating Sequential Processes multithreading toolkit[87] and its performance scales linearly with the number of available processors in the computer. Originally developed by Francesca Day during a summer project at the University of York but later significantly altered DataMining is on the supplementary DVD.

B.3 Genetic Algorithm Software: ChromoServe

To carry out the parameter search using the Genetic Algorithm ChromoServe was developed in Java. ChromoServe is able to generate a starting population of Chromosomes each encoding a set of parameters to be tested. The starting population can either be generated randomly, from a starting chromosome which is then mutated to make the entire population, or from a text file listing all the chromosomes. ChromoServe is then able to submit the required number of jobs to a Sun Grid Engine (SGE) managed compute cluster via the Java Distributed Resource Management Application API (DRMAA). Or to a waiting population of simulations running on a computer.

Using DRMAA ChromoServe is able to keep track of the progress of SGE jobs. Once all the simulations in a generation are complete ChromoServe is able to generate the new generation of Chromosomes using the algorithm described in chapter 4 section 4.1.1. This generation is submitted automatically using DRMAA. The communication between ChromoServe and the simulations running on the SGE grid or remote computers is done using the Remote Method Invocation (RMI) libraries distributed with Java.

B.3.1 Analysis of Rough Cluster Data

Due to occasional problems with simulations on the cluster failing to complete a simple set of scripts were developed to analysis the log files produced from the simulations running on the cluster. Although as the runs are not complete useful information can sometimes still be gained from from it.

B.3.2 ChromoServe2

ChromoServeV2 was developed to address a number of short comings with ChromoServe. These include; the ability to run without the GUI. SGE jobs that have failed, or are suspected of failure, are resubmitted automatically. The GA being used is pluggable, it is easy to implement other algorithms. This version of the ChromoServe is not on the DVD as it currently requires careful use. The intention it to release this software properly to the community at a later date, and will be the only version of ChromoServe released.

B.4 Run time alteration of Source and Sink Cells: SourceSink

The ability to activate a new source or sink in a running simulation will be required for future work. As simulations are often run on remote computers or clusters have a GUI built into the simulator for changing the sources and sinks was a good solution. Therefore SourceSink was developed. Using remote method invocation (RMI) running simulations can be contacted, both over a computer network or on a local machine, and the source/sink status of all cells can be changed. SourceSink is available on the supplementary DVD.

B.5 Tools for Template Generation

A number of Java programs and some scripts have been made for generating templates for the second generation simulator. These include generating templates automatically from images of tissue sections and also the ability to generate the templates to build irregular 3D cells. These tools were developed by Francesca Day during a summer project at the University of York and are on the supplementary DVD.

B.6 Tools for producing time series data

The first generation simulator can be put into a mode for generating time series data of PIN, auxin and AUX/LAX concentration. In this mode the simulator saves a text file containing the concentration of the agents every 150 time steps. The Java application TimeSeries can read these files and produce a single text

file for each cell in the tissue containing the changing agent concentrations for the duration of the run. TimeSeries is available on the supplementary DVD.

Bibliography

[1] Paul S. Andrews, Fiona A. C. Polack, Adam T. Sampson, Susan Stepney, and Jon Timmis. The CoSMoS process version 0.1: A process for the modelling and simulation of complex systems. Technical Report YCS-2010-453, Department of Computer Science, University of York, 2010.

[2] Paul S. Andrews, Adam T. Sampson, Fiona Polack, Susan Stepney, and Jon Timmis. CoSMoS development lifecycle, version 0. Technical report, University of York, 2008. (in preparation).

[3] Richard Clayton Arno Steinacher, Ottoline Leyser. Computational modelling of auxin/pH dynamics on a single plant cell level. In preperation, February 2011.

[4] A. Bandyopadhyay, J. J. Blakeslee, O. R. Lee, J. Mravec, M. Sauer, B. Titapiwatanakun, S. N. Makam, R. Bouchard, M. Geisler, E. Martinoia, J. Friml, W. A. Peer, and A. S. Murphy. Interactions of PIN and PGP auxin transport mechanisms. *Biochem Soc Trans*, 35(Pt 1):137–141, Feb 2007.

[5] Emmanuelle M Bayer, Richard S Smith, Therese Mandel, Naomi Nakayama, Michael Sauer, Przemyslaw Prusinkiewicz, and Cris Kuhlemeier. Integration of transport-based models for phyllotaxis and midvein formation. *Genes Dev*, 23(3):373–384, Feb 2009.

[6] René Benjamins and Ben Scheres. Auxin: the looping star in plant development. *Annu Rev Plant Biol*, 59:443–465, 2008.

[7] Tom Bennett, Tobias Sieberer, Barbara Willett, Jon Booker, Christian Luschnig, and Ottoline Leyser. The *Arabidopsis* MAX pathway controls shoot branching by regulating auxin transport. *Curr Biol*, 16(6):553–563, Mar 2006.

[8] J. Bentley. Programming pearls. *Communications of the ACM*, 29(8):711–721, 1986.

[9] Ikram Blilou, Jian Xu, Marjolein Wildwater, Viola Willemsen, Ivan Paponov, Jirí Friml, Renze Heidstra, Mitsuhiro Aida, Klaus Palme, and Ben Scheres. The PIN auxin efflux facilitator network controls growth and patterning in arabidopsis roots. *Nature*, 433(7021):39–44, Jan 2005.

[10] Jonathan Booker, Steven Chatfield, and Ottoline Leyser. Auxin acts in xylem-associated or medullary cells to mediate apical dominance. *Plant Cell*, 15(2):495–507, Feb 2003.

[11] John W Chandler. Local auxin production: a small contribution to a big field. *Bioessays*, 31(1):60–70, Jan 2009.

[12] Peter Checkland and Jim Scholes. *Soft Systems Methodology in Action.* Wiley, 1990.

[13] Morris Cline. Apical dominance. *The Botanical Review*, 57(4):318–358, October 1991.

[14] Enrico Coen, Anne-Gaëlle Rolland-Lagan, Mark Matthews, J. Andrew Bangham, and Przemyslaw Prusinkiewicz. The genetics of geometry. *Proceedings of the National Academy of Sciences*, 101(14):4728–4735, 2004.

[15] Pierre Barbier de Reuille, Isabelle Bohn-Courseau, Karin Ljung, Halima Morin, Nicola Carraro, Christophe Godin, and Jan Traas. Computer simulations reveal properties of the cell-cell signaling network at the shoot apex in *Arabidopsis*. *Proc Natl Acad Sci U S A*, 103(5):1627–1632, Jan 2006.

[16] Carolin Delker, Anja Raschke, and Marcel Quint. Auxin dynamics: the dazzling complexity of a small molecule's message. *Planta*, 227(5):929–941, Apr 2008.

[17] Pankaj Dhonukshe, Fernando Aniento, Inhwan Hwang, David G Robinson, Jozef Mravec, York-Dieter Stierhof, and Jirí Friml. Clathrin-mediated constitutive endocytosis of PIN auxin efflux carriers in *Arabidopsis* . *Curr Biol*, 17(6):520–527, Mar 2007.

[18] Pankaj Dhonukshe, Hirokazu Tanaka, Tatsuaki Goh, Kazuo Ebine, Ari Pekka Mähönen, Kalika Prasad, Ikram Blilou, Niko Geldner, Jian Xu, Tomohiro Uemura, Joanne Chory, Takashi Ueda, Akihiko Nakano, Ben Scheres, and Jirí Friml. Generation of cell polarity in plants links endocytosis, auxin distribution and cell fate decisions. *Nature*, 456(7224):962–966, Dec 2008.

[19] S Douady and Y Couder. Phyllotaxis as a physical self-organized growth process. *Physical Review Letters*, 68(13):2098–2101, 1992.

[20] S. Efroni, D. Harel, and I. R. Cohen. Toward rigorous comprehension of biological complexity: modeling, execution, and visualization of thymic T-cell maturation. *Genome. Res.*, 13(11):2485–97, jul 2003.

[21] Elena Feraru and Jirí Friml. PIN polar targeting. *Plant Physiol*, 147(4):1553–1559, Aug 2008.

[22] Francois G Feugier, A. Mochizuki, and Y. Iwasa. Self-organization of the vascular system in plant leaves: inter-dependent dynamics of auxin flux and carrier proteins. *J Theor Biol*, 236(4):366–375, Oct 2005.

[23] A. Finney and M. Hucka. Systems biology markup language: Level 2 and beyond. *Biochem Soc Trans*, 31(Pt 6):1472–1473, Dec 2003.

[24] Jasmin Fisher and Thomas A. Henzinger. Executable cell biology. *Nat Biotechnol*, 25(11):1239–1249, Nov 2007.

[25] Jasmin Fisher, Nir Piterman, Alex Hajnal, and Thomas A. Henzinger. Predictive modeling of signaling crosstalk during c. elegans vulval development. *PLoS Comput Biol*, 3(5):e92, May 2007.

[26] Jasmin Fisher, Nir Piterman, E. Jane Albert Hubbard, Michael J. Stern, and David Harel. Computational insights into *Caenorhabditis elegans* vulval development. *Proc Natl Acad Sci U S A*, 102(6):1951–1956, Feb 2005.

[27] M. Fowler. *UML Distilled*. Addison-Wesley, 2nd edition, Jan. 1999.

[28] Jirí Friml. Subcellular trafficking of PIN auxin efflux carriers in auxin transport. *Eur J Cell Biol*, 89(2-3):231–235, 2010.

[29] Jirí Friml, Anne Vieten, Michael Sauer, Dolf Weijers, Heinz Schwarz, Thorsten Hamann, Remko Offringa, and Gerd Jürgens. Efflux-dependent

auxin gradients establish the apical-basal axis of *Arabidopsis* . *Nature*, 426(6963):147–153, Nov 2003.

[30] Jirí Friml, Xiong Yang, Marta Michniewicz, Dolf Weijers, Ab Quint, Olaf Tietz, René Benjamins, Pieter B F Ouwerkerk, Karin Ljung, Göran Sandberg, Paul J J Hooykaas, Klaus Palme, and Remko Offringa. A PINOID-dependent binary switch in apical-basal pin polar targeting directs auxin efflux. *Science*, 306(5697):862–865, Oct 2004.

[31] Philip Garnett, Arno Steinacher, Susan Stepney, Richard Clayton, and Ottoline Leyser. Computer simulation: the imaginary friend of auxin transport biology. *Bioessays*, 32(9):828–835, Sep 2010.

[32] Philip Garnett, Susan Stepney, and Ottoline Leyser. Integrative hybrid modelling of plant shoot branching. In *5th International Workshop on Functional Structural Plant Models: FSPM07, New Zealand*, pages 9.1–9.5, 2007.

[33] Philip Garnett, Susan Stepney, and Ottoline Leyser. Towards an executable model of auxin transport canalisation. In Peter Welch. Susan Stepney, Fiona Polack, editor, *CoSMoS 2008*, pages 63–91. Luniver Press, 2008.

[34] Markus Geisler and Angus S Murphy. The ABC of auxin transport: the role of p-glycoproteins in plant development. *FEBS Lett*, 580(4):1094–1102, Feb 2006.

[35] N. Geldner, J. Friml, Y. D. Stierhof, G. Jürgens, and K. Palme. Auxin transport inhibitors block PIN1 cycling and vesicle trafficking. *Nature*, 413(6854):425–428, Sep 2001.

[36] Niko Geldner, Nadine Anders, Hanno Wolters, Jutta Keicher, Wolfgang Kornberger, Philippe Muller, Alain Delbarre, Takashi Ueda, Akihiko Nakano, and Gerd Jürgens. The *Arabidopsis* GNOM ARF-GEF mediates endosomal recycling, auxin transport, and auxin-dependent plant growth. *Cell*, 112(2):219–230, Jan 2003.

[37] Victoria Gomez-Roldan, Soraya Fermas, Philip B Brewer, Virginie Puech-Pagès, Elizabeth A Dun, Jean-Paul Pillot, Fabien Letisse, Radoslava Matusova, Saida Danoun, Jean-Charles Portais, Harro Bouwmeester, Guillaume Bécard, Christine A Beveridge, Catherine

Rameau, and Soizic F Rochange. Strigolactone inhibition of shoot branching. *Nature*, 455(7210):189–194, Sep 2008.

[38] Paul P. Green. Expression of pattern in plants: combining molecular and calculus-based biophysical paradigms. *American Journal of Botany*, 86(8):1059–1076, 1999.

[39] Verônica A. Grieneisen, Jian Xu, Athanasius F. M. Marée, Paulien Hogeweg, and Ben Scheres. Auxin transport is sufficient to generate a maximum and gradient guiding root growth. *Nature*, 449(7165):1008–1013, Oct 2007.

[40] Inman Harvey. The microbial genetic algorithm. In *European Conference on Artifical Life*. In Press, 2009.

[41] Marcus G Heisler, Carolyn Ohno, Pradeep Das, Patrick Sieber, Gonehal V Reddy, Jeff A Long, and Elliot M Meyerowitz. Patterns of auxin transport and gene expression during primordium development revealed by live imaging of the *Arabidopsis* inflorescence meristem. *Curr Biol*, 15(21):1899–1911, Nov 2005.

[42] Wilhelm Hofmeister. *Allgemeine Morphologie der Gewächse*. Handbuch der physiologischen Botanik. Engelmann, Leipzig, 1868.

[43] M. Hucka, A. Finney, S. Hoops, S. Keating, and N. Novere. Systems biology markup language (SMBL) level 2: Structures and facilities for model definitions. *Nature Precedings*, 58(2), 2007.

[44] M. Hucka, A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano, A. P. Arkin, B. J. Bornstein, D. Bray, A. Cornish-Bowden, A. A. Cuellar, S. Dronov, E. D. Gilles, M. Ginkel, V. Gor, I. I. Goryanin, W. J. Hedley, T. C. Hodgman, J-H Hofmeyr, P. J. Hunter, N. S. Juty, J. L. Kasberger, A. Kremling, U. Kummer, N. Le Novère, L. M. Loew, D. Lucio, P. Mendes, E. Minch, E. D. Mjolsness, Y. Nakayama, M. R. Nelson, P. F. Nielsen, T. Sakurada, J. C. Schaff, B. E. Shapiro, T. S. Shimizu, H. D. Spence, J. Stelling, K. Takahashi, M. Tomita, J. Wagner, and J. Wang. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, Mar 2003.

[45] Henrik Jönsson, Marcus G Heisler, Bruce E Shapiro, Elliot M Meyerowitz, and Eric Mjolsness. An auxin-driven polarized transport

model for phyllotaxis. *Proc Natl Acad Sci U S A*, 103(5):1633–1638, Jan 2006.

[46] N. Kam, I. R. Cohen, and D. Harel. The immune system as a reactive system. In *Proc Visual Lang. and Formal Methods*. IEEE, 2001.

[47] Jürgen Kleine-Vehn and Jirí Friml. Polar targeting and endocytic recycling in auxin-dependent plant development. *Annu Rev Cell Dev Biol*, 24:447–473, 2008.

[48] Jürgen Kleine-Vehn, Johannes Leitner, Marta Zwiewka, Michael Sauer, Lindy Abas, Christian Luschnig, and Jirí Friml. Differential degradation of PIN2 auxin efflux carrier by retromer-dependent vacuolar targeting. *Proc Natl Acad Sci U S A*, 105(46):17812–17817, Nov 2008.

[49] Jürgen Kleine-Vehn, Pankaj Dhonukshe, Michael Sauer, Philip B Brewer, Justyna Wiśniewska, Tomasz Paciorek, Eva Benková, and Jirí Friml. ARF GEF-dependent transcytosis and polar delivery of PIN auxin carriers in *Arabidopsis* . *Curr Biol*, 18(7):526–531, Apr 2008.

[50] Eric M. Kramer. PIN and AUX/LAX proteins: their role in auxin accumulation. *Trends Plant Sci*, 9(12):578–582, Dec 2004.

[51] Eric M. Kramer. Computer models of auxin transport: a review and commentary. *J Exp Bot*, 59(1):45–53, Apr 2008.

[52] Eric M Kramer. Auxin-regulated cell polarity: an inside job? *Trends Plant Sci*, 14(5):242–247, May 2009.

[53] Eric M Kramer and Malcolm J Bennett. Auxin transport: a field in flux. *Trends Plant Sci*, 11(8):382–386, Aug 2006.

[54] Eric M Kramer, Nicholas L Frazer, and Tobias I Baskin. Measurement of diffusion within the cell wall in living roots of *Arabidopsis thaliana*. *J Exp Bot*, 58(11):3005–3015, 2007.

[55] C. Kuhlemeier and D. Reinhardt. Auxin and phyllotaxis. *Trends in Plant Science*, 6(5):187–189, 2001.

[56] Cris Kuhlemeier. Phyllotaxis. *Trends Plant Sci*, 12(4):143–150, Apr 2007.

[57] Cris Kuhlemeier, Przemyslaw Prusinkiewicz, and Richard S. Smith. Inhibition fields for phyllotactic pattern formation: a simulation study. *Canadian Journal of Botany*, 84:1635–1649(15), 1 November 2006.

[58] D KWIATKOWSKA. The relationships between the primary vascular system and phyllotactic patterns of anagallis-arvensis (primulaceae). *AMERICAN JOURNAL OF BOTANY*, 79(8):904–913, AUG 1992.

[59] Philip R. Larson. Development and organization of the primary vascular system in *Populus deltoides* according to phyllotaxy. *American Journal of Botany*, 62(10):1084–1099, 1975.

[60] A. C. Leopold and O. F. Hall. Mathematical model of polar auxin transport. *Plant Physiol*, 41(9):1476–1480, Nov 1966.

[61] Ottoline Leyser. The fall and rise of apical dominance. *Curr Opin Genet Dev*, 15(4):468–471, Aug 2005.

[62] Ottoline Leyser. The control of shoot branching: an example of plant information processing. *Plant Cell Environ*, 32(6):694–703, Jun 2009.

[63] Chun-Jian Li and Fritz Bangerth. Autoinhibition of indoleacetic acid transport in the shoots of two-branched pea (pisum sativum) plants and its relationship to correlative dominance. *Physiologia Plantarum*, 106(4):415–420, 1999.

[64] V. A. Likhoshvai, N. A. Omel'ianchuk, V. V. Mironova, S. I. Fadeev, E. D. Melsness, and N. A. Kolchanov. [mathematical model of auxin distribution in the plant root]. *Ontogenez*, 38(6):446–456, 2007.

[65] K. Ljung, R. P. Bhalerao, and G. Sandberg. Sites and homeostatic control of auxin biosynthesis in *Arabidopsis* during vegetative growth. *Plant J*, 28(4):465–474, Nov 2001.

[66] Karin Ljung, Anna K Hull, Mariusz Kowalczyk, Alan Marchant, John Celenza, Jerry D Cohen, and Göran Sandberg. Biosynthesis, conjugation, catabolism and homeostasis of indole-3-acetic acid in *Arabidopsis thaliana*. *Plant Mol Biol*, 49(3-4):249–272, 2002.

[67] Min Lu, Jin Chai, and Dax Fu. Structural basis for autoregulation of the zinc transporter YiiP. *Nat Struct Mol Biol*, 16(10):1063–1067, Oct 2009.

[68] Nenad Malenica, Lindy Abas, René Benjamins, Saeko Kitakura, Harald F Sigmund, Kim S Jun, Marie-Theres Hauser, Jirí Friml, and Christian Luschnig. Modulator of PIN genes control steady-state levels of *Arabidopsis* PIN proteins. *Plant J*, 51(4):537–550, Aug 2007.

[69] Stefano Mancuso, Anna Maria Marras, Volker Magnus, and Frantisek Baluska. Noninvasive and continuous recordings of auxin fluxes in intact root apex with a carbon nanotube-modified and self-referencing microelectrode. *Anal Biochem*, 341(2):344–351, Jun 2005.

[70] Roeland M H Merks, Yves Van de Peer, Dirk Inzé, and Gerrit T S Beemster. Canalization without flux sensors: a traveling-wave hypothesis. *Trends Plant Sci*, 12(9):384–390, Sep 2007.

[71] Stefan Hoops Sarah M. Keating Sven Sahle Darren J. Wilkinson Michael Hucka, Frank Bergmann. The systems biology markup language (SBML): Language specification for level 3 version 1 core. *http://sbml.org/Documents/Specifications*, 2010.

[72] Melanie Mitchell. *An Introduction to Genertic Algorithms*. MIT Press, 1998.

[73] G. Mitchison. A model for vein formation in higher plants. *Pro. R. Soc. Lond.*, 207:79–109, 1980.

[74] G. Mitchison. The polar transport of auxin and vein patterns in plants. *Phil. Trans. R. Soc. Lond.*, 295:461–471, 1981.

[75] GJ MITCHISON. The dynamics of auxin transport. *PROCEEDINGS OF THE ROYAL SOCIETY OF LONDON SERIES B-BIOLOGICAL SCIENCES*, 209(1177):489–511, 1980.

[76] D. Morris, J. Friml, and E. Zamimalova. Auxin transport. In P. Davies, editor, *Plant Hormones: Biosynthesis, Signal Transduction, Action!*, chapter E1, pages 437–470. Kluwer Academic Publishers, 2004.

[77] Jozef Mravec, Petr Skůpa, Aurélien Bailly, Klára Hoyerová, Pavel Krecek, Agnieszka Bielach, Jan Petrásek, Jing Zhang, Vassilena Gaykova, York-Dieter Stierhof, Petre I Dobrev, Katerina Schwarzerová, Jakub Rolcík, Daniela Seifertová, Christian Luschnig, Eva Benková, Eva Zazímalová, Markus Geisler, and Jirí Friml. Subcellular homeostasis of

phytohormone auxin is mediated by the ER-localized PIN5 transporter. *Nature*, 459(7250):1136–1140, Jun 2009.

[78] Ayako Nakamura, Kanako Higuchi, Hideki Goda, Makoto T Fujiwara, Shinichiro Sawa, Tomokazu Koshiba, Yukihisa Shimada, and Shigeo Yoshida. Brassinolide induces IAA5, IAA19, and DR5, a synthetic auxin response element in *Arabidopsis*, implying a cross talk point of brassinosteroid and auxin signaling. *Plant Physiol*, 133(4):1843–1853, Dec 2003.

[79] Alan C Newell, Patrick D Shipman, and Zhiying Sun. Phyllotaxis: cooperation and competition between mechanical and biochemical processes. *J Theor Biol*, 251(3):421–439, Apr 2008.

[80] J. Odell, H. Parunak, and B. Bauer. Extending UML for agents. In *AOIS Worshop at AAAI*, 2000.

[81] OMG. Maintainer of the UML standards. *http://www.omg.org*, 2008.

[82] Veronica Ongaro and Ottoline Leyser. Hormonal control of shoot branching. *J Exp Bot*, 59(1):67–74, 2008.

[83] Oracle. Developers of the Java programming language. *http://www.java.com/en/*, 2010.

[84] Oracle. Netbeans java IDE. *http://netbeans.org/*, 2010.

[85] Tomasz Paciorek, Eva Zazímalová, Nadia Ruthardt, Jan Petrásek, York-Dieter Stierhof, Jürgen Kleine-Vehn, David A. Morris, Neil Emans, Gerd Jürgens, Niko Geldner, and Jirí Friml. Auxin inhibits endocytosis and promotes its own efflux from cells. *Nature*, 435(7046):1251–1256, Jun 2005.

[86] Geraint Parry, Alan Marchant, Sean May, Ranjan Swarup, Kamal Swarup, Nick James, Neil Graham, Trudie Allen, Tony Martucci, Antony Yemm, Richard Napier, Ken Manning, Graham King, and Malcolm Bennett. Quick on the uptake: Characterization of a family of plant auxin influx carriers. *Journal of Plant Growth Regulation*, 20(3):217–225, September 2001.

[87] James Moores Kevin Chalmers Bernhard Sputh Peter Welch, Neil Brown. Integrating and extending JCSP. In *Communicating Process Architectures 2007*, pages 349–370, 2007.

[88] Jan Petrásek, Jozef Mravec, Rodolphe Bouchard, Joshua J Blakeslee, Melinda Abas, Daniela Seifertová, Justyna Wisniewska, Zerihun Tadele, Martin Kubes, Milada Covanová, Pankaj Dhonukshe, Petr Skupa, Eva Benková, Lucie Perry, Pavel Krecek, Ok Ran Lee, Gerald R Fink, Markus Geisler, Angus S Murphy, Christian Luschnig, Eva Zazímalová, and Jirí Friml. PIN proteins perform a rate-limiting function in cellular auxin efflux. *Science*, 312(5775):914–918, May 2006.

[89] T. K. Prasad, X. Li, A. M. Abdel-Rahman, Z. Hosokawa, N. P. Cloud, C. E. Lamotte, and M. G. Cline. Does Auxin Play a Role in the Release of Apical Dominance by Shoot Inversion in Ipomoea nil? *Ann Bot*, 71(3):223–229, 1993.

[90] Przemyslaw Prusinkiewicz, Scott Crawford, Richard S Smith, Karin Ljung, Tom Bennett, Veronica Ongaro, and Ottoline Leyser. Control of bud activation by an auxin transport switch. *Proc Natl Acad Sci U S A*, Sep 2009.

[91] Przemyslaw Prusinkiewicz and Anne-Gaëlle Rolland-Lagan. Modeling plant morphogenesis. *Curr Opin Plant Biol*, 9(1):83–88, Feb 2006.

[92] A. M. Rashotte, S. R. Brady, R. C. Reed, S. J. Ante, and G. K. Muday. Basipetal auxin transport is required for gravitropism in roots of *Arabidopsis*. *Plant Physiol*, 122(2):481–490, Feb 2000.

[93] J A Raven. Transport of indoleacetic acid in plant cells in relation to pH and electrical potential gradients, and its significance for polar IAA transport. *New Phytologist*, 74(2):163–172, 1975.

[94] Mark Read, Jon Timmis, Paul S. Andrews, and Vipin Kumar. A domain model of experimental autoimmune encephalomyelitis. In *CoSMoS 2009*, pages 9–44. Luniver Press, 2009.

[95] D. Reinhardt, T. Mandel, and C. Kuhlemeier. Auxin regulates the initiation and radial position of plant lateral organs. *Plant Cell*, 12(4):507–518, 2000.

[96] Didier Reinhardt. Regulation of phyllotaxis. *Int J Dev Biol*, 49(5-6):539–546, 2005.

[97] Didier Reinhardt, Eva-Rachele Pesce, Pia Stieger, Therese Mandel, Kurt Baltensperger, Malcolm Bennett, Jan Traas, Jirí Friml, and Cris Kuhlemeier. Regulation of phyllotaxis by polar auxin transport. *Nature*, 426(6964):255–260, Nov 2003.

[98] Stéphanie Robert, Jürgen Kleine-Vehn, Elke Barbez, Michael Sauer, Tomasz Paciorek, Pawel Baster, Steffen Vanneste, Jing Zhang, Sibu Simon, Milada Čovanová, Kenichiro Hayashi, Pankaj Dhonukshe, Zhenbiao Yang, Sebastian Y Bednarek, Alan M Jones, Christian Luschnig, Fernando Aniento, Eva Zažímalová, and Jirí Friml. ABP1 mediates auxin inhibition of clathrin-dependent endocytosis in *Arabidopsis*. *Cell*, 143(1):111–121, Oct 2010.

[99] P. H. Rubery and A. R. Sheldrake. Carrier-mediated auxin transport. *Planta*, 118(2):101–121, 1974.

[100] T. SACHS. The Role of the Root in the Induction of Xylem Differentiation in Peas. *Ann Bot*, 32(2):391–399, 1968.

[101] T. SACHS. Polarity and the Induction of Organized Vascular Tissues. *Ann Bot*, 33(2):263–275, 1969.

[102] T. SACHS. Patterned differentiation in plants. *Differentiation*, 11(1-3):65–73, 1978.

[103] T Sachs. The control of the patterned differentiation of vascular tissues. *Advances in Botanical Research Incorporating Advances in Plant Pathpology*, 9:151–262, Jan. 1981.

[104] Patrik Sahlin, Bo Söderberg, and Henrik Jönsson. Regulated transport as a mechanism for pattern generation: capabilities for phyllotaxis and beyond. *J Theor Biol*, 258(1):60–70, May 2009.

[105] Enrico Scarpella, Danielle Marcos, Jirí Friml, and Thomas Berleth. Control of leaf vascular patterning by polar auxin transport. *Genes Dev*, 20(8):1015–1027, Apr 2006.

[106] Enrico Scarpella and Annemarie H Meijer. Pattern formation in the vascular system of monocot and dicot plant species. *New Phytologist*, 164:202–242, 2004.

[107] S. J. Singer and G. L. Nicolson. The fluid mosaic model of the structure of cell membranes. *Science*, 175(23):720–731, Feb 1972.

[108] Richard S Smith and Emmanuelle M Bayer. Auxin transport-feedback models of patterning in plants. *Plant Cell Environ*, May 2009.

[109] Richard S Smith, Soazig Guyomarc'h, Therese Mandel, Didier Reinhardt, Cris Kuhlemeier, and Przemyslaw Prusinkiewicz. A plausible model of phyllotaxis. *Proc Natl Acad Sci U S A*, 103(5):1301–1306, Jan 2006.

[110] Ian Sommerville. *Software Engineering*. Pearson Education, 9 edition, Feb 2010.

[111] Szymon Stoma, Mikael Lucas, Jérôme Chopard, Marianne Schaedel, Jan Traas, and Christophe Godin. Flux-based transport enhancement as a plausible unifying mechanism for auxin transport in meristem development. *PLoS Comput Biol*, 4(10):e1000207, Oct 2008.

[112] Mina Tanaka, Kentaro Takei, Mikiko Kojima, Hitoshi Sakakibara, and Hitoshi Mori. Auxin controls local cytokinin biosynthesis in the nodal stem in apical dominance. *Plant J*, 45(6):1028–1036, Mar 2006.

[113] K. V. Thimann and F. Skoog. Studies on the growth hormone of plants: III. the inhibiting action of the growth substance on bud development. *Proc Natl Acad Sci U S A*, 19(7):714–716, Jul 1933.

[114] Mikihisa Umehara, Atsushi Hanada, Satoko Yoshida, Kohki Akiyama, Tomotsugu Arite, Noriko Takeda-Kamiya, Hiroshi Magome, Yuji Kamiya, Ken Shirasu, Koichi Yoneyama, Junko Kyozuka, and Shinjiro Yamaguchi. Inhibition of shoot branching by new terpenoid plant hormones. *Nature*, 455(7210):195–200, Sep 2008.

[115] Anne Vieten, Michael Sauer, Philip B Brewer, and Jirí Friml. Molecular and cellular aspects of auxin-transport-mediated development. *Trends Plant Sci*, 12(4):160–168, Apr 2007.

[116] Anne Vieten, Steffen Vanneste, Justyna Wisniewska, Eva Benková, René Benjamins, Tom Beeckman, Christian Luschnig, and Jirí Friml. Functional redundancy of PIN proteins is accompanied by auxin-dependent cross-regulation of PIN expression. *Development*, 132(20):4521–4531, Oct 2005.

[117] Krzysztof Wabnik, Jürgen Kleine-Vehn, Jozef Balla, Michael Sauer, Satoshi Naramoto, Vilém Reinöhl, Roeland M H Merks, Willy Govaerts, and Jirí Friml. Emergence of tissue polarization from synergy of intracellular and extracellular auxin signaling. *Mol Syst Biol*, 6:447, Dec 2010.

[118] K. Webb and T. White. UML as a cell and biochemistry modeling language. *BioSystems*, 80:283–302, 2005.

[119] FW Went. Polarity of auxin transport in inverted tagetes cuttings. *Botanical Gazette*, 103(2):386–390, 1941.

[120] Justyna Wisniewska, Jian Xu, Daniela Seifertová, Philip B Brewer, Kamil Ruzicka, Ikram Blilou, David Rouquié, Eva Benková, Ben Scheres, and Jirí Friml. Polar PIN localization directs auxin flow in plants. *Science*, 312(5775):883, May 2006.

[121] L. M. Young, M. L. Evans, and R. Hertel. Correlations between gravitropic curvature and auxin movement across gravistimulated roots of zea mays. *Plant Physiol*, 92:792–796, 1990.

[122] Yunde Zhao. The role of local biosynthesis of auxin and cytokinin in plant development. *Curr Opin Plant Biol*, 11(1):16–22, Feb 2008.