

This item was submitted to Loughborough University as a PhD thesis by the author and is made available in the Institutional Repository (<https://dspace.lboro.ac.uk/>) under the following Creative Commons Licence conditions.



For the full text of this licence, please go to:
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

Advances in Modelling of Epithelial to Mesenchymal Transition

by

Tariq Abdulla

A Doctoral Thesis

Submitted in partial fulfilment of the requirements for the award of PhD Electronic, Electrical and Systems Engineering of Loughborough University.

June 2013

© by Tariq Abdulla 2013



Acknowledgements

I have been extremely fortunate to be surrounded by positive and constructive people at every stage of my PhD. They have stimulated and supported me to do a better job. I would like to express my sincere gratitude to the following people (in no particular order):

Ron Summers

Jean-Marc Schleich

Roy Kalawsky

Ewart Carson

Jean-Louis Dilenseger

Lucile Houyel

Fanny Bajolle

Chenyi Yang

The team at Mendeley

Luis Luna-Zurita

José Luis de la Pompa

James Glazier

Randy Heiland

Maciej Swat

Jocelyn Griselle

Dylan Banks

Mohamed Elgendi

Mehul P. Sampat

Maxim Tankaria

Pedro Mendes

Daniel Cook

Herbert M. Sauro



List of Publications

Journal Papers

- [1] T. Abdulla, L. Luna-Zurita, J.L. de la Pompa, J.M. Schleich and R. Summers, "Epithelial to mesenchymal transition – the roles of cell morphology, labile adhesion and junctional coupling," *Computer Methods and Programs in Biomedicine*, 2013, vol 11, no. 2, pp. 435-446, 2013.
- [2] T. Abdulla, R. Imms, J.L. Dillenseger, J.M. Schleich, and R. Summers, "Computational modelling of epithelial to mesenchymal transition," *IRBM*, vol. 32, no. 5, pp. 306-310, Nov. 2011.
- [3] R. Summers, T. Abdulla, L. Houyel, and J.M. Schleich, "Progress with a Multiscale Systems Engineering Approach to Cardiac Development," *Automatika*, vol. 52, no. 1, pp. 49-57, 2011.
- [4] R. Summers, T. Abdulla, and J.M. Schleich, "Progress with Multiscale Systems," *Measurement + Control*, vol. 44, no. 6, pp. 180-185, 2011.

Peer Reviewed Conference Papers

- [1] R. Summers, T. Abdulla, and J.-M. Schleich, "Progress On Multiscale Representation Of Cardiac Valve Morphogenesis," in *XX IMEKO World Congress*, 2012.
- [2] T. Abdulla, J.M. Schleich, and R. Summers, "Hybrid Modelling of *in vitro* Epithelial to Mesenchymal Transition," in *VPH 2012*, 2012.
- [3] M. Tankaria, T. Abdulla, and R. Summers, "Multiscale Modelling of Delta-Notch Pathways," in *VPH 2012*, 2012.
- [4] T. Abdulla, J.M. Schleich, and R. Summers, "Multiscale Modelling of Notch-Mediated Lateral Induction," in *Proceedings of the IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI 2012)*, 2012, pp. 257-260.
- [5] T. Abdulla, R. A. Imms, J.M. Schleich, and R. Summers, "Towards Multiscale Systems Modeling of Endocardial to Mesenchymal Transition," in *33rd Annual International Conference of the IEEE EMBS*, 2011, pp. 449-452.
- [6] T. Abdulla, R. A. Imms, J.M. Schleich, and R. Summers, "Composite Annotation for Heart Development," in *International Conference on Biomedical Ontology*, 2011, pp. 47-54.
- [7] T. Abdulla, R. A. Imms, J.L. Dillenseger, J.M. Schleich, and R. Summers, "Computational Modelling of Epithelial to Mesenchymal Transition," in *Recherche en Imagerie et Technologies pour la Santé*, 2011, p. 582.
- [8] R. Summers, T. Abdulla, R. A. Imms, and J.M. Schleich, "3D Simulation of an *in vitro* Epithelial to Mesenchymal Transition," in *5th European IFMBE Conference, IFMBE Proceedings 37*, 2011, pp. 1287-1290.
- [9] R. Summers et al., "Multiscale systems modeling of the tetralogy of Fallot," in *32nd Annual International Conference of the IEEE EMBS*, 2010, pp. 752-5.
- [10] T. Abdulla, R. A. Imms, J.M. Schleich, and R. Summers, "Systems modelling of EMT cell signalling pathways in heart valve development," in *VPH 2010*, 2010, pp. 160-162.
- [11] R. Summers, T. Abdulla, L. Houyel, and J.M. Schleich, "Multiscale modelling of congenital heart disease," in *VPH 2010*, 2010.
- [12] T. Abdulla, R. Imms, J.M. Schleich, and R. Summers, "Multiscale information modelling for heart morphogenesis," in *Journal of Physics: Conference Series*, 2010, vol. 238, p. 012062.
- [13] J.M. Schleich, R. Summers, T. Abdulla, and R. A. Imms, "Multiscale modelling of the Tetralogy of Fallot," in *17th World Congress in Cardiac Electrophysiology*, 2010.

Magazine Articles

- [1] R. Summers, T. Abdulla, and J.M. Schleich, "Project Fallot: Multiscale Modelling of Heart Development," *VPH NoE News No.5*, pp. 24-25, 2011.

Abstract

Epithelial to Mesenchymal Transition (EMT) is a cellular transformation process that is employed repeatedly and ubiquitously during vertebrate morphogenesis to build complex tissues and organs. Cellular transformations that occur during cancer cell invasion are phenotypically similar to developmental EMT, and involve the same molecular signalling pathways. EMT processes are diverse, but are characterised by: a loss of cell-cell adhesion; a gain in cell-matrix adhesion; an increase in cell motility; the secretion of proteases that degrade basement membrane proteins; an increased resistance to apoptosis; a loss of polarisation; increased production of extracellular matrix components; a change from a rounded to a fibroblastic morphology; and an invasive phenotype.

This thesis focuses explicitly on endocardial EMT, which is the EMT that occurs during vertebrate embryonic heart development. The embryonic heart initially forms as a tube, with myocardium externally, endocardium internally, with these tissue layers separated by a thick extracellular matrix termed the cardiac jelly. Some of the endocardial cells in specific regions of the embryonic heart tube undergo EMT and invade the cardiac jelly. This causes cellularised swellings inside the embryonic heart tube termed the endocardial cushions. The emergence of the four chambered double pump heart of mammals involves a complex remodelling that the endocardial cushions play an active role in. Even while heart remodelling is taking place, the heart tube is operating as a single-circulation pump, and the endocardial cushions are performing a valve-like function that is critical to the survival of the embryo (Nomura-Kitabayashi et al. 2009). As the endocardial cushions grow and remodel, they become the valve leaflets of the foetal heart. The endocardial cushions also contribute tissue to the septa (walls) of the heart. Their correct formation is thus essential to the development of a fully functional, fully divided, double-pump system. It has been shown that genetic mutations that cause impaired endocardial EMT lead to the development of a range of congenital heart defects (Fischer et al. 2007).

An extensive review is conducted of existing experimental investigations into endocardial EMT. The information extracted from this review is used to develop a multiscale conceptual model of endocardial EMT, including the major protein signalling pathways involved, and the cellular phenotypes that they induce or inhibit. After considering the requirements for computational simulations of EMT, and reviewing the various techniques and simulation packages available

for multi-cell modelling, cellular Potts modelling is selected as having the most appropriate combination of features. The open source simulation platform CompuCell3D is selected for model development, due to the flexibility, range of features provided and an existing implementation of multiscale models; that include subcellular models of reaction pathways.

Based on the conceptual model of endocardial EMT, abstract computational simulations of key aspects are developed, in order to investigate qualitative behaviour under different simulated conditions. The abstract simulations include a 2D multiscale model of Notch signalling lateral induction, which is the mechanism by which the embryonic heart tube is patterned into cushion and non-cushion forming regions. Additionally, a 3D simulation is used to investigate the possible role of contact-inhibited mitosis, upregulated by the VEGF protein, in maintaining an epithelial phenotype.

One particular *in vitro* investigation of endocardial EMT (Luna-Zurita et al. 2010) is used to develop quantitative simulations. The quantitative data used for fitting the simulations consist of cell shape metrics that are derived from simple processing of the imaging results. Single cell simulations are used to investigate the relationship between cell motility and cell shape in the cellular Potts model. The findings are then implemented in multi-cell models, in order to investigate the relationship between cell-cell adhesion, cell-matrix adhesion, cell motility and cell shape during EMT.

Table of Contents

Acknowledgements.....	ii
List of Publications.....	iv
Abstract.....	v
List of Figures.....	x
List of Tables	xvii
List of Acronyms	xviii
1. Introduction.....	1
1.1. Background Context.....	1
1.1.1. Modelling and Simulation in Biomedicine	1
1.1.2. Epithelial to Mesenchymal Transition	3
1.2. Aims and Objectives.....	7
1.3. Plan of Thesis	8
2. Literature Review	9
2.1. Introduction	9
2.2. Heart Development.....	10
2.2.1. Congenital Heart Diseases	10
2.2.2. Anatomy of Heart Development	12
2.2.3. Epithelial to Mesenchymal Transition in Heart Development	16
2.2.4. Notch Signalling in Heart Development.....	21
2.3. Cell Level Modelling Techniques.....	29
2.4. Existing Simulations of EMT	34
2.4.1. 3D Cellular Potts Simulation of Endocardial Cushion Growth.....	34
2.4.2. 2D Cellular Potts Simulation of Active Cell Migration & Mechanical Equilibrium.....	36
2.4.3. Multiscale 3D Cell Centre Model of Cancer Cell Invasion	38
2.4.4. Multiscale 3D Cellular Potts Simulation Cancer Cell Growth and Invasion.....	40
2.4.5. 2D Cellular Potts Simulation of the Role of ECM in Glioma Invasion	42
2.4.6. Multiscale 3D Cell Centre Model of the Human Epidermis and Wound Healing	44
2.4.7. 2D Cellular Potts Model of Intercellular Adhesion and Cancer Invasion	46
2.5. Existing Simulations of EMT Pathways	46
2.5.1. Notch Signalling	46

2.5.2.	TGF- β and BMP Signalling	47
2.5.3.	Wnt/ β -catenin and E-cadherin Signalling.....	49
2.6.	Challenges for Multiscale Modelling in Biomedicine	49
2.6.1.	Annotation of Multiscale Models.....	52
2.7.	Existing Multiscale Simulation Platforms.....	54
2.7.1.	VPH/Physiome Project.....	54
2.7.2.	CompuCell3D	56
2.7.3.	Chaste.....	57
2.7.4.	FLAME	59
2.7.5.	SemGen	59
2.7.6.	MoBi and PK-Sim	60
2.8.	Summary.....	62
3.	Methods	63
3.1.	Introduction.....	63
3.2.	Method Selection	64
3.3.	Image Processing.....	67
3.4.	Cellular Potts Modelling.....	72
3.4.1.	3D Simulations.....	76
3.4.2.	2D Simulations.....	77
3.4.3.	Fitting Space and Time Scale of Simulations.....	79
3.4.4.	Cell Morphology Metrics	80
3.4.5.	Cell Migration Metrics	80
3.4.6.	CompuCell3D-Bionetsolver Implementation.....	81
3.5.	Verification and Validation	85
3.5.1.	Verification.....	86
3.5.2.	Validation	86
3.6.	Summary.....	87
4.	Results	88
4.1.	Introduction	88
4.2.	3D Simulation of <i>in vitro</i> EMT.....	89
4.3.	3D Simulation of Contact Inhibited Proliferation	94
4.4.	2D Simulations of Cell Morphology.....	98
4.4.1.	Image Processing Results	98

4.4.2. Single Cell Simulation Results.....	100
4.4.3. Multicell Simulation Results	102
4.4.4. Verification and Validation	108
4.5. Multiscale Modelling of Notch Mediated Lateral Induction	109
4.6. Composite Annotation for Heart Development.....	116
4.7. Summary.....	123
5. Discussion and Conclusions	124
5.1. Review of the Aims and Objectives	124
5.2. Contributions to Original Knowledge.....	125
5.3. Future Research Opportunities	128
5.3.1. Verification and Validation	130
5.3.2. Potential Model Applications.....	132
References.....	135
Appendix A.	152
Modelling and Simulation Methods in Further Detail	152

List of Figures

Figure 1.1 Schematic of the systems biology approach (Merks & Glazier 2005)	1
Figure 1.2 The cycle of epithelial cell plasticity (Thiery & Sleeman 2006)	4
Figure 1.3 Schematic of multiscale modelling applied to EMT in heart development. OFT: Outflow Tract, AVC: Atrioventricular Canal, VEGF: Vascular Endothelial Growth Factor	5
Figure 1.4 Structure of thesis.....	8
Figure 2.1 A malformation may originate from different mechanisms. Common arterial trunk (CAT), Tetralogy of Fallot (TOF) or any other disease resulting from abnormal remodelling of the outflow tract, may result from the participation of a defect secondary heart field and/or a migration defect of the neural crest cells, and/or a rotation defect of the myocardial outflow tract and/or a formation defect of the endocardial cushions (Bajolle et al. 2009). Key: TOF: Tetralogy of Fallot, TOF&PA: Tetralogy of Fallot and Pulmonary Atresia, IAA: Interrupted Aortic Arch, CAT: Common Arterial Trunk, VSD: Ventricular Septal Defect, DORV: Double Outlet Right Ventricle, TGA: Transposition of the Great Arteries.....	11
Figure 2.2 Heart looping and wedging between embryonic day 20 (E20) and E32, or Carnegie stages 9 to 13. After (Kirby 2007) and (Hill 2011).....	12
Figure 2.3 Detail of endocardial cushion growth and fusion during heart looping and wedging, after (Kirby 2007)	13
Figure 2.4 Modified Van Praagh diagram, showing rotation of the outflow tract, after (Donnelly & Higgins 1996). As the outflow tract septates into the pulmonary artery (P) and aorta (A) it rotates 150 degrees in normal development, with different degrees of rotation corresponding to different congenital heart diseases. Key: PTA: Persistent Truncus Arteriosus, TOF: Tetralogy of Fallot, DORV: Double Outlet Right Ventricle, d-TGA: dextro-Transposition of the Great Arteries, l-TGA: levo-Transposition of the Great Arteries	14
Figure 2.5 Illustration of human cardiac morphogenesis and the redistribution of tissues (Srivastava & Olson 2000)	17
Figure 2.6 a) TGF β and BMP signalling from the myocardium induce Snail expression in the endocardium, which inhibits expression of the endothelial adhesion molecule VE-Cadherin. Activated intracellular Notch has the same effect, and these factors combine to stimulate the endocardial cell to lose its adhesion and adopt a mesenchymal phenotype. b) The cushion forming regions are established by restricted areas of gene expression in both in the myocardium and endocardium.	18
Figure 2.7 Model for VEGF in control of heart valve endothelial cell proliferation (Armstrong & Bischoff 2004)	19

Figure 2.8 Major protein interactions during EMT in endocardial cushion growth. After (Luna-Zurita et al. 2010) and (Wagner & Siddiqui 2007). Gradients show expression pattern for Notch1 (blue, in the endocardium) and VEGF (green in the cardiac jelly and red in the myocardium).....	20
Figure 2.9 Basic operation of the Notch signalling pathway. The key players are a Delta-type ligand, a Notch receptor and the CSL transcription factor (Lai 2004).	22
Figure 2.10 Mechanisms of Notch juxtacrine signalling between two adjacent cells, and examples of resulting tissue level patterning for (a) lateral inhibition and (b) lateral induction (N: Notch receptor, D: Delta-type ligand). In the central figures, Notch signalling is represented by black arrows. In lateral inhibition, initially equivalent (purple) cells are resolved into distinct fates (blue and pink); cells with the highest Notch activity (receptor cells) coloured pink. Confocal scan (upper right) illustrates the case of ommatidium, with cell membranes stained green and Notch expression pink. In lateral induction, a boundary is established to segregate and/or organise two groups of cells. Black arrows indicate direction of Notch signalling, with pink cells having Notch activation. The confocal image (lower right) is of the fly wing primordium, where Notch activity (pink) is measured at the boundary of ligand (Serrate) expressing cells (green). After (Bray 2006)	24
Figure 2.11 Developmental patterning of the myocardium of the embryonic heart tube by Notch and Hairy-related transcription factors (Hey1 and Hey2) (Rutenberg et al. 2006).	25
Figure 2.12 Downstream targets of Snail genes. Molecules and processes shown in red are downregulated, while those in green are upregulated (Barrallo-Gimeno & Nieto 2005).....	26
Figure 2.13 Qualitative network of Notch and TGF β synergy in activating Snail gene expression in endothelial cells, derived from results of (Niessen et al. 2008). Snail1 is a direct target of TGF β while Snail2 is a direct target of Notch. Notch and TGF β synergistically upregulate Snail1, and Snail2 inhibits TGF β mediated Snail1 activation.	27
Figure 2.14 Mechanism for the integration of TGF β and Notch signalling. By binding to SMAD3, NICD recruits the SMAD3/SMAD4 complex to the CSL domain. The two pathways cooperatively activate Notch target genes (Heitzler 2010).....	28
Figure 2.15 Cell centre, cell vertex and cellular Potts models.....	30
Figure 2.16 Pugh matrix for cell to tissue level modelling methods for EMT, based on modelling realism, computational efficiency and ease of implementation. Green + = strength, red - = weakness, yellow S= neither strength nor weakness.....	31
Figure 2.17 3D cellular Potts model of endocardial EMT (Neagu et al. 2010). (A) EMT is assumed to be restricted to a circular area with a diameter of 30 cells, and takes place during the first 3×10^4 MCS. Probabilities are assigned for matrix production and cell proliferation. (B) The final state assumed at 5×10^4 MCS; showing full view, epithelium removed, and	

extracellular matrix removed respectively. Key: Epithelium: green, ECM: grey, mesenchymal cells: red. 35

Figure 2.18 The ‘piston test’ demonstrates the shortcomings of strictly local dynamics of the cellular Potts formalism. Cells (red) adhere to an ECM surface (green) which is shifted 1 pixel every 100 MCS. (A) With the standard cellular Potts model, cells cannot adjust to the changing boundary and holes appear near the attachment surface. The holes expand, and eventually the cells separate from the adhesive surface. (B) By including mechanical relaxation, whereby the aggregate of cells is given the same dynamics as each individual cell, the cells are able to adjust; matching the behaviour of real cell aggregates (Szabó et al. 2012). 37

Figure 2.19 Bulk cell movement at the ECM-cavity interface is unrealistic in the standard Potts model. While cells that adhere to the ECM are able to invade the space between ECM filaments; cells in the bulk are unable to adjust fast enough. Thus gaps form, and the average cell density is reduced at the cavity-matrix interface (Szabó et al. 2012). 37

Figure 2.20 2D extended cellular Potts model of cell invasion from an aggregate (Szabó et al. 2012). Cells with spontaneous motility invade an anisotropic, aligned and immutable ECM environment. (A) Morphology diagram of the invasion process for various values of the parameters β (cell-cell adhesion) and γ (cell-ECM adhesion). Configurations shown were obtained at $t = 1000$ MCS, which represents 17 hours. (B) Cell density changes along the direction of invasion reveal a steady invasion speed. The colour code indicates cell density compared to confluency. Temporal and spatial ranges on vertical and horizontal axes respectively are 2000 MCS and 500 lattice units. Plots annotated as 1, 2 and 3 correspond to the parameter values marked accordingly in (A). 38

Figure 2.21 3D cell centre simulation of cancer cell invasion, toward a source of morphogen to the right of the tumor (Ramis-Conde et al. 2008). Cells detach gradually as the intracellular concentration of β -catenin is upregulated (light grey). Unit of time is given in minutes. 39

Figure 2.22 Comparison showing the difference in β -catenin detachment wave simulations between centre based model of Ramis-Conde et al. (2008) (left figure) and the CompuCell3D-Bionetsolver model of Andasari et al. (2012) (right figure). 40

Figure 2.23 Results multicellular tumour spheroid simulation. The tumour grows from a single cell placed in the middle of the lattice. After 400 MCS, cells begin to detach and invade from the surface of the tumour. Cell colour represents β -catenin concentration (Andasari et al. 2012). 41

Figure 2.24 (a,b,c) Spheroid development at 14 days (2016 MCS) as a function of the number of 52 μ m long collagen threads, with 577, 2308 and 3462 threads respectively. Greyscale levels have no meaning, and are just to allow visualisation of individual cells. (d,e,f) Spheroid development as a function of time on a lattice containing 7500 12 μ m collagen threads at 0, 7 and 14 days respectively. Key: proliferative cells: purple, quiescent cells: blue, necrotic cells: green (Rubenstein & Kaufman 2008). 43

Figure 2.25 With a small virtual wound with normal proliferation and migration rates, epidermal cells migrate to cover the denuded area; shown at (A) 0, (B) 50, (C) 120, (D) 500 iterations. With a large virtual wound with the same migration and proliferation rates, keratinocytes on the wound bed begin to differentiate into committed cells before the epidermal cells have migrated to cover the large denuded area; shown at (E) 0, (F) 200, (G) 400, (H) 800 iterations. The level of TGF- β 1 in each cell was a function of the position of the agent in the stratified layers or in contact with the matrix (I) (J). Key: keratinocyte stem cells (blue), transit amplifying cells (light green), committed cells (dark green), corneocytes (brown), provisional matrix (dark red), secondary matrix (green). Cell agent diameter = 10 μ m (Sun et al. 2009).	44
Figure 2.26 2D cellular Potts simulations of part of cancer invading front; (left) without cell proliferation, (right) with cell proliferation (Turner & Sherratt 2002).	46
Figure 2.27 Notch, TGF- β and BMP activity in endocardial cushion formation (Wagner & Siddiqui 2007)	48
Figure 2.28 Relationship between models/data and annotations of models/data using ontologies. Models will be annotated to terms using a standard annotation scheme. These annotations will be stored in a central repository to enable fast querying (Baldock et al. 2010).	53
Figure 2.29 Spatial and temporal scales encompassed by the Human Physiome Project. Markup languages developed at Auckland Bioengineering Institute for different levels of scale are indicated, as well as the types of mathematical model appropriate to each level (Hunter et al. 2002).	55
Figure 2.30 Example simulations of an intestinal crypt developed in Chaste using (a) the cell centre model and (b) the cell vertex model. Proliferating cells: yellow, non-proliferating cells: pink (Osborne et al. 2010).	58
Figure 2.31 SemGen annotator tool (top) is used to create composite annotations for model variables, and the merger tool (bottom) is used to suggest resolution points between models (Neal 2010).	61
Figure 3.1 Overview of the computational methods applied to particular biological mechanisms, and the main justifications for the use of each method	64
Figure 3.2 Image processing workflow in ImageJ for wildtype cells. a) Starting image b) Nuclei extracted and overlaid in white, background removed. c) Find maxima used to output segmented particles d) Outlines used to add cell shapes to ROI manager.	69
Figure 3.3 Closeup inspection of overlay for wildtype condition suggests that cell outlines provide a reasonably good fit.	70

Figure 3.4 Image processing workflow in ImageJ for Notch1 activated cells undergoing EMT. a) Starting image b) Myocardium removed. c) Adjust Threshold to create binary image d) Analyse particles used to add cell shapes to ROI manager. 71

Figure 3.5 CPM representation of an index-copy attempt for two cells on a 2D square cell lattice – The “white” pixel (source) attempts to replace the “grey” pixel (target). The probability of accepting the index copy is given by $\min(1, e^{\Delta H/T})$. Thus any change reducing the entropy is accepted, while those increasing the entropy depend on the change in H and the T parameter (Swat et al. 2012). 73

Figure 3.6 Data points expressing the relationship between cadherin expression level and aggregate surface tension fall almost exactly on a straight line ($R^2=0.9965$) that passes very close to the graph’s origin, intersecting the Y axis (representing zero N-cad expression) at the very low surface tension value of 0.32 erg/cm² (Foty & Steinberg 2005). 74

Figure 3.7 Cells interact at their surfaces (adjacent lattice sites) in cellular Potts models. a) At the start of the simulation, cell types are mixed. b) After 8000 Monte Carlo Steps (MCS), the cells are sorted, with the less cohesive type (red) forming an outer layer. 75

Figure 3.8 Sorting out of subclones differing only in expression level of a given cadherin. Two N-cad-transfected L cell subclones, expressing N-cad at their surfaces in the ratio of 2.4:1, were stained with red and green fluorescent membrane-intercalating dyes, mixed in equal proportions and cultured as hanging drops. (a) Confocal optical section through an aggregate after 4 hours of incubation, showing initial cell mixture. (b) Confocal optical section through another aggregate after 24 hours of incubation. As predicted by the DAH, the cell line expressing the lower level of N-cad (surface tension $\sigma = 2.4$ erg/cm²), labeled red, segregates from and envelops the cell line expressing higher amounts of N-cad (surface tension $\sigma = 5.6$ erg/cm²), labeled green. This demonstrates that cell sorting does not require (although it does, of course, permit) qualitative differences in cell–cell ‘recognition specificity’ (Steinberg 2007) 75

Figure 3.9 Skeleton of Bionetwork API implementation within a CompuCell3D Python steppable 82

Figure 3.10 UML class diagram of multiscale cellular Potts model, as implemented in CompuCell3D 84

Figure 4.1 *In vitro* endocardial explants, comprising endocardium (e) and myocardium (m). a) Wildtype (WT) endocardial tissue remains in a monolayer. b) Notch activated (Tie2-Cre;N1ICD) endocardial cells migrate on the surface of the collagen gel. c) BMP2 treatment causes wildtype cells to both migrate on the surface and invade the gel (Luna-Zurita et al. 2010). 89

Figure 4.2 Cellular Potts simulations of *in vitro* EMT. a) Endothelial monolayer on the surface of collagen gel. b) With reduced EC-EC adhesion, cells migrate on the surface, but do not invade the gel (partial EMT). c) With reduced EC-EC adhesion and increased EC-Gel adhesion, some cells invade the gel in 3D (full EMT). 92

Figure 4.3 UML state machine diagram for contact inhibited mitosis model.....	95
Figure 4.4 Mitosis simulations. a) Cells separate in 2D under set B parameters. b) monolayer prevails under set B if contact-inhibited mitosis is included. c) Cells migrate in 2D and 3D under set C parameters. d) Including contact-inhibited mitosis rescues monolayer integrity for set C. Daughter cells are illustrated in a different shade in order to highlight the effects of mitosis.....	97
Figure 4.5 Shape metric comparison of wild type and N1ICD murine ventricular endocardial cells. a) Wildtype cells have a mean circularity of 0.536, and N1ICD cells 0.187 ($p < 0.001$). b) Wild type cells have a mean aspect ratio of 1.735 and N1ICD cells 2.452 however this difference is not significant. Error bars show standard deviation.	99
Figure 4.6 a) Circularity of simulated cells falls with increasing Target Surface. b) This is accompanied by increased speed (and hence motility). Error bars show standard deviation, and caps show the range, from 10 simulation replicas.....	101
Figure 4.7 a) Speed of simulated cells increases as a result of increasing 'Temperature'. b) This is accompanied by a reduction in circularity. Error bars show standard deviation, and caps show the range, from 10 simulation replicas.....	101
Figure 4.8 Separation force measurements for Ecad cells held in contact for 0.5-60min (Chu et al. 2004).	102
Figure 4.9 Cells with fibroblastic morphology. Simulation snapshots at 50000 MCS, for different levels of plastic coupling λ [20, 5] and cell-cell contact energy [-15, 0] (units are $10^{-15} \text{kg}^1 \text{s}^{-2}$).....	104
Figure 4.10 a) Monolayer with no epithelial adhesion and b) monolayer with both labile and junctional adhesion, after 50000 MCS. c) Labile adhesions, C, and junctional coupling, FPP, do not significantly affect average cell roundness ($p > 0.05$, $n=316$ in all cases).....	104
Figure 4.11 Under conditions of increased motility (Temp=500) and moderate cell-medium adhesion (-20), cells scatter under conditions with a loss of junctional coupling, but not with a loss of labile adhesion alone. In all cases cells maintain a rounded morphology (units are $10^{-15} \text{kg}^1 \text{s}^{-2}$).	105
Figure 4.12 Combined loss of cell adhesion, gain of cell-medium adhesion and increased motility simulation a) Snapshot at 50000 MCS b,c,d) Mean and standard deviation of simulated (CM-200) cell characteristics at 50000 MCS ($n=316$) compared with those obtained from <i>in vitro</i> Notch activated endocardial cells (N1ICD, $n=54$) (Luna-Zurita et al. 2010) or MCF-7 cells transfected with vimentin ($n=5$) (Mendez et al. 2010).	107
Figure 4.13 Addition of an elongation parameter to the model, and reduction of cell-medium adhesion from -200 to -100 ($10^{-15} \text{kg}^1 \text{s}^{-2}$), provides a better fit, a) Snapshot at 50000 MCS b,c,d) Mean and standard deviation of simulated cell characteristics at 50000 MCS ($n=316$)	

compared with those obtained from <i>in vitro</i> notch activated endocardial cells (N1ICD, n=54) (Luna-Zurita et al. 2010) or MCF-7 cells transfected with vimentin (n=5) (Mendez et al. 2010).	107
Figure 4.14 Notch Lateral induction model in JDesigner and reaction equations. ni: active Notch (NICD), di: free (unbound) Delta. Other variables are defined in the text.	111
Figure 4.15 Notch activation is highly correlated with Delta expression in individual cells, as expected. Data taken from simulation as shown in Figure 4.16b over the first 800 MCS. Pearson correlation coefficient $r = 0.953$	114
Figure 4.16 Notch activation after 800 MCS for (a) a tissue starting with a radius of 20 pixels and (b) a tissue starting with a radius of 30 pixels.....	115
Figure 4.17 Average Notch activation over time in a tissue starting from (a) 20 pixel radius and (b) 30 pixel radius	115
Figure 4.18 Spatial and temporal scales of heart morphogenesis modelling. The modelling framework encompasses spatial scales from 10^{-9} m (proteins) to 10^{-3} m (the primitive heart tube), and temporal scales from 10^{-6} s (molecular events) to 10^6 s (weeks of heart development). Spatial and Temporal ontologies that apply to different levels of scale are also shown. PRO: Protein Ontology, ChEBI: Chemical Entities of Biological Interest, CL: Cell Type Ontology, FMA: Foundational Model of Anatomy, GO-CC: Gene Ontology Cellular Component, EHDA: Edinburgh Human Developmental Anatomy, GO-MF: Gene Ontology Molecular Function, CBO: Cell Behaviour Ontology, MP: Mouse Phenotype Ontology, OPB: Ontology of Physics for Biology, PATO: Phenotype And Trait Ontology, GO-BP: Gene Ontology Biological Process	117
Figure 4.19 Schema for creating composite annotations from terms in multiple reference ontologies.	120
Figure 5.1 Major steps in the drug discovery process. A typical path for drug discovery is presented. Compounds in a chemical library are screened to identify those molecules that interact with the intended target. Molecules that are positive in this assay “Hits” begin the process of lead identification (Hit-to-Lead) and Lead optimization. <i>In vitro</i> toxicity screening as well as screens designed to identify ADME, genotoxicity, and cardiac toxicity should be done early in this process in order to identify high risk molecules early (McKim 2010).	133

List of Tables

Table 1.1 Modelling approaches, tools, existing repositories and example systems for different levels of spatial scale.....	6
Table 4.1 Surface energy parameters J ($10^{-15}\text{Kg}^1\text{s}^{-2}$) for <i>in vitro</i> EMT simulations.	91
Table 4.2 Surface energy parameters J ($10^{-15}\text{Kg}^1\text{s}^{-2}$) for contact-inhibited mitosis simulations.	95
Table 4.3 Parameter values of the lateral induction model (N. J. Savill & J. A. Sherratt 2003)	110
Table 4.4 OBO and OWL representation of composite annotation.....	120

List of Acronyms

ADME	Absorption Distribution Metabolism and Excretion
AEPC	Association for European Paediatric Cardiology
API	Application Programming Interface
AVC	Atrioventricular Canal
BAV	Bicuspid Aortic Valve
BFO	Basic Formal Ontology
BMP	Bone Morphogenetic Protein
CAT	Common Arterial Trunk
CBML	Cell Behaviour Markup Language
CBO	Cell Behaviour Ontology
CellML	Cell Markup Language
Chaste	Cancer Heart and Soft Tissue Environment
CHD	Congenital Heart Defect
CheBI	Chemical Entities of Biological Interest
CL	Cell Type Ontology
COPASI	Complex Pathway Simulator
CPM	Cellular Potts Model
CT	Conotruncus
d-TGA	Dextro-Transposition of the Great Arteries
DORV	Double Outlet Right Ventricle
ECM	Extra Cellular Matrix
E	Embryonic Day
E-Cadherin	Epithelial Cadherin
EHDA	Edinburgh Human Developmental Anatomy
EMT	Epithelial to Mesenchymal Transition
EQ	Entity-Quality
FieldML	Field Markup Language
FLAME	Flexible Large Scale Modelling Environment
FMA	Foundational Model of Anatomy
FPP	Focal Point Plasticity
GO	Gene Ontology
GO-BP	Gene Ontology Biological Process
GO-CC	Gene Ontology Cellular Component
GO-MF	Gene Ontology Molecular Function
HPO	Human Phenotype Ontology
IAA	Interrupted Aortic Arch
I-TGA	Levo-Transposition of the Great Arteries
MathML	Mathematical Markup Language
MCS	Monte Carlo Step
MET	Mesenchymal to Epithelial Transition
MMP	Matrix metalloproteinase
MP	Mouse Phenotype
MRI	Magnetic Resonance Imaging
NICD	Notch Intracellular Domain
NFAT	Nuclear Factor of Activated T-Cells
OBO	Open Biomedical Ontologies
ODE	Ordinary Differential Equation
OFT	Outflow Tract
OMIM	Online Mendelian Inheritance in Man

OPB	Ontology of Physics for Biology
OWL	Web Ontology Language
PA	Pulmonary Atresia
PATO	Phenotype and Trait Ontology
PDE	Partial Differential Equation
PK	Pharmacokinetics
PTA	Persistent Truncus Arteriosus
SBML	Systems Biology Markup Language
SBW	Systems Biology Workbench
TGA	Transposition of the Great Arteries
TGF- β	Transforming Growth Factor beta
TOF	Tetralogy of Fallot
UO	Unit Ontology
VE-Cadherin	Vascular Endothelial Cadherin
VEGF	Vascular Endothelial Growth Factor
VPH	Virtual Physiological Human
VSD	Ventricular Septal Defect
XML	eXtensible Markup Language

Chapter 1

1. Introduction

1.1. Background Context

1.1.1. Modelling and Simulation in Biomedicine

Computational modelling and simulation are increasingly utilised as tools within biomedical research. Experimental findings about the biological systems under investigation become ever more detailed, complex and difficult to mentally connect. Thus there is a need to integrate and represent the information in ways that are comprehensible. Conceptual models – usually diagrammatic – play a key role in summarising the current knowledge, or hypotheses, about a biological process. However, the number of interactions involved in a biological process is so large that representing all of them in a diagram is impractical. Besides which, large numbers of interacting components can exhibit collective behaviours that can not necessarily be understood from the behaviour of individual components. These interactions can be integrated with mathematical models. However, these models are often too complex for analytical solution. They can be translated into sets of algorithms or computational models, which are then implemented as simulations. Simulation models provide a representation that can be tested and validated against experimental results. They can also provide suggestions for further experimentation; either by indicating missing knowledge or by providing experimental predictions, as suggested by the black box model below (Figure 1.1).

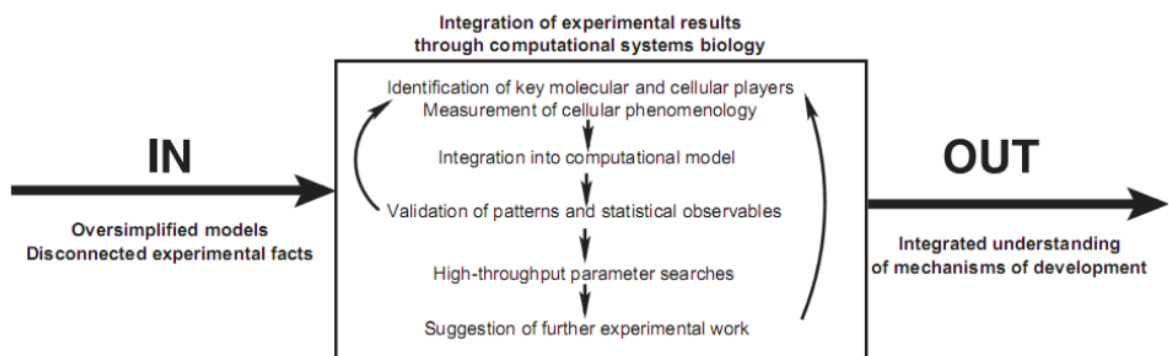


Figure 1.1 Schematic of the systems biology approach (Merks & Glazier 2005)

The majority of computational biology models are single-scale models. This is partly because different types of mathematics, and thus modelling techniques, are suitable for representing the behaviour of biological systems at different levels of scale. For example: biochemical reactions are best represented with ordinary differential equations, or some form of network model; diffusion of chemical fields and cellular physiology are best modelled with partial differential equations; while cellular behaviour might be represented by some form of agent based modelling, because this reduces the complex machinery of cells to a handful of core behaviours, encoded as rule-based decisions.

Peer-reviewed, single-scale models are now frequently deposited in online databases, in a format that can be used by many computer programs. These are generally in eXtensible Markup Language (XML) format. For example, biochemical reaction models are made available as System Biology Markup Language (SBML) files on the Biomodels.net database (Li et al. 2010). These can be reused, in whole or part, and simulated either deterministically or stochastically depending on the tool used. There are now over 150 software systems that are SBML compatible. Similarly, hundreds of single cell physiology models have been made available in the CellML model repository. These are modular, allowing for easier reuse of components. There is a stated aim to provide a catalogue of standard virtual subcellular parts (such as ion exchangers or signal transduction pathways) that can be composed into more complex models (Garny et al. 2008).

Agent based formalisms have tended to be more disparate, with each model being encapsulated in a specific problem. While there is a library of models of biochemical reactions, and a library of models of cell physiology, and even a small library of 3D geometrical models of human organs (provided in another XML specification, FieldML); there is no comparable repository of models of cellular ‘social’ behaviour in tissues. Currently the closest approximations to this are the modelling environments CHASTE (Pitt-Francis et al. 2009) and CompuCell3D (Cickovski et al. 2007), both of which are open source and provide functionality for multiscale modelling of cells and tissues. CompuCell3D achieves this through cellular Potts models, while CHASTE implements it with both cell centre and cell vertex models. These different techniques are explained further in chapter 2.

The Physiome consortium for multiscale physiological modelling has tended to ignore cells, treating them as parts of continuous tissues (Popel & Hunter 2009). In this way, detailed subcellular physiology and reaction networks (SBML or CellML) can be integrated with mechanical models of organs (FieldML) such as heart wall excitation and deformation. But this

approach is not suitable where the changing shape, rearrangement and migration of individual cells are a crucial part of the system we wish to simulate and understand. This is very much the case when modelling morphogenesis. While cell vertex and cell centre modelling do represent the behaviour of individual cells, and this can incorporate multiscale phenomena, such as for example in the Epitheliome project (Smallwood 2006); the cellular Potts formalism has the most flexibility in terms of directly representing changes in cell shape. As this thesis focuses on cardiac morphogenesis and epithelial to mesenchymal transition in particular, explicit representation of cellular shape and adhesion are important modelling features to include.

1.1.2. Epithelial to Mesenchymal Transition

Epithelial and mesenchymal cells differ in a number of phenotypic characteristics. Epithelial cells form layers of cells that are organised by a variety of junctional adhesions. These epithelial layers have apical-basolateral polarisation due to the localised distribution of different adhesion molecules. Epithelial cells are motile, and can move around *within* the layer. However, they do not *move away* from the layer under normal conditions. Mesenchymal cells do not form an organised layer, and contact other mesenchymal cells only by focal adhesions. In culture, mesenchymal cells have spindle-shaped (fibroblastic) morphology, and are highly motile, while epithelial cells grow in clusters, maintaining full cell-cell adhesion (Thiery & Sleeman 2006).

Epithelial cells can become mesenchymal cells in a process termed Epithelial to Mesenchymal Transition (EMT). This is broadly a process of the loss of epithelial characteristics and acquisition of mesenchymal characteristics. As such, there is a wide spectrum of types of EMT, in different tissues, regulated by different signalling pathways, which may involve the loss and gain of different phenotypic characteristics. For example, in a partial EMT, cells may lose tight junctions, but maintain adherens junctions or desmosomes (Figure 1.2 The cycle of epithelial cell plasticity (Thiery & Sleeman 2006)

). Alternatively, they may lose all types of epithelial adhesion, but remain non-invasive (Luna-Zurita et al. 2010). EMT, and the reverse process: Mesenchymal to Epithelial Transition (MET) are used throughout embryonic development to build complex tissues and organs. Cancer cells also undergo similar processes during metastasis, although there is some controversy over whether this is strictly speaking a reactivation of the embryonic programme (Savagner 2010). However, the process in cancer cells is certainly controlled by many of the same molecular pathways, and the phenotypic changes are similar.

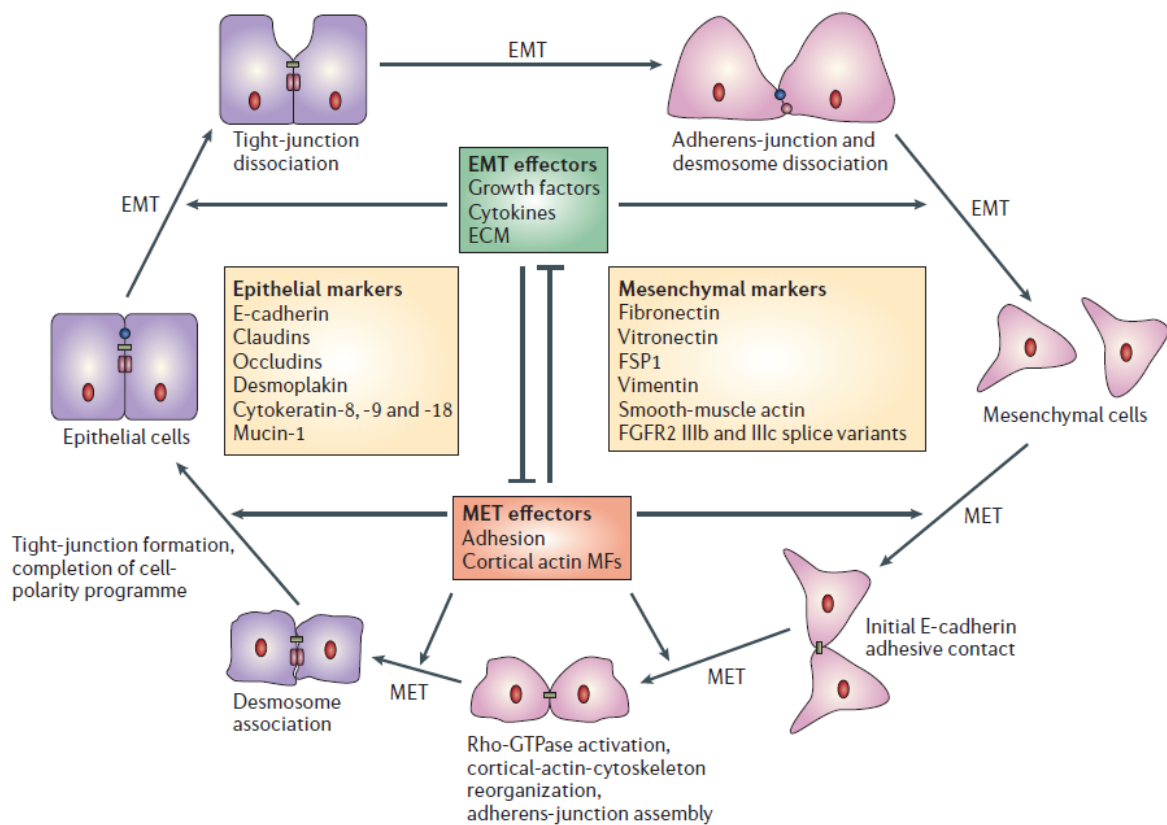


Figure 1.2 The cycle of epithelial cell plasticity (Thiery & Sleeman 2006)

In heart development, EMT underlies the growth of the endocardial cushions, which are localised tissue swellings on the inner surface of the embryonic heart tube. As the heart tube loops and twists into its final configuration, these endocardial cushions grow and fuse, to form much of the inner structure of the heart, including the heart valves and membranous septa. This EMT is restricted to the atrioventricular canal and the outflow tract regions. These areas are established through Notch signalling by lateral induction (Timmerman et al. 2004). Protein signals from the myocardium, including BMP2 (Bone Morphogenetic Protein 2) and TGF β (Transforming Growth Factor β) are also necessary for the fully invasive EMT (Luna-Zurita et al. 2010). The regions of the myocardium expressing these proteins are also controlled by Notch signalling and lateral induction, though the Notch expression pattern is the reverse of the endocardium. This is covered in more detail in the Literature Review.

Mutations in Notch pathway ligands and receptors are associated with a range of congenital heart defects in humans, including tetralogy of Fallot and Alagille syndrome. Deletion of Notch target genes causes congenital heart defects in mice, due to impaired EMT (Fischer et al.

2007). Due to its importance in heart development, there has been a wealth of *in vivo* and *in vitro* studies into the genetic and protein factors underpinning endocardial EMT. Many of these studies propose conceptual models (diagrams) of how the process is controlled, but it remains difficult to comprehend fully.

Thus a multiscale representation of Notch signalling in endocardial EMT is an excellent candidate for multiscale modelling analysis. This is one of the most thoroughly investigated examples of EMT, providing a starting point for building models that will be relevant, or may be adapted for EMT in other fields of development and disease. EMT is also one of the most intensively researched aspects of heart development, providing an entrance for modelling the complex mechanisms of congenital heart defects. The models also have the potential to be applied in a clinical setting; for example through refining the tissue engineering of replacement heart valves (Sewell-Loftin et al. 2011). Additionally it is possible to use this type of multiscale modelling as a tool within developmental toxicology (Shah & Wambaugh 2010).

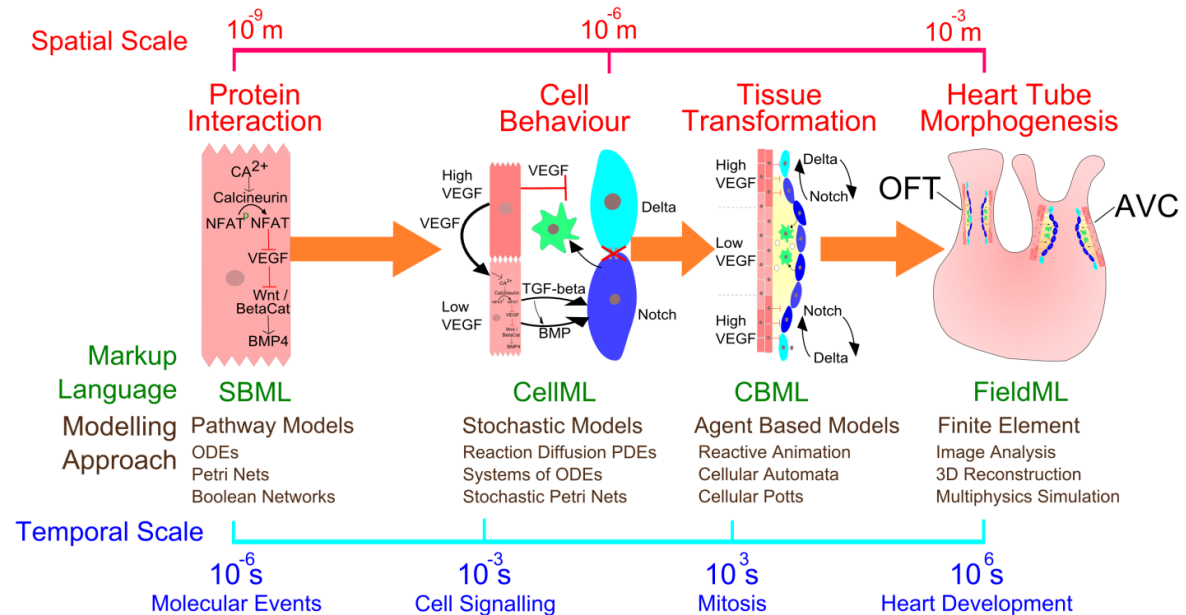


Figure 1.3 Schematic of multiscale modelling applied to EMT in heart development. OFT: Outflow Tract, AVC: Atrioventricular Canal, VEGF: Vascular Endothelial Growth Factor

A schematic of the multiscale modelling approach as applied to EMT in heart development is given in Figure 1.3. This indicates the range of temporal and spatial scales that are applicable to this process, as well as the modelling approaches that could be used at each level of scale. Important events taking place at each level during this process are illustrated. Protein interactions within cells determine cell behaviour, the interactions of which lead to tissue

transformations which govern heart morphogenesis. This is reviewed in more detail in the literature review.

As shown in Figure 1.3 the processes relevant to EMT during heart development encompass spatial scales from 10^{-9} m (proteins) to 10^{-3} m (the primitive heart tube), and temporal scales from 10^{-6} s (molecular events) to 10^6 s (weeks of heart development). The different modelling approaches applicable at each scale have the potential to be encoded in an XML format that aids in the sharing of such models between platforms. It would be unwieldy to incorporate all levels of scale within a single multiscale model; rather a given multiscale model would normally combine models from two levels. For example, an SBML model of a reaction pathway could be incorporated within each cell in agent based model (Andasari et al. 2012). Alternatively, a CellML model of cellular physiology can be interpolated at discrete points within a continuum organ or tissue model, encoded in FieldML (Hunter et al. 2008).

The modelling methods applicable to each level of spatial scale are outlined in more detail in Table 1.1, alongside some of the tools applied, and candidate systems. One point to note is that existing repositories of models are essentially only developed for the molecular and single cell level at the present time. Additionally, that information modelling (in the form of ontologies) is appropriate to all levels of scale, and thus may provide one means of scale-linking between models.

Table 1.1 Modelling approaches, tools, existing repositories and example systems for different levels of spatial scale.

Biological Scale	Modelling Approach		Modelling Tools	Example Systems	Model Repositories
Molecular	Information Modelling (Ontologies)	Network / Pathway Modelling -Petri Nets -ODEs -Boolean Networks -Rule Based Modelling	BiologicalNetworks Systems Biology Workbench* Labview Cell Illustrator Biotapestry BioNetGen	Notch/Delta Lateral Induction Crosstalk between TGF- β , Notch, BMP, VEGF pathways during EMT Specification of chamber forming regions of heart tube	Biomodels KEGG SigPath Reactome CellML JWS Online Science STKE BRENDA DOQCS
Single Cell		Systems of ODEs Reaction Diffusion Systems	VCell* OpenCell*	Electrophysiology Differentiation Apoptosis	
Cell – Tissue		Agent Based Modelling -Reactive Animation -Cellular Potts -Cell Centre -Cell Vertex	Statecharts NetLogo Xholon CompuCell3D* Chaste* FLAME*	Epithelial-Mesenchymal Transformation Heart Field Specification Neural Crest Cell Migration & Proliferation Vasculogenesis	
Anatomical Part – Organ		Finite Element Modelling Image Analysis -Segmentation -Visualisation -Simulation	Comsol* CMISS* CardiViz3D* openDICOM* GIMIAS*	Rotation and shortening of the Outflow Tract Growth of Heart Valves Atrial and Ventricular Septation Developing Conduction System	

*Recommended as part of VPH Toolkit (Garny et al. 2010)

1.2. Aims and Objectives

The aims of this work reflect the three aspects that are brought together to represent the thesis: EMT; computational simulation models; and multiscale appreciation of their interaction. They can be stated thus.

Aims

- A1: Integrate existing understanding of Epithelial to Mesenchymal Transition (EMT) at different levels of spatial scale.
- A2: Increase understanding of EMT by representing key features with computational simulation models.
- A3: Refine existing approaches for multiscale modelling of developmental processes.

To achieve these stated aims, several objectives become clear that build upon each other to improve understanding of the multiscale system of EMT through modelling and simulation.

Objectives

- O1: Understand EMT processes at various levels of spatial scale (A1).
- O2: Define conceptual multiscale model of EMT (A1)
- O3: Build cell and tissue level simulations of endocardial cells undergoing EMT (A2).
- O4: Explore key EMT signalling pathways through protein level modelling (A2).
- O5: Use existing experimental results for model validation (A2).
- O6: Explore methodological approaches for semantic integration of multiscale models (A3).
- O7: Investigate multiscale modelling and simulation, using the example of Delta-Notch protein signalling in a dynamic tissue field (A3).

1.3. Plan of Thesis

Chapter 1, **Introduction**, has given the background and context of the thesis, including a primer on simulation modelling in biomedicine, and multiscale modelling in particular; as well as epithelial to mesenchymal transition and its role in heart development. Chapter 2, **Literature Review**, extends on this by giving a more detailed background on heart development and EMT in development and disease. Existing EMT simulations, as well as EMT pathway simulations and multiscale biomedical models in general, are also reviewed in chapter 2. Chapter 3, **Methods**, describes the research methods used, including image processing cellular Potts modelling and pathway modelling. Chapter 4, **Results**, reports the findings of the simulation models, and validation with experimental imaging data. An approach for annotation of multiscale models is also reported. Chapter 5, **Discussion and Conclusions**, reviews and analyses these results and their implications, including the contribution to original knowledge made by the thesis. Future research opportunities and potential model applications are also discussed.

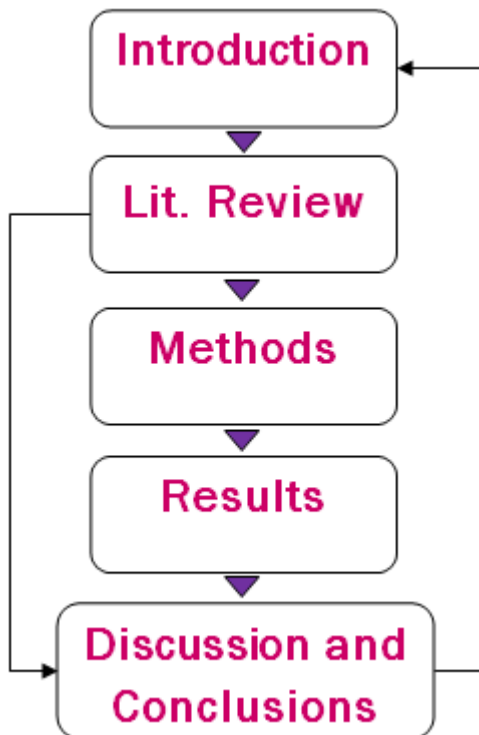


Figure 1.4 Structure of thesis.

Chapter 2

2. Literature Review

2.1. Introduction

This chapter opens with an overview of heart development (Section 2.2.). This tells the story of the role of EMT in heart development addressed from progressively smaller levels of spatial scale, with a progressively narrowing focus. The section begins with a brief overview of congenital heart diseases (Section 2.2.2.). This is followed by an in-depth review of the anatomy of heart development (Section 2.2.3.), which helps the reader understand how different types of congenital heart defect arise as a result of disrupted developmental processes. Section 2.2.4. narrows the focus to a single developmental process, and explores in detail the role of EMT in heart development, and the major protein signalling pathways that control it. This focus is narrowed again in Section 2.2.5, which focuses on the role of Notch signalling. The basic model for Notch signalling, and the difference between lateral inhibition and lateral induction is described. This is followed by a review of how Notch signalling delineates the cushion forming regions of the embryonic heart tube; and how Notch signalling interacts with other pathways during endocardial EMT.

Section 2.3. reviews the major cell-level modelling techniques that are available. These are continuum, cell centre, cell vertex and cellular Potts modelling. These different techniques are compared according to their suitability for features that are requirements in the modelling of EMT.

Following the review of different available modelling techniques, Section 2.4. reviews existing simulations of EMT. Several cell centre and cellular Potts models of EMT, with a range of different features are described in this section. Some are 2D and some are 3D. Some include subcellular models of protein dynamics. Some represent extracellular

matrix (ECM) threads explicitly. Some investigate cell-cell or cell-ECM adhesion, and some include both. Some models assume the cells actively migrate in a particular direction.

Section 2.5. reviews existing simulations of pathways that play a role in EMT (which were introduced in Section 2.2.). This section briefly outlines existing simulations of Notch signalling (Section 2.5.1.), TGF- β and BMP signalling (Section 2.5.2.) and Wnt/ β -catenin and E-cadherin signalling (Section 2.5.3.).

Section 2.6. Outlines some of the challenges for multiscale modelling in biomedicine, including the challenges of consistent semantic annotation of models at different levels of scale.

Section 2.7. then provides details on some existing platforms for multiscale simulation in biomedicine, and how they are individually approaching the challenges.

2.2. Heart Development

2.2.1. Congenital Heart Diseases

Congenital Heart Diseases (CHDs) are the most common type of birth defect, and the main cause of birth defect related mortality and morbidity. The incidence of moderate to severe forms of CHD is about 6/1000 live births or 19/1000 if bicuspid aortic valve (BAV) is included (Hoffman & Kaplan 2002). BAV is a CHD in which two aortic valve leaflets develop instead of three. BAV leads to aortic valve stenosis later in life, and most people born with BAV will eventually require valve surgery (Otto 2002). Therefore the higher figure of 19/1000 live births is appropriate when considering the lifetime effects of CHD. The effects of long term remodelling of the heart following surgical repair present a significant challenge in terms of therapy planning and disease evaluation (Mansi et al. 2011).

There have been a great number of studies linking particular genetic mutations to particular CHDs and syndromes in human populations, e.g. (Pierpont et al. 2007). Likewise, there is an abundance of research investigating the effects of particular

genetic mutations on particular heart development processes in animal models, e.g. (Niessen et al. 2008). Such research is, of necessity, highly specific and linear in scope. What is sought is the identification of one gene, its effect on one mechanism and its implication in one type of disease.

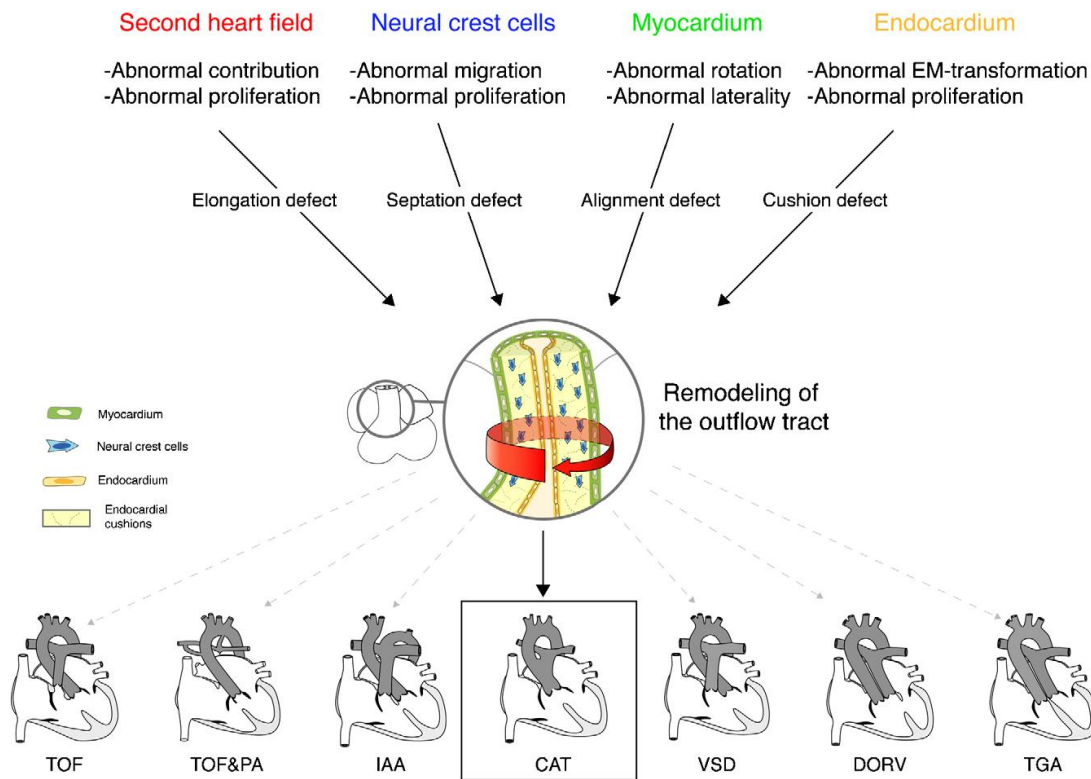


Figure 2.1 A malformation may originate from different mechanisms. Common arterial trunk (CAT), Tetralogy of Fallot (TOF) or any other disease resulting from abnormal remodelling of the outflow tract, may result from the participation of a defect secondary heart field and/or a migration defect of the neural crest cells, and/or a rotation defect of the myocardial outflow tract and/or a formation defect of the endocardial cushions (Bajolle et al. 2009). Key: TOF: Tetralogy of Fallot, TOF&PA: Tetralogy of Fallot and Pulmonary Atresia, IAA: Interrupted Aortic Arch, CAT: Common Arterial Trunk, VSD: Ventricular Septal Defect, DORV: Double Outlet Right Ventricle, TGA: Transposition of the Great Arteries

In reality, morphogenesis is highly nonlinear: several genes coordinate a single mechanism, one gene affects several mechanisms, and several mechanisms interact together in normal or abnormal development (Figure 2.1). Thus one type of CHD can be the result of many different genetic mutations, and one genetic mutation can lead to several different types of CHD, under different circumstances. For example, mutations in the Notch signalling pathway, such as the Notch1 receptor or the JAG1 ligand, correlate

with BAV as well as Alagille syndrome; and more specifically with tetralogy of Fallot and pulmonary valve stenosis (Butcher et al. 2011).

2.2.2. Anatomy of Heart Development

The development of the embryonic heart commences in week 2 and is fully formed by week 8. This process is well documented, e.g. (Kirby 2007). Week 2 of foetal life provides the first milestone of cardiac development as cells of the splanchnic mesoderm cluster to form two endocardial tubes (termed the cardiogenic crescent) at the cranial end of the embryo. At day 20, these two endocardial tubes join together at the median and ventral part of the embryo, thus forming the primitive heart tube. At this stage of development the first contractions occur, permitting actual blood circulation (Christoffels et al. 2010). At the end of week 3 the heart tube folds into an S-shape, looping to the right (D-loop). This repositioning constitutes a crucial step towards the morphology of the heart because it brings the future heart chambers and their inflow and outflow tracts into their relative spatial positions (Figure 2.2).

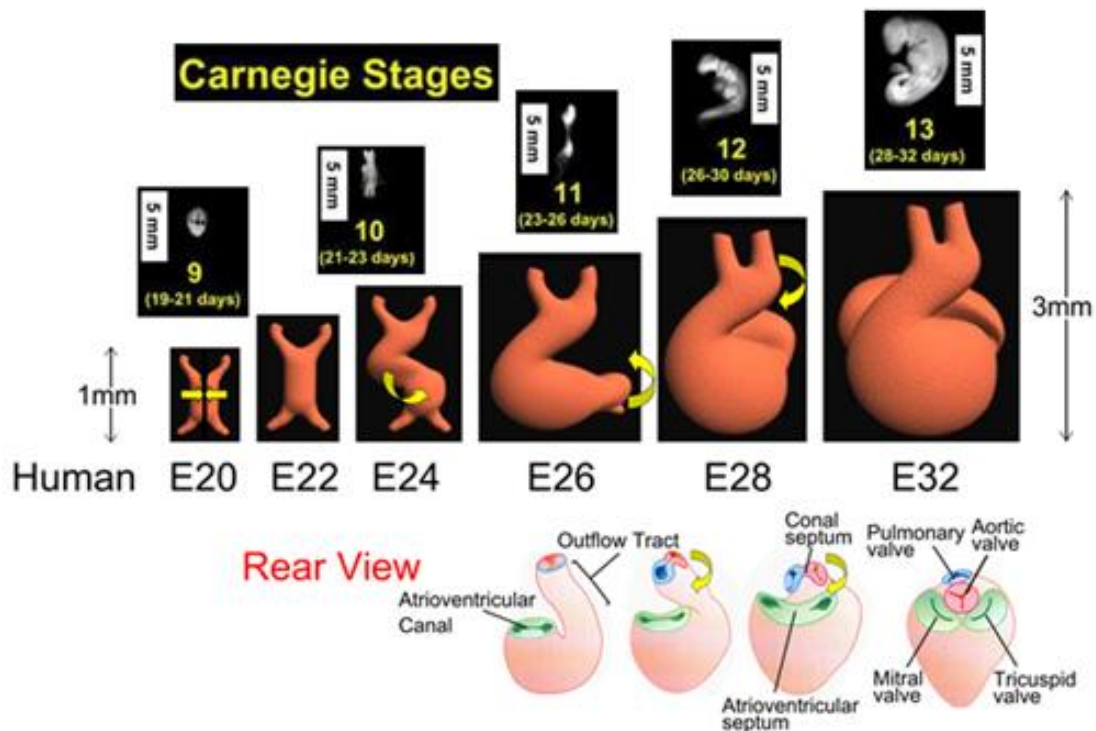


Figure 2.2 Heart looping and wedging between embryonic day 20 (E20) and E32, or Carnegie stages 9 to 13. After (Kirby 2007) and (Hill 2011).

In Figure 2.2, embryonic days are shown for human and mouse development. The heart grows from about 1mm to 3mm between E20 and E32 in human. The primitive atrium originates at the base of the heart, but due to looping and folding is later positioned at the apex. Later stages are shown in more detail from a rear view, sectioned at the level of developing valves. In normal development, the outflow tract rotates clockwise while growing endocardial cushions divide it into the aorta and pulmonary artery. Simultaneously, endocardial cushions are growing in the atrioventricular canal, which form the atrioventricular septum, and later the tricuspid and mitral valves (Figure 2.3). Throughout week 4 the ventricles grow considerably, in particular the right ventricle, by addition of myocardial cells from the second heart field. At the same time the apex of the ventricles balloon in sequence from the ventricular loop, leading to the development of the ventricular septum. Two processes in the development of the embryonic heart are crucial to the understanding of the conotruncal family of congenital heart diseases: looping and aortic wedging.

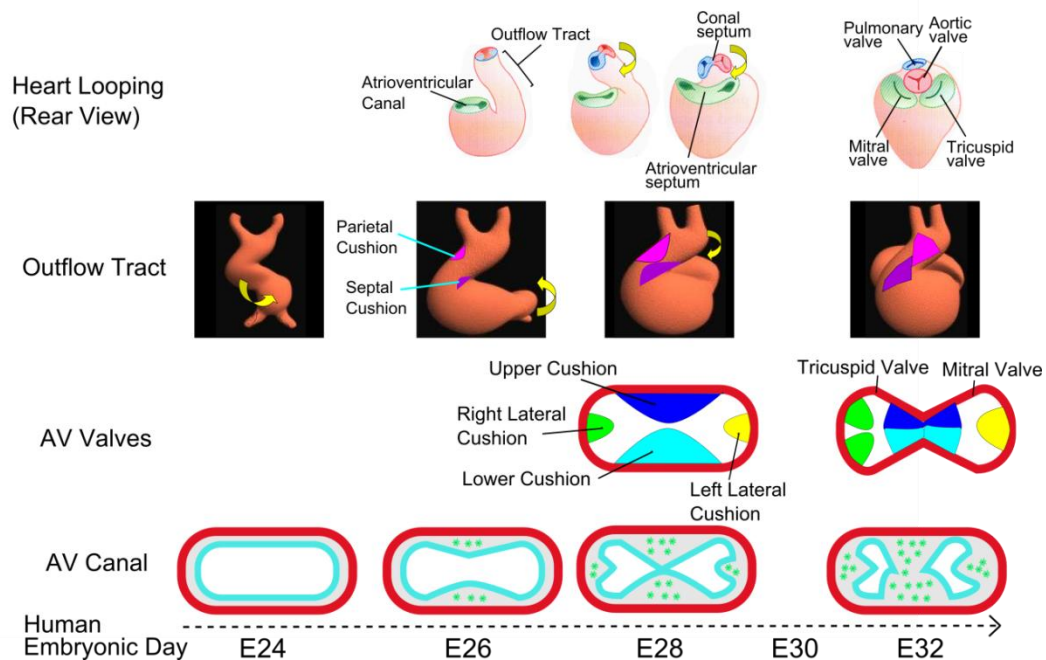


Figure 2.3 Detail of endocardial cushion growth and fusion during heart looping and wedging, after (Kirby 2007)

Looping is the first manifestation of the asymmetry of the embryo. Aortic wedging is due to the rotation of the myocardial wall of the outflow tract, itself secondary to remodeling of the inner curvature of the heart. This rotation leads the developing aortic valve, which is initially located at the right part of the conotruncus, downwards, posteriorly and to the left, to nestle between the two atrioventricular valves, thus establishing the mitral-aortic continuity. When the heart is fully formed at 8 weeks, the aorta is posterior to the pulmonary artery, with the degree of rotation about 30 degrees short of a complete juxtaposition of 180 degrees (Figure 2.4). At the same time, the conal septum develops by fusion of the endocardial cushions of the outflow tract and is taken along leftwards by the rotation of the developing aortic valve, to join the upper primitive ventricular septum between the two limbs of the septomarginal trabeculation. The conal (or outflow tract) septum is helical in shape, due to the rotation of the outflow tract while the endocardial cushions are growing and septation is occurring.

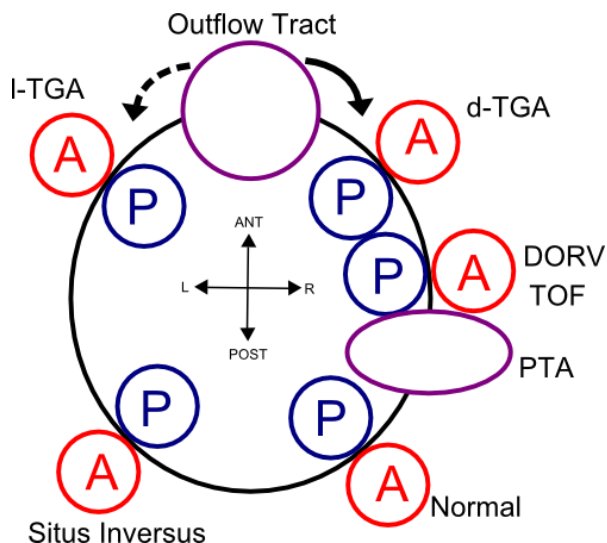


Figure 2.4 Modified Van Praagh diagram, showing rotation of the outflow tract, after (Donnelly & Higgins 1996). As the outflow tract septates into the pulmonary artery (P) and aorta (A) it rotates 150 degrees in normal development, with different degrees of rotation corresponding to different congenital heart diseases. Key: PTA: Persistent Truncus Arteriosus, TOF: Tetralogy of Fallot, DORV: Double Outlet Right Ventricle, d-TGA: dextro-Transposition of the Great Arteries, I-TGA: levo-Transposition of the Great Arteries

In normal heart development, the outflow tract should rotate by 150 degrees as shown in Figure 2.4. While this rotation, or 'wedging' is occurring, the endocardial cushions in the outflow tract are growing and fusing into the conal septum. This divides the outflow tract into the pulmonary artery and aorta. Thus different degrees of rotation correspond to different congenital heart defects. CHDs associated with abnormal rotation of the outflow tract are known as outflow tract (or conotruncal) defects. While these all appear to be related as disruptions of a common developmental mechanism, they can however be defined by very different anatomical characteristics.

For example, the tetralogy of Fallot (TOF) is defined as the association of four anatomic features: overriding aorta (less to the left than it should be), subpulmonary stenosis, ventricular septal defect, and right ventricle hypertrophy. Double Outlet Right Ventricle (DORV) is defined as a condition in which both the pulmonary artery and aorta arise from the right ventricle. However, both these defects correspond to a similar degree of outflow tract rotation (about 90 degrees), and thus could be overlapping in terms of definition and embryological cause. Furthermore, as shown previously, insufficient rotation of the outflow tract could be due to a shortened outflow tract (due to reduced contribution or proliferation of the second heart field), insufficient migration or proliferation of neural crest cells, disrupted left-right signalling or insufficient contribution from the endocardial cushions due to disrupted epithelial to mesenchymal transition (refer again to Figure 2.1).

Thus several developmental mechanisms could have been selected as a starting point for systems modelling of heart development. Epithelial to mesenchymal transition (EMT) was selected for several reasons. There has been a wealth of *in vitro* and *in vivo* experimentation attempting to characterise this system, but very few attempts to formally model it with computer simulations. This means that current knowledge about EMT is highly disjointed. By necessity, each 'wet lab' investigation is restricted to exploring the influence of one aspect of causation. The emergent effects of multiple types of regulation cannot be explored in this way. Thus formalising conceptual models into mathematical and computational models provides a way to represent the system, so that emergent effects may be explored. Furthermore, simulation allows the isolation of

effects, which may be impossible to uncouple in the real system. For example, each gene influences multiple cell level properties, such as adhesion, motility and invasiveness, thus the effect of these properties in EMT cannot be studied individually. When represented as parameters in a cell level simulation, they can be independently controlled. This makes it possible to do some experiments *in silico* that are not possible *in vitro* or *in vivo*.

EMT also provides an example with the potential to form a link between genetic disruption (e.g. a mutation in the Notch signalling pathway), cellular behaviour (EMT) and tissue morphogenesis (endocardial cushions, and subsequently heart valve leaflets and septa). Mutations in the Notch signalling pathway are linked to CHDs such as aortic valve disease, tetralogy of Fallot and Alagille syndrome, and it has been shown that these mutations cause CHD because of impaired EMT (Fischer et al. 2007). Multiscale modelling of the process therefore has the potential to be informative about the effect of genetic regulation on this developmental process, and the potential for disrupted morphogenesis.

2.2.3. Epithelial to Mesenchymal Transition in Heart Development

As outlined in the previous section, Epithelial to Mesenchymal Transition (EMT) underlies the growth of the endocardial cushions, which constitutes a crucial step in the development of the heart. As shown in Figure 2.5, tissue from the endocardial cushions in the Atrioventricular Canal (AVC, blue) becomes the mitral and tricuspid valves, while endocardial cushion tissue in the Conotruncus (CT, yellow, also termed outflow tract) becomes the semilunar valves and the membranous portion of the interventricular septum. Ventricular Septal Defects are the most common type of CHD, and most usually the defect is in the membranous portion (rather than the muscular portion) of the interventricular septum. Defects of the heart valves are among the most common types of CHD such as bicuspid aortic valve (BAV), pulmonary valve stenosis and pulmonary atresia (Armstrong & Bischoff 2004). While people with BAV may not require an intervention as children, calcification and prolapse of heart valves make repair and replacement a widespread intervention later in life.

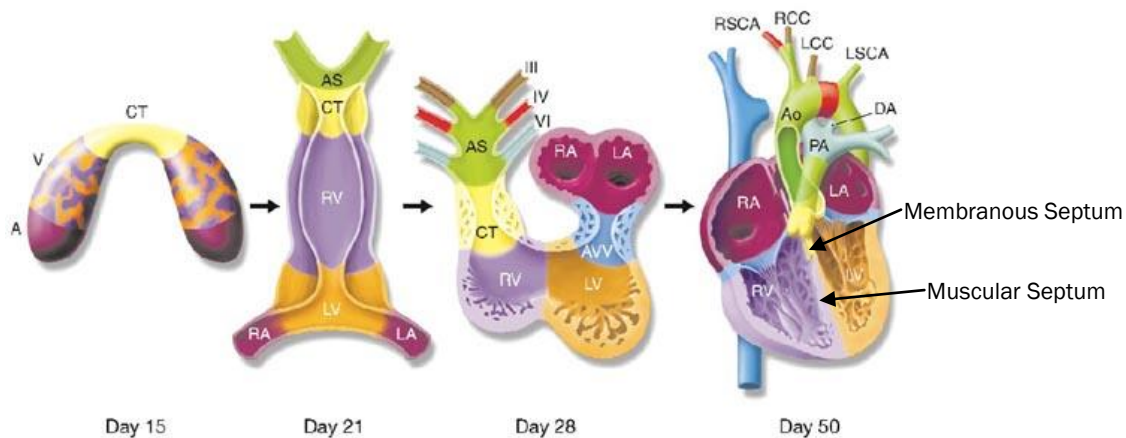


Figure 2.5 Illustration of human cardiac morphogenesis and the redistribution of tissues (Srivastava & Olson 2000)

As illustrated in Figure 2.5, development of tissues in early heart development results in altered structures in quite different places, due the complicated remodelling of the heart. The endocardial cushions, which grow by an EMT process, contribute to some of the most vital structures of a fully-formed functioning heart. These are also the structures that underpin the most common types of CHD, and those responsible for the highest rates of morbidity and mortality, such as ventricular septal defects, and abnormal or missing heart valves. Furthermore, endocardial cushion growth appears to contribute mechanistically to overall heart looping, due to the additional tissue generated at this stage (Bajolle et al. 2009).

EMT in the outflow tract has some differences to EMT in the AV canal. This is partly due to gene expression in the two areas. It is also partly due to the fact that neural crest cells migrate to the outflow tract, and contribute to septation there, but they do not migrate to the AV canal (Hutson & Kirby 2007). The differences between the two areas have recently been demonstrated. Knocking out the *Sur-8* gene in mice leads to reduced EMT in the atrioventricular canal, which leads to hypoplastic endocardial cushions (fewer cells); but the endocardial cushions of the outflow tract are unaffected (Yi et al. 2010).

The primitive heart tube consists of an outer layer of myocardium and an inner layer of endocardium, separated by a thick extracellular matrix termed the cardiac jelly. During endocardial EMT, epithelial cells delaminate from the endocardial layer and invade the

cardiac jelly, adopting a mesenchymal phenotype (Figure 2.6). The delamination is a result of cell signalling interactions between the myocardium, endocardium and cardiac jelly, ultimately resulting in a reduction in the vascular endothelial (VE) cadherin proteins that attach endocardial cells together. Though multiple signalling pathways are coordinated in this process, it is ultimately achieved via an increased expression of Snail family transcription factors (High & Epstein 2008).

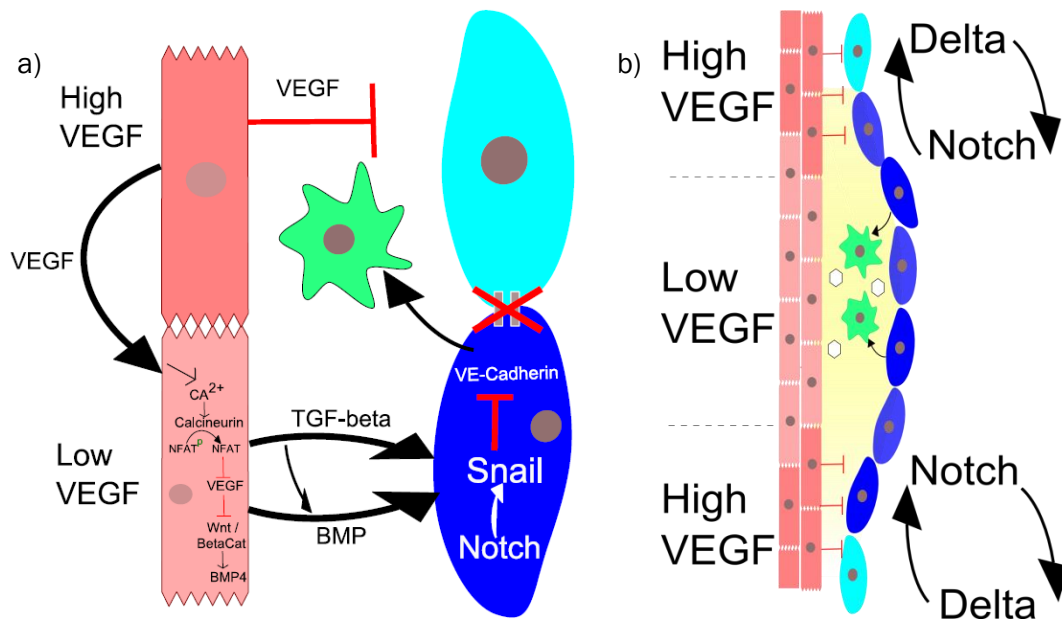


Figure 2.6 a) TGF β and BMP signalling from the myocardium induce Snail expression in the endocardium, which inhibits expression of the endothelial adhesion molecule VE-Cadherin. Activated intracellular Notch has the same effect, and these factors combine to stimulate the endocardial cell to lose its adhesion and adopt a mesenchymal phenotype. b) The cushion forming regions are established by restricted areas of gene expression in both in the myocardium and endocardium.

As shown in Figure 2.6, in the endocardium Notch signalling defines a field of cells that are predisposed to undergoing EMT. A key reason for this is that Notch increases the expression of Snail transcription factors, which inhibit the expression of VE-Cadherin (Vascular Endothelial Cadherin), leading to loss of adhesion between endocardial cells. At the same time in the myocardium VEGF (Vascular Endothelial Growth Factor) expression is controlled to tight levels that enable EMT to take place. It has been suggested that this tight control is necessary, because too high a concentration of VEGF leads to too high a level of endothelial proliferation, which sustains the epithelial phenotype and prevents EMT from happening (Armstrong & Bischoff 2004). On the other

hand, too low a level of VEGF means that there are insufficient new endocardial cells for prolonged EMT, and formation of sufficiently large and cellularised endocardial cushions, as shown in Figure 2.7.

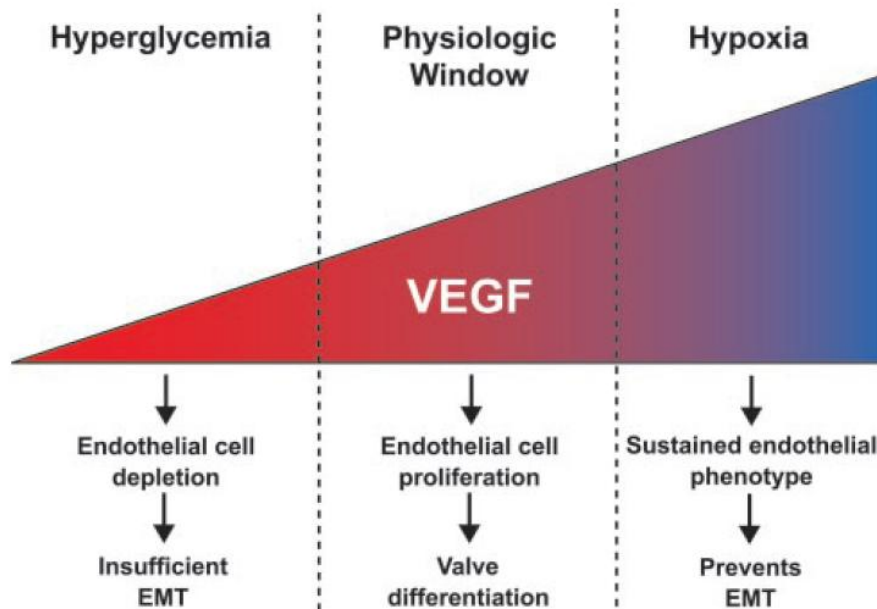


Figure 2.7 Model for VEGF in control of heart valve endothelial cell proliferation (Armstrong & Bischoff 2004)

One of the key mechanisms regulating VEGF signalling is feedback with the NFAT (Nuclear Factor of Activated T-Cells) transcription factors, first in the myocardium, and later in the endocardium. NFATc2, c3 and c4 repress VEGF in the myocardium at mouse E9 (human E24), which is essential for controlling endocardial proliferation and thus allowing EMT to proceed (Lambrechts & Carmeliet 2004). Later, at mouse E11 (human E32), VEGF expression increases in the myocardium, which terminates EMT. Now a second wave of NFATc1 signalling is required in the endocardium of the cushion forming regions to direct valve elongation and remodelling (Chang et al. 2004).

As suggested in Figure 2.8 below, there is a feedback between VEGF and NFAT, which is due to the role of VEGF in controlling the cell's influx of calcium. VEGF signalling increases the influx of calcium through connexin-45 gap junctions. Calcium activates calcineurin in the cytosol, which dephosphorylates NFAT, allowing translocation to the nucleus where it affects gene transcription (Armstrong & Bischoff 2004). There is

evidence for both positive and negative feedback in NFAT signalling. Positive feedback is mediated via the transcription of NFATc1. Negative feedback is mediated via the transcription of DSCR1; which also binds to calcineurin, and therefore competes with phosphorylated NFAT for the calcineurin available (Crabtree & Olson 2002).

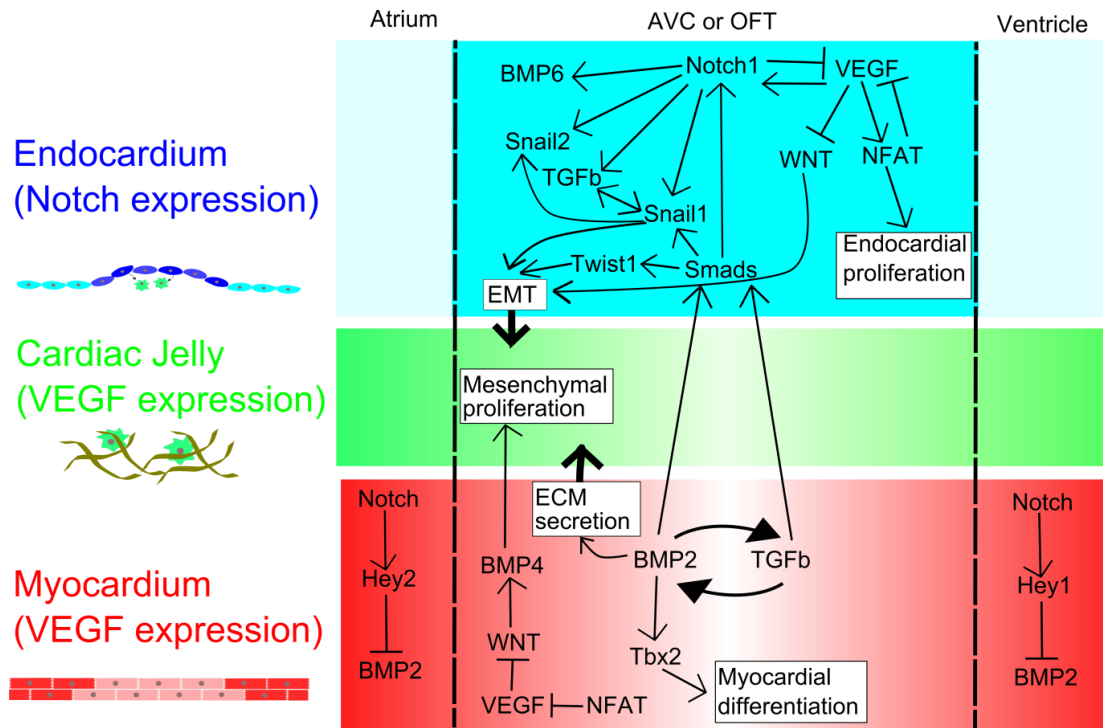


Figure 2.8 Major protein interactions during EMT in endocardial cushion growth. After (Luna-Zurita et al. 2010) and (Wagner & Siddiqui 2007). Gradients show expression pattern for Notch1 (blue, in the endocardium) and VEGF (green in the cardiac jelly and red in the myocardium).

Notch signalling controls the boundary of the cushion forming region in the myocardium as well as the endocardium, however here the expression pattern is reversed, with active Notch being expressed only outside of the cushion forming regions (Rutenberg et al. 2006). As shown in Figure 2.8, the Notch target genes Hey1, in the ventricular region, and Hey2, in the atrial region, act to repress BMP2. Expression of BMP2 upregulates TGFβ protein, which means that both BMP2 and TGFβ are secreted by the myocardium in the cushion forming (Notch-inactive) areas. The BMP2 and TGFβ proteins secreted by the myocardium are essential for the fully invasive EMT phenotype (Luna-Zurita et al. 2010). This demonstrates how endocardial and myocardial signalling are integrated to drive

EMT. Underlying both of these complex signalling networks is a spatial patterning in both tissues, created by Notch signalling.

As is evident from Figure 2.8, which by necessity illustrates only some of the signalling pathways that regulate this EMT process, there is a great deal of complexity and crosstalk between different pathways. This provides a degree of redundancy in the signalling and developmental mechanisms. For example, there is evidence that indicates that while, Snail1 (Snail) protein is a direct target of TGF- β signalling, Snail2 (Slug) is a direct target of the Notch signalling pathways. Furthermore, that Slug acts to repress Snail, most likely through repression of TGF- β related pathways (Niessen et al. 2008). The result of this is that in Snail2 knockout mice, Snail1 expression increases, which compensates for the loss of Snail2 expression; hence there is redundancy built in to the system. This ensures that complex embryonic developmental processes nearly always proceed robustly, and with precise sequencing. It does, however, make understanding the precise function of each gene more challenging, as the effects of silencing a single gene is partly compensated by others.

Notch signalling is a further example of a pathway with built-in redundancy as there are multiple ligands and receptors for Notch signalling. Mammals have four Notch receptors and seven ligands (Bray 2006). Because these are expressed and active in different combinations in different tissues, a genetic deletion of, for example the Jag1 ligand, affects some of the tissues where Jag1 is expressed but not others, due to the redundancies provided by other ligands and receptors (Bolíós et al. 2007).

2.2.4. Notch Signalling in Heart Development

The Notch signalling cascade is evolutionarily highly conserved. Notch-like molecules have been identified in a multitude of diverse species, from *C. elegans* to humans, and appear to play conserved functional roles in development. 'Notch' is the broad term for the transmembrane receptor proteins of this pathway. All vertebrates and mammals have four receptors (Notch1-4) while *D. melanogaster* has only one (Notch), and *C. elegans* has two (LIN-12 and GLP-1). Vertebrates have seven ligands for Notch (Delta1-4, Serrate and Jagged1-2), while *D. melanogaster* has two and *C. elegans* has four.

However, the basic model is the same throughout (Borggreffe & Oswald 2009). Notch signalling differs from most other signalling pathways in the respect that most of its ligands (as well as the receptors) are also transmembrane proteins. This means that signalling is restricted to neighbouring cells that are physically adjacent, making this the canonical example of juxtacrine signalling. What this allows for is very fine control of developmental patterning in tissues, down to the level of single cells.

The basic model of Notch signalling is illustrated in Figure 2.9. A Delta-type ligand binds to a Notch receptor. This activates the receptor, and results in two proteolytic cleavages of the receptor; termed the S2 and S3 cleavage, which involve different enzymes. These cleavages release the Notch intracellular domain (Notch^{Intra}/NICD) which translocates to the nucleus, and interacts with the DNA-binding protein CSL. It achieves this by binding to a co-activator (Co-A) which then replaces the co-repressor (Co-R). In the absence of Notch^{Intra}, CSL associates with a co-repressor, which actively represses the transcription of 'Notch target genes' (Lai 2004).

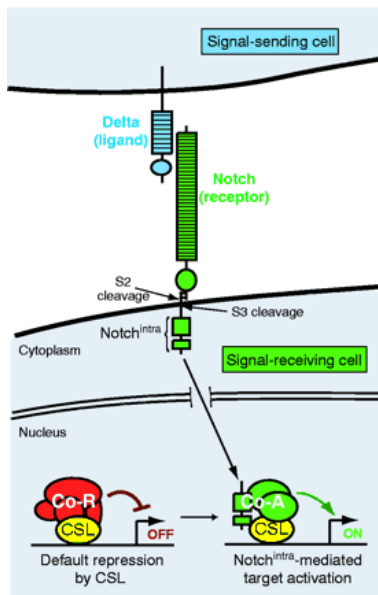


Figure 2.9 Basic operation of the Notch signalling pathway. The key players are a Delta-type ligand, a Notch receptor and the CSL transcription factor (Lai 2004).

While the basic model illustrated in Figure 2.9 appears to be quite simple, it is used in a number of different ways, to achieve different types of developmental patterning.

Furthermore, there are some subtle complexities. These include crosstalk with other pathways, such as the TGF β pathway (Blokzijl et al. 2003), as well as positive and negative feedback loops (Agrawal et al. 2009) and the effects of trafficking ligands and receptors between the cytosol and cell membrane (Bray 2006). Furthermore, while there are only a handful of Notch target genes, the question of what makes a particular gene into a target in a given cell type, in a particular context, is not well understood.

Although Notch signalling operates in a wide variety of contexts, with different combinations of ligands, receptors and target genes involved, we can define some basic modes of signalling that produce different types of tissue patterning. A basic distinction in conceptual models of Notch signalling is made between lateral inhibition and lateral induction. In lateral inhibition, one effect of Notch signalling (and the activation of CSL) is to inhibit the expression of one or more Notch ligands. The result of this is that cells which are initially equivalent will become specialised to either be signal receiving (Notch expressing) or signal sending (Delta-type ligand expressing) cells. A pattern will emerge of signal sending cells surrounded by receptor cells. This happens because lateral inhibition amplifies small differences between cells. The more 'Notch signal' a cell receives, the less signal it can transmit, and thus it becomes specialised to be a Notch receptor cell. Conversely, the more Notch signal a cell sends, the more likely its neighbouring cells are to have Notch receptors on their surfaces, and the more potential the signalling cell has to transmit further signal. The signalling cell also receives a weaker inhibitory signal from its neighbours. Conversely, in lateral **induction**, one of the effects of Notch signalling is to upregulate the expression of a ligand for Notch. This positive feedback generates a contiguous field of cells with the same fate, with Notch signalling and simultaneous expression of receptor and ligand occurring throughout the field. Conversely, loss of Notch signalling leads to downregulation of the ligand throughout the field. When inductive signalling operates between areas of initially non-equivalent cells, the boundaries between these areas are reinforced (Artavanis-Tsakonas et al. 1995).

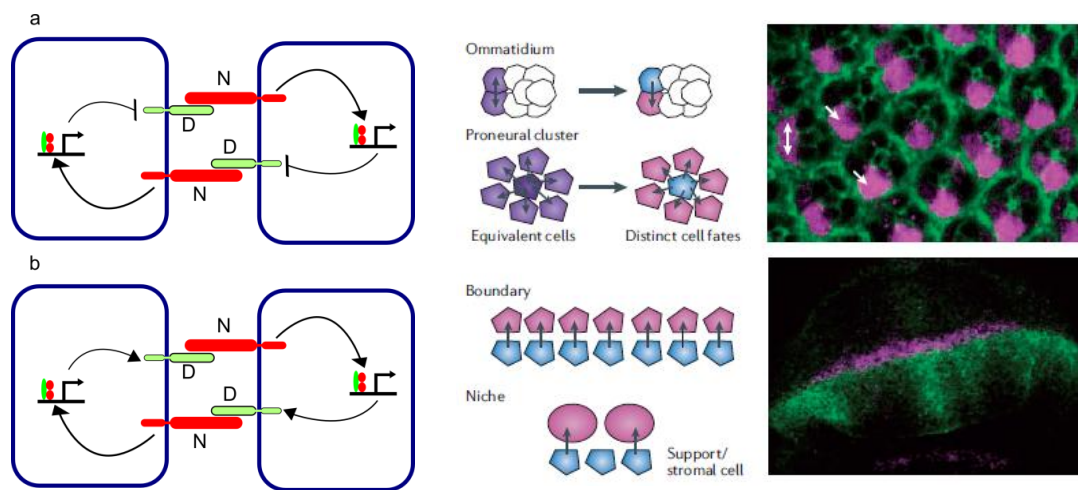


Figure 2.10 Mechanisms of Notch juxtacrine signalling between two adjacent cells, and examples of resulting tissue level patterning for (a) lateral inhibition and (b) lateral induction (N: Notch receptor, D: Delta-type ligand). In the central figures, Notch signalling is represented by black arrows. In lateral inhibition, initially equivalent (purple) cells are resolved into distinct fates (blue and pink); cells with the highest Notch activity (receptor cells) coloured pink. Confocal scan (upper right) illustrates the case of ommatidium, with cell membranes stained green and Notch expression pink. In lateral induction, a boundary is established to segregate and/or organise two groups of cells. Black arrows indicate direction of Notch signalling, with pink cells having Notch activation. The confocal image (lower right) is of the fly wing primordium, where Notch activity (pink) is measured at the boundary of ligand (Serrate) expressing cells (green). After (Bray 2006)

Examples of fine-grained patterning mediated by Notch lateral inhibition include vertebrate neurogenesis (Cau & Blader 2009), control of stem cell function (Lewis 1998) and the development of sensory hairs in the mammalian inner ear (Lanford et al. 1999). Lateral induction of Jagged1 is involved in differentiation of lens fibre cells (Saravanamuthu et al. 2009). There is evidence to suggest that lateral induction also plays a role in the assembly of the arterial walls, via upregulation of Jag1 in successive layers of smooth muscle cells (Hoglund & Majesky 2012). Lateral induction also plays a role in synchronising the oscillations between cells during somitogenesis (Giudicelli et al. 2007), and in boundary formation of margins of the *Drosophila* wing (Bray 1998). While most sources suggest that the role of Notch signalling in somitogenesis is to synchronise the oscillations of presomitic mesoderm cells, it has been suggested that the Notch pathway itself may also act as an oscillator in this context. Furthermore, that it can produce either a continuous or transient signal in different contexts (Agrawal et al. 2009). These possibilities are out of the scope of this review.

In some cases, lateral inhibition and lateral induction are known to be combined for developmental patterning. For example, in the chick inner ear, prosensory patches are initially specified by lateral induction. Subsequently lateral inhibition within each patch generates a fine grained pattern of hair cells, surrounded by supporting cells. In such cases, different ligands or receptors may be used in the different patterning mechanisms. In the case of the chick inner ear, the ligand Serrate2 is expressed throughout the prosensory patches, while Delta2 is later expressed a fine-grained pattern of individual hair cells within the patch (Daudet & Lewis 2005).

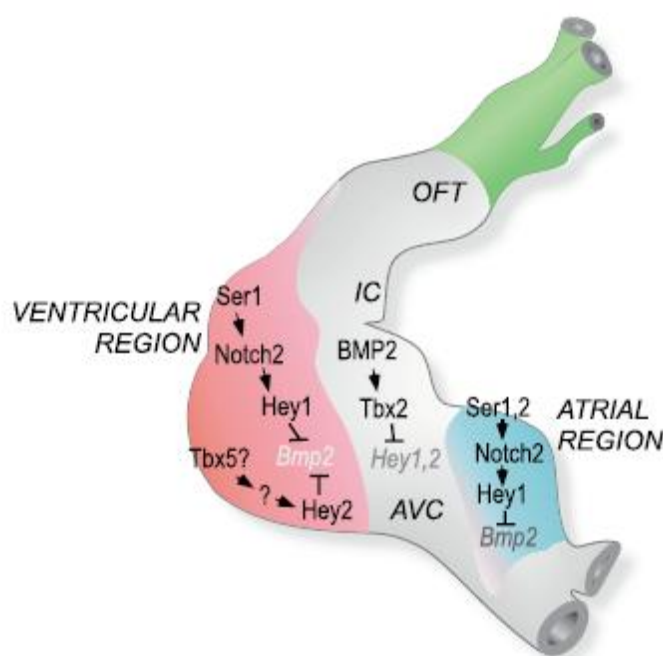


Figure 2.11 Developmental patterning of the myocardium of the embryonic heart tube by Notch and Hairy-related transcription factors (Hey1 and Hey2) (Rutenberg et al. 2006).

It has been found experimentally that lateral induction is the Notch signalling that patterns the embryonic heart, at least at the stage of the endocardial cushions (Timmerman et al. 2004). The ligands expressed throughout the endocardium are Jagged1 and Delta4, and the receptors are Notch 1, 2 and 4. Though it is not yet fully clear the exact roles played by these different ligands and receptors, it is clear that Delta4/Notch1 interactions lead to the expression of Snail, which downregulates adhesion (Bolós et al. 2007). The mechanism of lateral induction initially patterns the

heart into three sections: a ventricular region, an atrial region and a region that comprises the outflow tract (OFT) and atrioventricular canal (AVC), which together are the cushion-forming regions of the heart tube, as shown in Figure 2.11. This occurs both in the myocardium; with Notch2 signalling suppressing BMP2 expression in non-cushion-forming regions, and in the endocardium, with Notch1 inducing the expression of Snail family proteins.

The Snail family of genes can be seen as central determinants in the process of EMT. The downstream targets of Snail gene expression lead to loss of epithelial markers, gain of mesenchymal markers, as well as changes in cell shape and movements that characterise invasion. These include upregulation of Matrix Metalloproteinases (MMPs), which are capable of degrading proteins in the extracellular matrix (ECM) when secreted by cells, enabling them to invade tissues. Snail genes also regulate cell proliferation and cell death, such that cells tend to proliferate less but survive longer; allowing mesenchymal cells to migrate to distant sites in some cases (such as the neural crest cells). These effects are summarised in Figure 2.12 (Barrallo-Gimeno & Nieto 2005).

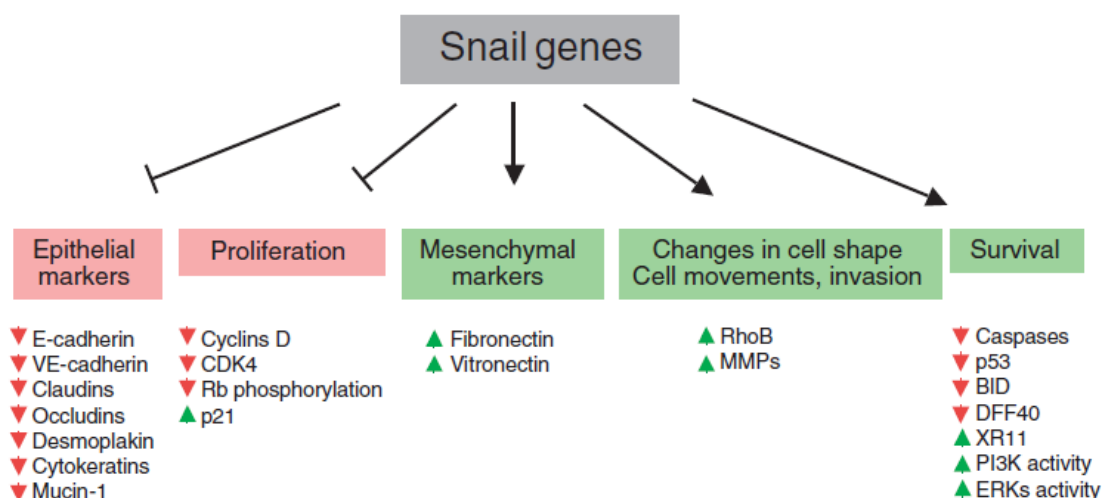


Figure 2.12 Downstream targets of Snail genes. Molecules and processes shown in red are downregulated, while those in green are upregulated (Barrallo-Gimeno & Nieto 2005).

The Snail family function as repressors and the phenotypic outcomes are a combination of direct (repression) or indirect (repression of a repressor) regulation of target genes. For example, the EMT that leads to formation of the endocardial cushions includes:

repression of VE-cadherin (loss of endocardial cohesion), repression of repressors of integrins such as fibronectin (increase in cell-ECM adhesion and strengthening of the cytoskeleton) and repression of repressors of MMPs (ECM degradation).

The Snail family genes Snail1 (also known as Snail) and Snail2 (also known as Slug) are expressed in both endocardial and mesenchymal cells in the locations where the endocardial cushions develop, in both mouse and human (Niessen et al. 2008). Snail1 and Snail2 are likely to have somewhat different effects in terms of the extent to which they upregulate and downregulate different molecules, however it is not yet known what these differences might be. It has been demonstrated that a deletion of Snail2 in mice is compensated for by an increase in the expression of Snail1 (Niessen et al. 2008). This suggests that Snail2 acts to suppress the TGF β pathway activation of Snail1. As previously indicated, Snail2 is a direct of Notch signalling, while Snail1 is a direct target of TGF β signalling. Snail1 is also synergistically induced in endothelial cells when both Notch activation and TGF β stimulation are present (Niessen et al. 2008). This tentatively suggests the qualitative network shown in Figure 2.13.

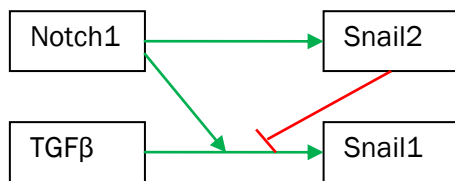


Figure 2.13 Qualitative network of Notch and TGF β synergy in activating Snail gene expression in endothelial cells, derived from results of (Niessen et al. 2008). Snail1 is a direct target of TGF β while Snail2 is a direct target of Notch. Notch and TGF β synergistically upregulate Snail1, and Snail2 inhibits TGF β mediated Snail1 activation.

Lateral induction Notch signalling operates to define both the endocardial regions predisposed to EMT, and the myocardial regions that secrete the necessary proteins (including TGF β and BMP2). Notch signalling lateral induction is thus fundamental to the process of EMT in the developing heart. In the endocardium, Notch1 signalling specifies the domain of cells capable of undergoing EMT (by inducing Snail2 expression). Simultaneously in the myocardium, Notch2 signalling represses the expression of BMP2 and TGF β *outside* of the cushion forming regions. The TGF β secreted by the myocardium *inside* the cushion forming region induces Snail1 expression in the endocardium. BMP2

has a common pathway with TGF β signalling via the mediator SMAD4 (Wagner & Siddiqui 2007). However, the effects of BMP2 have been shown to be somewhat different from those of TGF β . BMP2 induces 3D invasion into a collagen gel for *in vitro* endocardial cells, while TGF β treatment leads merely to 2D migration on the surface of the collagen gel (similar to the effects of Notch activation) (Luna-Zurita et al. 2010). Luna-Zurita et al. (2010) suggest that this is because BMP2 induces Twist1, which inhibits GSK3 β , resulting in stabilisation of Snail1. This suggests that it is the sustained expression of Snail1 that induces migratory behaviour. Though it is not clear what mechanisms lie behind this, possible actions include a delayed upregulation of cell-ECM adhesion or delayed expression of MMPs for matrix degradation. Delayed response of cell-ECM adhesion is quite plausible, as the strength cell-ECM adhesion increases rapidly over the initial hours of integrin binding, before reaching a steady state (Gallant et al. 2005).

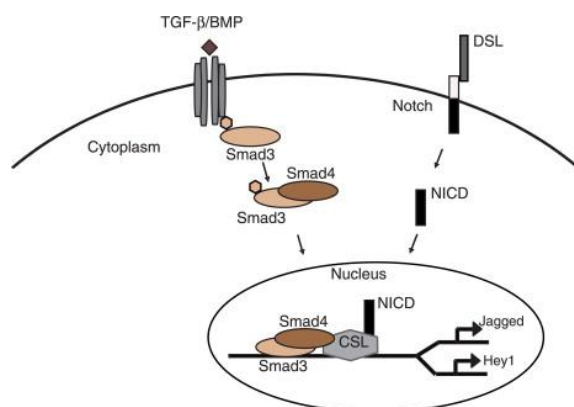


Figure 2.14 Mechanism for the integration of TGF β and Notch signalling. By binding to SMAD3, NICD recruits the SMAD3/SMAD4 complex to the CSL domain. The two pathways cooperatively activate Notch target genes (Heitzler 2010).

TGF β is known to have both synergistic and antagonistic interactions with Notch signalling in different cellular contexts (Blokzijl et al. 2003; Zavadil et al. 2004; Sun et al. 2005). In endothelial cells, it has been found that Notch and TGF β signalling synergistically induce the expression of Snail genes (Niessen et al. 2008). Notch Intracellular Domain (NICD) has been shown to have interaction with TGF β signalling by binding to the SMAD3/SMAD4 complex, which then cooperatively activates Notch target genes (Blokzijl et al. 2003). This model is illustrated in Figure 2.14. Notch has different effects on the target genes of TGF β in endothelial cells. Specifically, Notch activation

decreases the expression of Smad1 and Smad2 and their target genes, but increases the expression of Smad3 and its target genes (Fu et al. 2009).

In summary, endocardial and myocardial signalling are integrated during endocardial EMT, and underlying both is an expression pattern formed by Notch lateral induction (Notch1 and Notch2 for the endocardium and myocardium respectively). Mutations in Notch1, Notch2 and Jag1 are all known to cause Alagille syndrome in humans, which includes the heart defects tetralogy of Fallot and pulmonary valve stenosis (Butcher et al. 2011; McDaniel et al. 2006). There is thus an unambiguous link between genetic mutation, impaired EMT and the development of congenital heart defects. This complete link has been demonstrated through an *in vivo* investigation in which mice with knockouts of Notch target genes were shown to develop congenital heart defects; that impaired EMT was apparent in the endocardial cushions of knockout mice; and that EMT of atrioventricular explants from knockout mice was impaired *in vitro* (Fischer et al. 2007).

2.3. Cell Level Modelling Techniques

There is some precedent for bulk cell migrations, particularly those that occur during cancer metastasis, to be represented with continuum models (M. A. J. Chaplain & Lolas 2006; Painter et al. 2010). There are certain advantages to this approach. Continuum models are generally quite simple, and contain only a few parameters. A large area can be represented at relatively low computational cost. It is possible to include cell proliferation and death, and to have this depend on a spatially varying field. This could be quite appropriate for modelling EMT at the tissue level, such as the growth of the endocardial cushions. However, the focus of this thesis lies in modelling cellular properties, such as cell size, morphology and motility, explicitly. These features are essential to the EMT process, and thus it is evidently more appropriate to use some form of agent based modelling, in which cells are represented as discrete entities.

There are three main types of agent based modelling that can be used for representing cell behaviour: cell centre, cell vertex and cellular Potts models (Figure 2.15). In cell centre models, cell centres are represented as points in space. Forces are defined as acting between these centres, such as damped spring forces. Connectivity between cells,

as well as cell shapes, can be represented by Voronoi tessellation (Meineke et al. 2001). In cell vertex models cells are represented as adjacent polygons. The dynamics of each cell are represented by the movement of its vertices. This movement can either be determined by explicit forces on each vertex or by a free energy function as a result of a balance of forces between cells. It is thus possible to have direct representation of mechanical forces such as plasticity, elasticity and viscoelasticity of cells, as well as incorporating differential adhesion terms (Walter 2009). Finally, in cellular Potts models cells are represented as multiple sites on a lattice. Cell dynamics are represented by pixel copy attempts between adjacent lattice sites belonging to different cells. These are accepted stochastically with a probability that depends on a free energy function, allowing differential adhesion and properties of cell morphology and migration to be modelled.

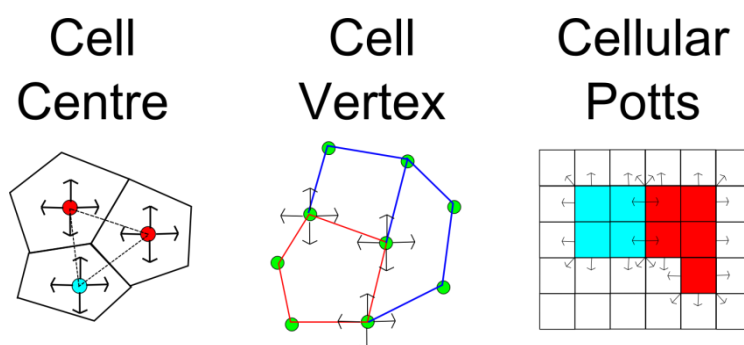


Figure 2.15 Cell centre, cell vertex and cellular Potts models

The most essential requirements for cell-level modelling of endocardial EMT are a good representation of cell morphology, adhesion, movement, and the ability to extend these features into a 3D model. Other important features include: the ease of developing multiscale models that include subcellular reaction models within each cell; computational efficiency (including consideration for the efficiency of multiscale or 3D simulations); the ease of programming each model; and the explicit representation of biomechanical behaviour such tension and pressure forces and the stiffness of cell walls. Nice-to-have features include polarisation of epithelial cells (so that loss of polarisation might be included), cell growth, and the representation of intracellular filaments. To summarise the key properties of each of the three cell-level modelling techniques, as well as continuum modelling, these are presented as a Pugh matrix in

Figure 2.16, with an importance rating assigned to each of the features. This illustrates the point that some of the modelling techniques are clearly more appropriate to this task than others; but no model is perfect, or better in terms of all the requirements than every other technique.

	Importance Rating	Continuum	Cell Centre	Cell Vertex	Cellular Potts
Cell Morphology	10	-	S	+	+
Adhesion	10	-	S	+	+
Movement	10	S	S	S	+
Extension to 3D	10	+	S	-	+
Division	9	-	+	+	+
Ease of Multiscale Coupling	9	-	S	S	+
Ease of Programming	8	S	S	S	+
Computational Efficiency	8	+	-	S	+
Biomechanical Behaviour	7	-	-	+	-
Polarisation	5	-	-	+	-
Growth	4	-	+	+	+
Intracellular Elements	2	-	-	S	-
Weighted Score		-38	-9	35	64

Figure 2.16 Pugh matrix for cell to tissue level modelling methods for EMT, based on modelling realism, computational efficiency and ease of implementation. Green + = strength, red - = weakness, yellow S= neither strength nor weakness.

The advantages of the cell centre model include that it is able to model cell proliferation and migration, as well as the social behaviour of cells such as contact inhibition to some extent (Walker et al. 2010). However, the representation of cell shape by Voronoi tessellation is both computationally expensive and unrealistic. It is computationally expensive due to the need to construct a new Voronoi diagram during each time step. This would especially be the case in a 3D model, and indeed most 3D cell centre models do not investigate cell shape, and represent cells as spheres in three dimensional space (Ramis-Conde et al. 2008; Walker et al. 2010). The question of how to represent the

shape of cells that lose connectivity to other cells (as occurs during migration in EMT) would also need to be resolved. Voronoi tessellation lacks realism because the shapes of cells, and their number of vertices, do not change smoothly during a simulation. Furthermore, there is no way to set parameters that directly control cell morphology in cell centre models. The validity of representing cell-cell interactions as spring forces might also be questioned, as adhesion forces are not dependent on the contact area between cells. However, there are examples of cell centre models that do set the adhesion force as a function of the contact area between cells (Galle et al. 2005). There have been examples of multiscale cell centre models (Ramis-Conde et al. 2008; Walker et al. 2010). The Chaste simulation package (Bernabeu et al. 2009) provides a platform for developing multiscale cell centre and cell vertex models. Implementing multiscale EMT models in Chaste would still require fairly extensive development, however, as the multiscale functionality currently provided focuses on incorporating CellML for multiscale models of cardiac electrophysiology.

The advantages of the cell vertex model include a good representation of differential cell-cell adhesion. Cell shapes and tissue patterns change smoothly and provide a good representation of real cell morphology (Honda et al. 2004). It is possible to use either explicitly calculated forces on each vertex, or to have cells move down gradients of a free energy function (Walter 2009). Where using explicit forces, it is possible to include a cell pressure and membrane tension force that will ensure cells tend to a particular volume and surface area (Weliky & Oster 1990). Where using a free energy function, it is possible to include a cell-cell adhesion energy, which is proportional to the contact area between cells (e.g. the length of the edge between two cells). Deformation energies can also be included so that cells tend to a particular volume or surface area (Nagai & Honda 2006). The free energy function approach is comparable to the dynamics of the cellular Potts model; however vertices move deterministically rather than stochastically. While cell vertex models have been extended to 3D, this has only been used for representing epithelial cells on a hollow 3D surface (Trichas et al. 2012) or as space filling polyhedra in a cell aggregate (Honda et al. 2004). The issue of using the vertex model to represent cells that migrate away from other cells has not been addressed. This means that the existing 3D extensions are not fit for our purposes. Vertex models have the advantage

that biomechanical forces acting on the cell walls can be modelled and measured explicitly. Extensions of the vertex model, such as the subcellular element model (Newman 2005), have the capability to also represent intracellular filaments. Although cell vertex models have the advantage of having no need for a Voronoi diagram or Delaunay triangulation to be produced at each timestep, they often contain more information, and so can be more computationally expensive. They are also comparatively computationally expensive compared to cellular Potts models, as vertex models are lattice-free, while cell movements in Potts models are restricted to a lattice. This can mean that cell movements in vertex models are smoother than Potts models.

The Potts model has the advantages that it is simple to program and computational inexpensive to run, due to it being a lattice-based model. This also means, however, that a high spatial resolution is required for realistic cell shapes (Walter 2009). Extension into three dimensions is very straightforward. A number of multiscale cellular Potts models have been developed recently, notably using the CompuCell3D simulation platform (Hester et al. 2011; Swat et al. 2012). This provides a framework for loading SBML models to individual cells and solving them as ODEs during a cellular Potts simulation. Cell level parameters can be set as functions of SBML model variables, and vice versa, providing a strong coupling between the models, while keeping their declarative specifications separate. It is quite straightforward to represent all of the main requirements of 2D and 3D EMT modelling with cellular Potts models. Terms can be added to the free energy function such that cells will tend toward particular volumes or surface areas. Differential adhesion is included as a core mechanism in the dynamics of cell movement in the model. As the 'medium' (*in vitro* collagen gel or extracellular matrix) is also represented as a set of lattice points, cells are able to move independently through the medium. Thus a representation of cell migration and transformation is uncomplicated. The main disadvantage of the cellular Potts model is no physical realisation of time, as time is measured by the number of Monte-Carlo steps (MCS). This means that a little more effort is required to fit the MCS of a particular model to an equivalent 'biological time' and to ensure that the space and time scales correspond to this (Hester et al. 2011). The cellular Potts model also does not have a representation of biomechanical forces acting on individual cell walls. As the cellular Potts model meets all

of the essential requirements for modelling EMT (Figure 2.16) this formalism is used for the cell and tissue level simulations in this thesis as reported in Chapters 3 and 4. A combination of 2D, 3D and 2D multiscale cellular Potts models are used for different mechanisms within endocardial EMT. Furthermore, by adding a plastic coupling term to the free energy function as a breakable spring force between cell centres, an aspect of the cell centre model is included. This allows investigation of the interactions between cell morphology, weak labile adhesion and strong plastic coupling (assumed to act between cell centres). This methodology is explained further in Chapter 3.

Various types of epithelial to mesenchymal transition have been previously modelled using variations of the three different agent based simulation methods shown previously in Figure 2.15. Wound healing EMT has been modelled using a cell centre model (T. Sun et al. 2009), as has cancer metastasis EMT (Ramis-Conde et al. 2009). These models both incorporated subcellular reaction models, and represent cell adhesion as a breakable spring force between cell centres. Endocardial EMT has been modelled using the cellular Potts method (Neagu et al. 2010). However, this latter example did not investigate the role of cell morphology during EMT, which is an issue addressed in this thesis. A review of existing cell and tissue level models of EMT follows in the next section.

2.4. Existing Simulations of EMT

2.4.1. 3D Cellular Potts Simulation of Endocardial Cushion Growth

A number of simulations of EMT have been developed, using a variety of cell level modelling techniques. The majority of these simulations focus on EMT as reactivated during cancer metastasis. To our knowledge, there is only one existing example of a simulation model of endocardial EMT (Neagu et al. 2010). This is a 3D cellular Potts model, and includes four ‘particle types’: endocardial cells, mesenchymal cells, extracellular matrix and ‘medium’. In this case, the medium represents the lumen of an embryonic heart tube.

This model provides a conceptual representation of an *in vivo* endocardial cushion. Due to the complexity of processes involved, it is not directly fitted to experimental data, but rather provides qualitative results of the interplay of mechanisms that drive endocardial cushion growth. The mechanisms included within the model are: cell movement (driven

stochastically by differential adhesion), EMT (conversion of an endocardial cell to a mesenchymal cell), mesenchymal cell proliferation, endocardial cell proliferation and production of ECM by mesenchymal cells.

Rather than investigating the role of cell-level factors, such changes in adhesion or cell motility, that drive EMT; this model focuses on the role of EMT and other factors (cell proliferation, adhesion and matrix production) in driving endocardial cushion growth. Thus it assigns probabilities of EMT occurring, and of proliferation or matrix production occurring during a simulation step. The model does not suggest mechanisms that control these probabilities, but rather the parameter space is explored to determine the relative importance of different mechanisms. Thus cell shape and deformability are not important within the model, and cells are represented as individual voxels. The model indicates that an increase of cell-ECM adhesion is more efficient at driving endocardial cushion growth than a decrease in cell-cell adhesion. Interestingly, this conclusion matches that of cellular Potts simulations described in this thesis (see Chapter 4), which take an approach of modelling the influences of cell shape, motility and adhesion in driving EMT.

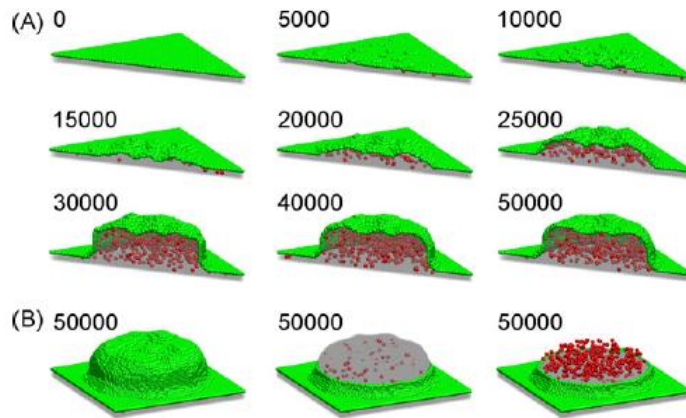


Figure 2.17 3D cellular Potts model of endocardial EMT (Neagu et al. 2010). (A) EMT is assumed to be restricted to a circular area with a diameter of 30 cells, and takes place during the first 3×10^4 MCS. Probabilities are assigned for matrix production and cell proliferation. (B) The final state assumed at 5×10^4 MCS; showing full view, epithelium removed, and extracellular matrix removed respectively. Key: Epithelium: green, ECM: grey, mesenchymal cells: red.

The myocardium is not included explicitly within this model, although is included implicitly by the assumption of signalling proteins that would induce EMT within a regionally restricted area of the endocardium. The model demonstrates that the combination of regionally restricted EMT, cell proliferation, differential adhesion and ECM production are sufficient to produce structures consistent with the development of endocardial cushions, as shown in Figure 2.17.

2.4.2. 2D Cellular Potts Simulation of Active Cell Migration & Mechanical Equilibrium

A recently published extended cellular Potts model focussed on cell invasion from an aggregate (Szabó et al. 2012). Rather than specifying the developmental or disease context, the model described the behaviour emerging from autonomous cell motility, changes in cell-cell adhesion, contact guidance by ECM filaments and the ability of cells to degrade the ECM. Rather than using random motility, the model assumed a positive feedback to cell polarity from cell displacements exists; on the basis that cell invasion from multicellular aggregates has been observed to be persistent *in vitro* (Rupp et al. 2008).

The authors first explain the adaptations of the cellular Potts model to enforce mechanical equilibrium within a cell mass. While the standard cellular Potts model represents each individual cell as a droplet of viscoelastic fluid, the adapted model accounts for bulk dynamics by additionally treating an aggregate of cells as a larger mass of viscoelastic fluid. They demonstrate that only accounting for local mechanics produces unrealistic results in the context of a cell mass adhering to a moving adhesive surface (Figure 2.18). They further demonstrate that unrealistic results are obtained in the case of cells adhering to simulated immutable ECM fibres. While cells are able to invade the ECM filaments, they leave gaps behind, as cells in the bulk are unable to adjust quickly enough to the changing conditions (Figure 2.19).

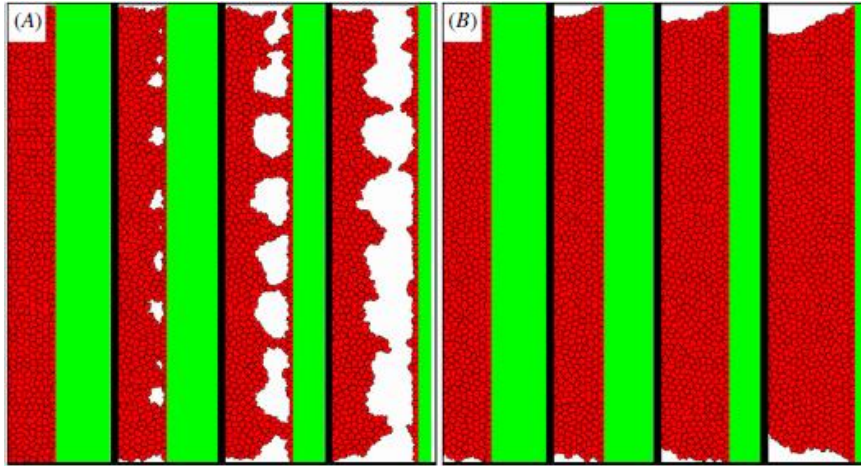


Figure 2.18 The 'piston test' demonstrates the shortcomings of strictly local dynamics of the cellular Potts formalism. Cells (red) adhere to an ECM surface (green) which is shifted 1 pixel every 100 MCS. (A) With the standard cellular Potts model, cells cannot adjust to the changing boundary and holes appear near the attachment surface. The holes expand, and eventually the cells separate from the adhesive surface. (B) By including mechanical relaxation, whereby the aggregate of cells is given the same dynamics as each individual cell, the cells are able to adjust; matching the behaviour of real cell aggregates (Szabó et al. 2012).

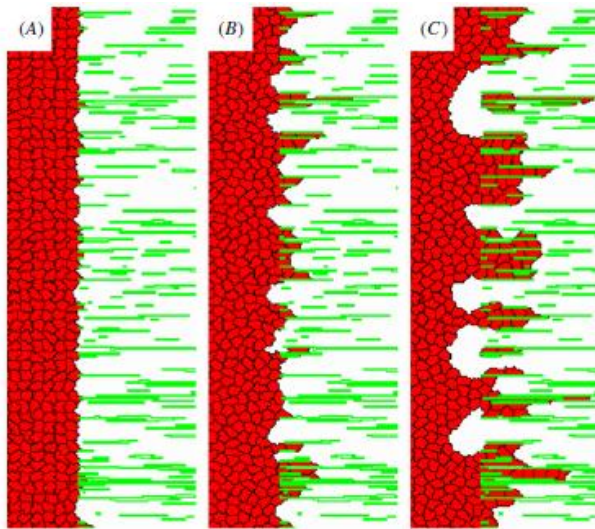


Figure 2.19 Bulk cell movement at the ECM-cavity interface is unrealistic in the standard Potts model. While cells that adhere to the ECM are able to invade the space between ECM filaments; cells in the bulk are unable to adjust fast enough. Thus gaps form, and the average cell density is reduced at the cavity-matrix interface (Szabó et al. 2012).

The authors then use the extended model to investigate the effects of cell-cell adhesion, cell-ECM adhesion, persistent cell motion, the presence of an oriented matrix, and matrix degradation on cell invasion. While the assumptions of a highly oriented and rigid ECM is

unrealistic, and the dimensions are not fitted to a realistic geometry, the model provides a means to evaluate the effectiveness of contact guidance under idealised conditions.

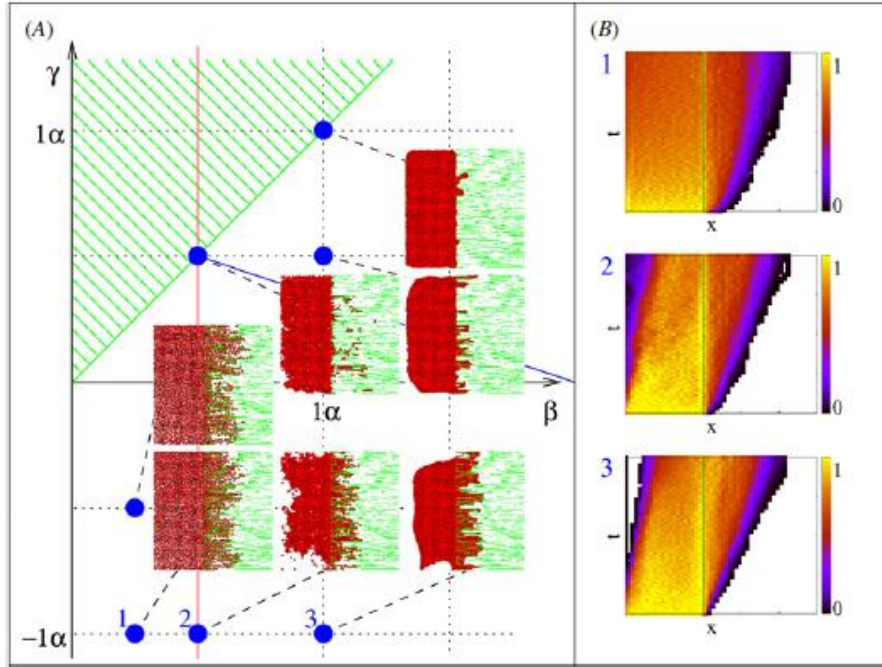


Figure 2.20 2D extended cellular Potts model of cell invasion from an aggregate (Szabó et al. 2012). Cells with spontaneous motility invade an anisotropic, aligned and immutable ECM environment. (A) Morphology diagram of the invasion process for various values of the parameters β (cell-cell adhesion) and γ (cell-ECM adhesion). Configurations shown were obtained at $t = 1000$ MCS, which represents 17 hours. (B) Cell density changes along the direction of invasion reveal a steady invasion speed. The colour code indicates cell density compared to confluency. Temporal and spatial ranges on vertical and horizontal axes respectively are 2000 MCS and 500 lattice units. Plots annotated as 1, 2 and 3 correspond to the parameter values marked accordingly in (A).

The model demonstrates that the presence of an oriented ECM indeed increases cell invasion, as does the level of cell-ECM adhesion; while a high level of cell-cell adhesion increases the probability of cells invading in chords, as shown in Figure 2.20. The ability of cells to degrade ECM enables them to form new channels in the matrix, increasing their potential for invasion.

2.4.3. Multiscale 3D Cell Centre Model of Cancer Cell Invasion

This model illustrates how cell adhesion can be regulated by interactions between E-cadherin and β -catenin, and in turn how cell adhesion is related to cell migration (Ramis-Conde et al. 2008). It is a multiscale model, and uses the cell centre formalism in combination with a system of differential equations that are assigned to each cell. The

subcellular model embodies the molecular kinetics of E-cadherin and β -catenin. The model makes a fundamental assumption that an increase in free β -catenin is a primary determinant in activating cell migration. Dissociation of E-cadherin adhesive bonds between cells releases β -catenin, while the formation of E-cadherin bonds sequesters β -catenin to the adherens junction. This suggests a mechanism whereby loss of cell-cell adhesion induces an increase in cell migration. In this scenario, a loss of cell-cell adhesion is mediated by the disassembly of E-cadherin based junctions, which activates downstream targets via β -catenin/Wnt signalling, including mechanisms that increase cell migration. The model does not suggest specifically what these mechanisms are (e.g. an increase in motility or an increase in cell-ECM adhesion), but represents them as chemotaxis (migration towards a chemical source of attractant).

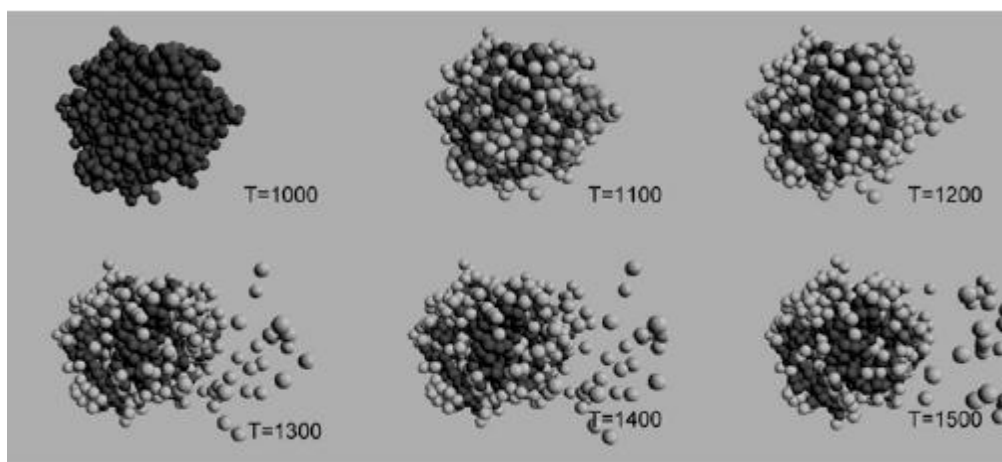


Figure 2.21 3D cell centre simulation of cancer cell invasion, toward a source of morphogen to the right of the tumor (Ramis-Conde et al. 2008). Cells detach gradually as the intracellular concentration of β -catenin is upregulated (light grey). Unit of time is given in minutes.

The model is applied to a generalised case of EMT in an epithelial layer. This illustrates that a single cell undergoing detachment induces an outwardly moving detachment wave in neighbouring cells. As E-cadherin junctions break, β -catenin is released, inducing migration towards the attractant. It is also applied to a model of cancer cell invasion (Figure 2.21). This illustrates that cancer cells at the periphery of a tumour, having fewer E-cadherin bonds and thus a higher concentration of soluble β -catenin, have the potential to detach and migrate towards a source. This suggests that cancer cell

migration is a gradual process, with subsequent layers of cells detaching from the tumour surface.

2.4.4. Multiscale 3D Cellular Potts Simulation Cancer Cell Growth and Invasion

The model of Ramis-Conde et al. (2008) described in the previous section has been re-implemented as a multiscale 3D cellular Potts model (Andasari et al. 2012). This utilises CompuCell3D for the lattice based cell level modelling and the Bionetsolver API for integrating SBML encoded models of the E-cadherin/ β -catenin dynamics. Re-creating a model using a different cell-level modelling formalism provides a means of cross validation, by ensuring that different techniques provide at least a qualitative agreement. In this case, the wave of detachment produced in the Potts implementation is uneven, due to the stochastic nature of cell dynamics in cellular Potts models, and fluctuations in contact area between cells. This contrasts with the cell centre implementation in which a very regular wave of detachment is produced. The particular nature of the detachment waves (regular or randomised) is an artefact of each modelling technique. However both methods at least produce qualitative results that include waves of detachment. Reproducing results in this way helps to demonstrate which effects are artefacts of the model and which are true properties of the system being represented.

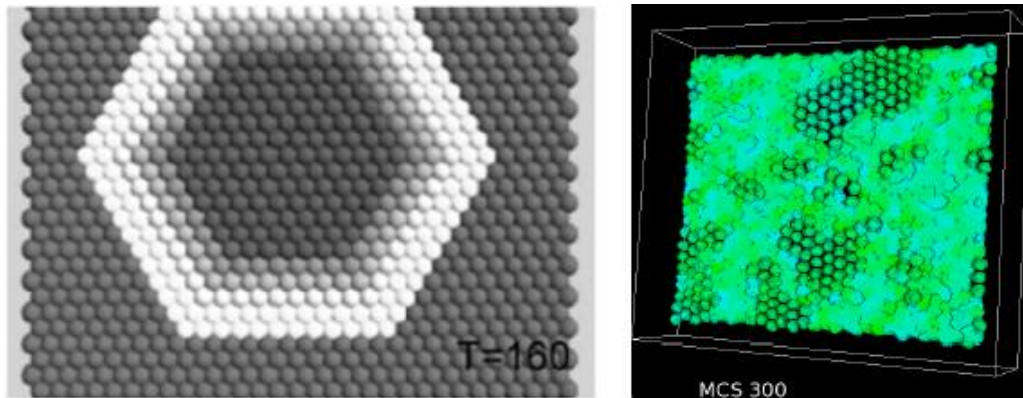


Figure 2.22 Comparison showing the difference in β -catenin detachment wave simulations between centre based model of Ramis-Conde et al. (2008) (left figure) and the CompuCell3D- Bionetsolver model of Andasari et al. (2012) (right figure).

Like the model of Ramis-Conde et al. (2008), this model is applied to generalised detachment and EMT from an epithelial layer, as well as the representation of tumour growth and cancer cell invasion (Figure 2.23). A radial chemoattractant gradient is used

to cause cells to migrate outwards in all directions, following detachment. However, this simulated chemotaxis serves as an analogue for any combination of cell characteristics that induce migration: such as active cell motility or increased cell-ECM adhesion. The model does not include any representation of changes in cell shape, which have been widely observed to occur during EMT (Mendez et al. 2010).

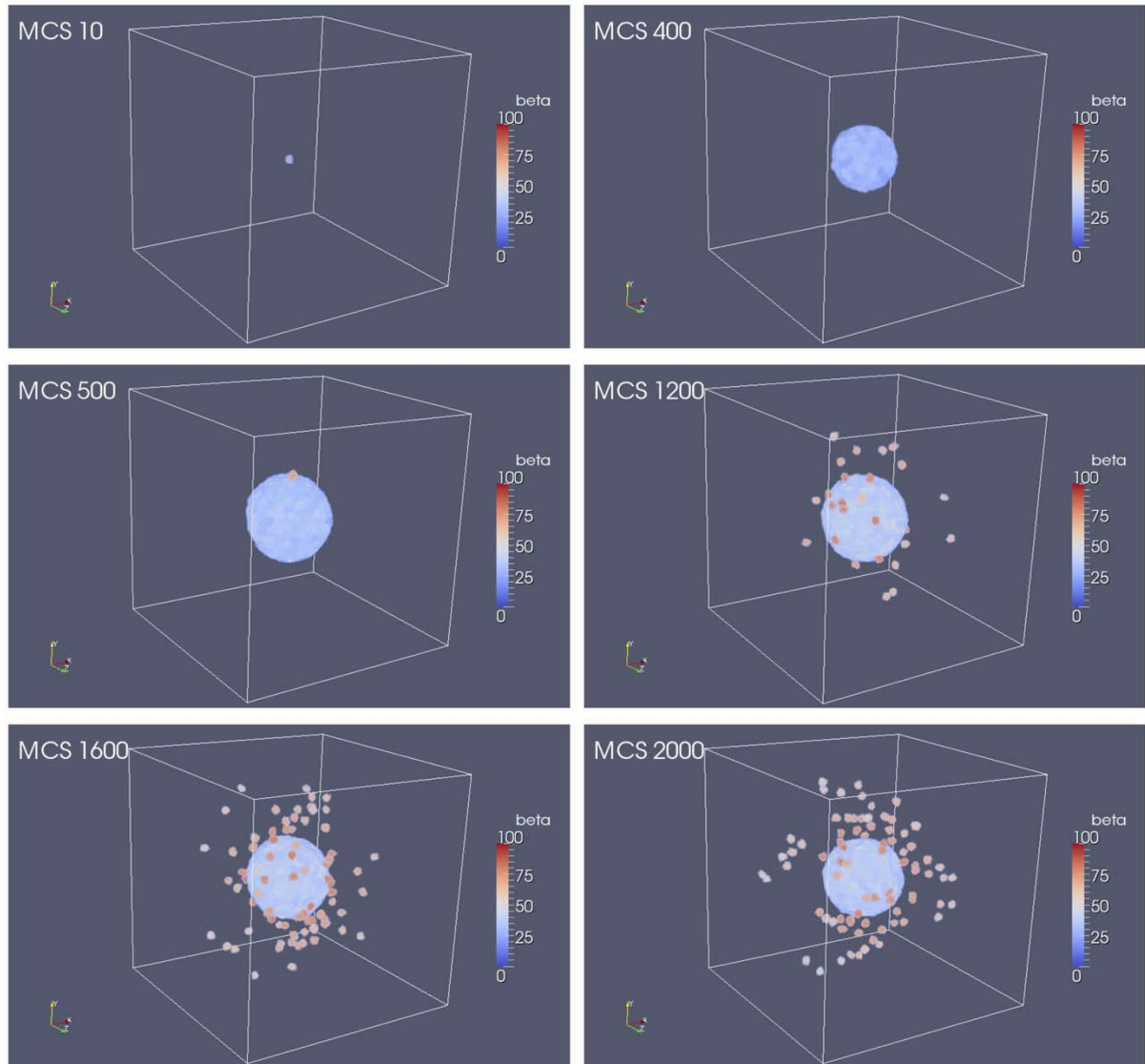


Figure 2.23 Results multicellular tumour spheroid simulation. The tumour grows from a single cell placed in the middle of the lattice. After 400 MCS, cells begin to detach and invade from the surface of the tumour. Cell colour represents β -catenin concentration (Andasari et al. 2012).

2.4.5. 2D Cellular Potts Simulation of the Role of ECM in Glioma Invasion

This cellular Potts based model is used to analyse the relative importance of cell-cell and cell-ECM adhesion in glioma invasion (Rubenstein & Kaufman 2008). In order to capture the dynamics of this type of tumour, three cell types are defined: proliferative, quiescent and necrotic. These types are widely used in continuum and agent based tumour modelling as cancer cells require nutrients to grow and divide. Thus as the tumour expands, cells on the inner layers become first quiescent (non-dividing) and then necrotic (dead). The ECM is represented with two components: a fibrous component that forms a scaffold and a non-fibrous component that is homogenous at the cellular scale. The inclusion of a two component ECM is a novel aspect of the model. It is intended to recapitulate the structure of typical collagen I gels, widely used for *in vitro* studies. Additionally, it allows for investigating the optimal density of ECM fibres for invasion, following *in vitro* results that cells are most invasive in intermediate density collagen gels.

The model qualitatively captures the invasive patterns that occur with *in vitro* experiments of gliomas embedded in collagen I gels. The authors acknowledge the limitation of accurate geometric representation of collagen fibres in the model. While *in vitro* collagen fibres generally range between 100nm and 1 μ m, in the model they are given the width of a single site (2 μ m) as this is the minimum permitted by the lattice scale used (Rubenstein & Kaufman 2008).

The model uses a target perimeter term in the Hamiltonian equation, and a parameter for the strength of this constraint, which they term 'elasticity'. Elasticity is varied in the model in order to explore the role of cell membrane rigidity on tumour growth and cancer cell invasion. This is used alongside 'Temperature' (cell motility) to explore the role of membrane rigidity in cancer cell invasion. The results suggest that, at very high collagen densities, cell shape alterations are not sufficient to overcome the small pore size, or disinclination of cells to detach from collagen fibres in order to migrate. This is in agreement with recent experimental findings (Wolf et al. 2007).

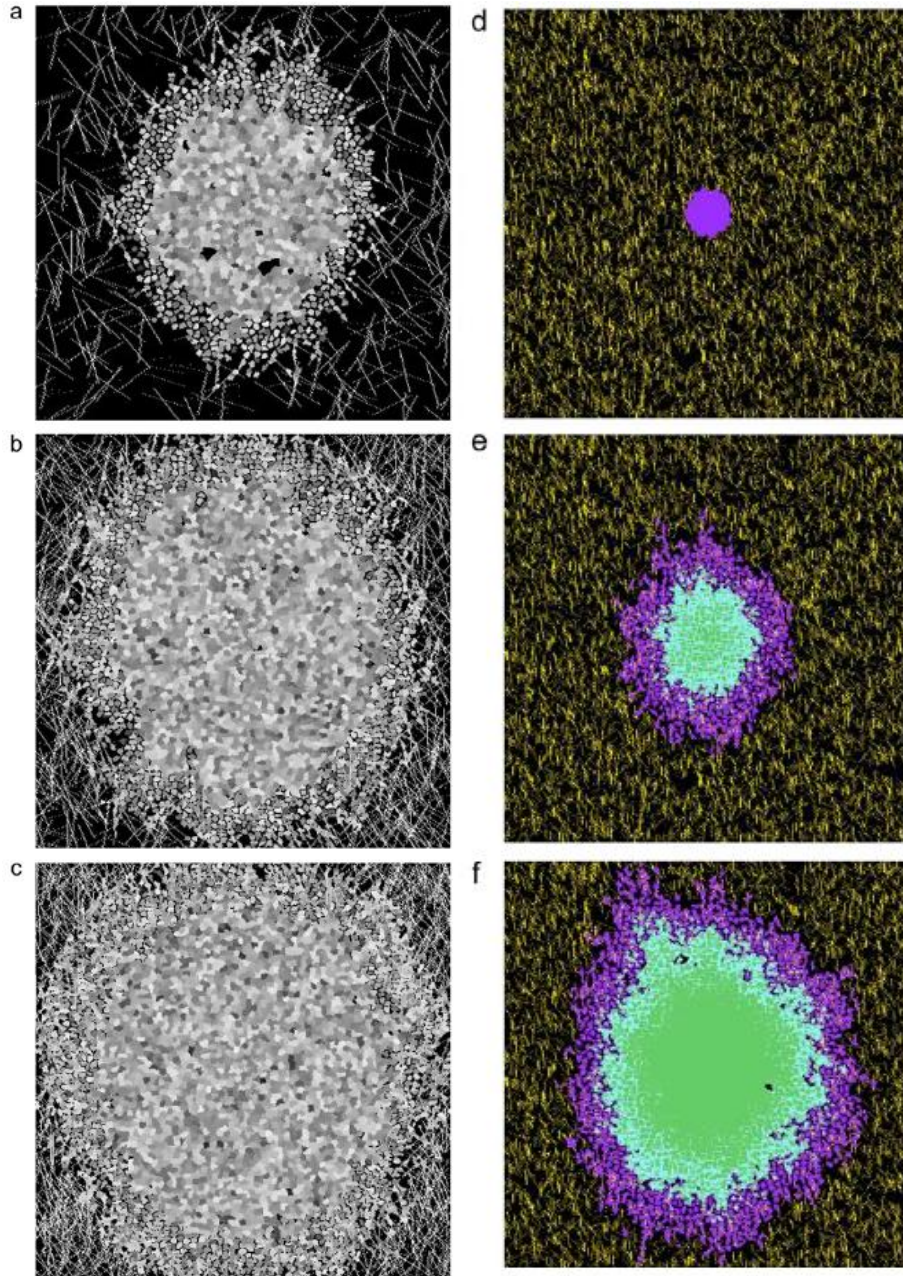


Figure 2.24 (a,b,c) Spheroid development at 14 days (2016 MCS) as a function of the number of $52\mu\text{m}$ long collagen threads, with 577, 2308 and 3462 threads respectively. Greyscale levels have no meaning, and are just to allow visualisation of individual cells. (d,e,f) Spheroid development as a function of time on a lattice containing 7500 $12\mu\text{m}$ collagen threads at 0, 7 and 14 days respectively. Key: proliferative cells: purple, quiescent cells: blue, necrotic cells: green (Rubenstein & Kaufman 2008).

2.4.6. Multiscale 3D Cell Centre Model of the Human Epidermis and Wound Healing

This 3D multiscale model integrates an agent based model of the epithelium with a model of TGF- β 1 signalling (Adra et al. 2010). This was also used to explore hypotheses of the functions of TGF- β 1 signalling during epidermal wound healing (Sun et al. 2009). The model provides a visualisation tool of the contradictory roles of TGF- β 1 signalling in keratinocytes: that it stimulates migration but inhibits proliferation. This has implications for the normal migration and turnover of keratinocytes in the epidermis, as well as the behaviour of cells when migrating and proliferating to cover a wound area. The integrated model recapitulates some qualitative aspects of wound healing. This includes the emergent effect that a small wound area is successfully covered by a fully renewed epidermis, due to the migration of epidermal cells; while in a large wound area, keratinocytes become committed before epidermal cells have migrated to cover the large area (Figure 2.25).

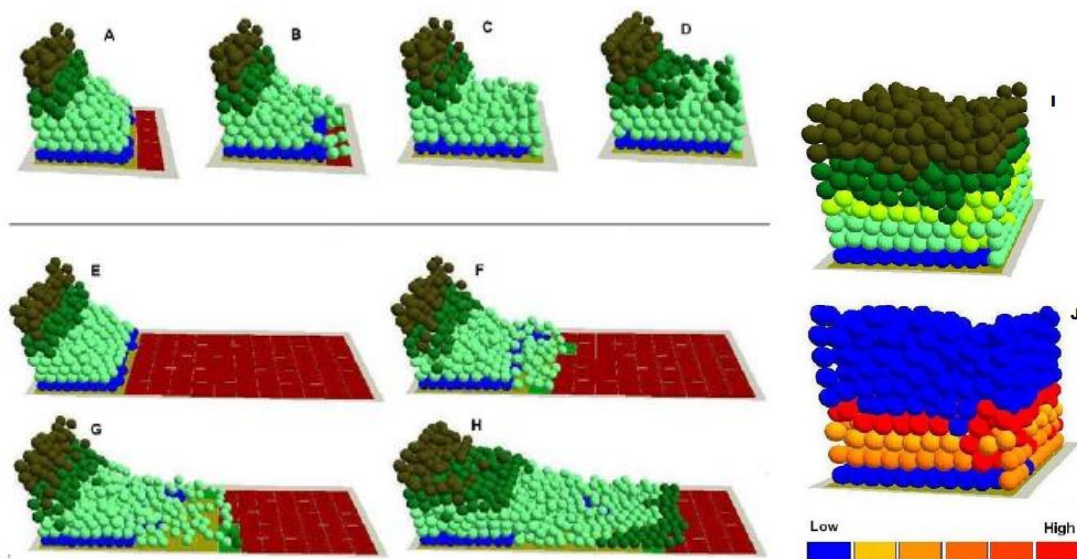


Figure 2.25 With a small virtual wound with normal proliferation and migration rates, epidermal cells migrate to cover the denuded area; shown at (A) 0, (B) 50, (C) 120, (D) 500 iterations. With a large virtual wound with the same migration and proliferation rates, keratinocytes on the wound bed begin to differentiate into committed cells before the epidermal cells have migrated to cover the large denuded area; shown at (E) 0, (F) 200, (G) 400, (H) 800 iterations. The level of TGF- β 1 in each cell was a function of the position of the agent in the stratified layers or in contact with the matrix (I) (J). Key: keratinocyte stem cells (blue), transit amplifying cells (light green), committed cells (dark green), corneocytes (brown), provisional matrix (dark red), secondary matrix (green). Cell agent diameter = 10 μ m (Sun et al. 2009).

The concentration of TGF- β 1 in cell agents in the model is dependent upon their position within the stratified layers of epidermis, their cell type, and contact with the matrix. For the subcellular model, an existing SBML model was solved during model iterations using the Complex Pathway Simulator (COPASI). In the integrated model, cell level behaviours were dependent upon the concentration of TGF- β 1 in the pathway model. This determined the propensity for cells to migrate, proliferate and differentiate. The TGF- β 1 variable also determined cell-cell and cell-substrate attractive forces, which represent adhesion. Cells in the model are set to migrate preferentially towards exposed areas of secondary matrix (wounds). While in reality this is driven by cell-matrix adhesion, as well as active cell movement, these features are reduced to a preferential directed migration in the model. The model demonstrated that TGF- β 1 plays an important role in maintaining the balance between migration and proliferation in normal epidermal wound healing. Any disruption to TGF- β 1 expression or signalling in the model could lead to chronic or hypertrophic wounds, which corresponds qualitatively with *in vitro* research (Fitsialos et al. 2007).

The multiscale approach of coupling an SBML model solver to an agent based modelling platform has similarities to the approach of developing multiscale models in CompuCell3D. It is interesting to note that the approach of the epidermal wound healing model makes use of a stochastic subcellular model solver, COPASI, with a deterministic agent based model simulator, FLAME. This is the reverse situation to using stochastic cellular Potts models with a deterministic subcellular model solver. It has been argued that there is little need for representing the stochastic nature of biochemical models within multiscale cellular Potts models, because the stochastic fluctuations of the Potts model are large compared to the deterministic approximation of subcellular ODEs (Hester et al. 2011).

2.4.7. 2D Cellular Potts Model of Intercellular Adhesion and Cancer Invasion

This 2D cellular Potts model is used to explore the relationship between cell-cell adhesion, cell-ECM adhesion, secretion of proteolytic enzymes and the rate of cell proliferation (Turner & Sherratt 2002). The model reproduces the qualitative emergence of ‘fingering’ and the invasive front of part of a tumour. The model suggests that the rate of mitosis can, under certain conditions, inhibit the invasiveness of cancer cells, by increasing the amount of surface area between cells at the invading front. As the cells adhere to each other, this is a limiting factor in the distance migrated (Figure 2.26).

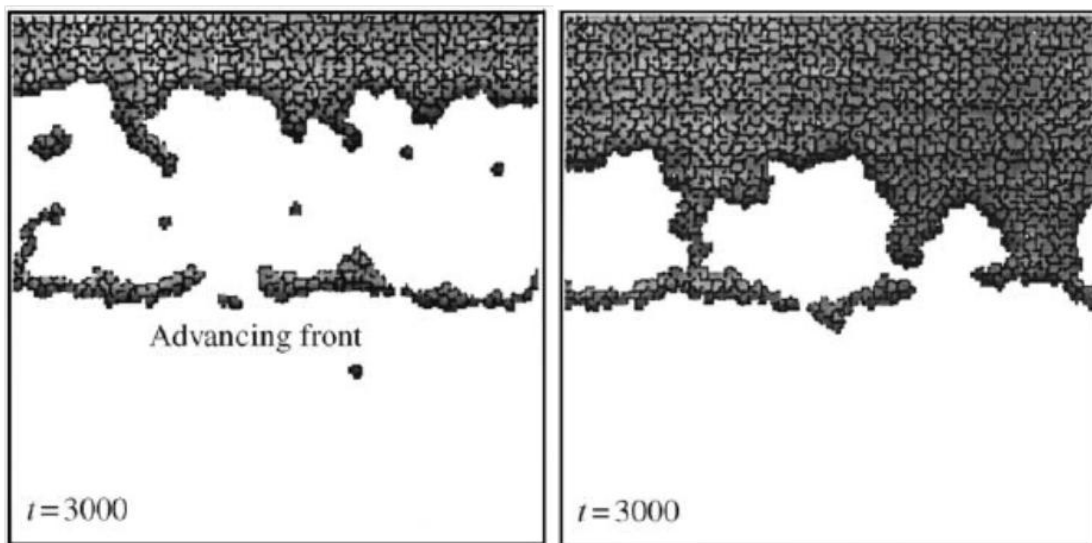


Figure 2.26 2D cellular Potts simulations of part of cancer invading front; (left) without cell proliferation, (right) with cell proliferation (Turner & Sherratt 2002).

2.5. Existing Simulations of EMT Pathways

2.5.1. Notch Signalling

The Notch signalling cascade is evolutionarily highly conserved. Notch-like molecules have been identified in a multitude of diverse species, from *C. elegans* to humans, and appear to play conserved functional roles in development (Borggrefe & Oswald 2009). The classic model of Notch signalling is the lateral inhibition model (Collier et al. 1996). In lateral inhibition, a cell adopts a fate, and then inhibits its neighbours from doing likewise. Thus fine grained patterns of gene expression may emerge from an initially homogenous tissue, allowing the evenly spaced differentiation of, for example, neural

cells from epidermal cells. Most simulation studies of Notch signalling are extensions on the Collier et al. model, and focus on the lateral inhibition mechanism. Refinements to the model of Collier et al. (1996) have included: implementing it as a homogenised reaction diffusion model (O'Dea & King 2011); investigating the role of differential adhesion and apoptosis in fine-grained pattern formation through a multiscale cellular Potts model (Podgorski et al. 2007); a multiscale model of Notch and VEGF signalling in tip cell selection during angiogenesis sprout formation (Bentley et al. 2008); an investigation of the role of mutual inactivation between Notch receptors and ligands (Sprinzak et al. 2011); and an investigation of the role of structured noise in driving pattern formation (Cohen et al. 2010).

In heart development, Notch signalling operates by lateral induction rather than lateral inhibition, to specify the regions of endocardial cells predisposed to EMT, as well as the regions of myocardium that secrete transforming proteins in these same regions (Timmerman et al. 2004). Only one model of Notch signalling lateral induction has been identified (Owen et al. 2000). This is a generic system of differential equations, with indications for ranges of parameters under which spatial patterning, uncontrolled feedback and homogenous equilibrium occur.

2.5.2. TGF- β and BMP Signalling

Transforming Growth Factor beta (TGF- β) superfamily (which includes Bone Morphogenetic Protein (BMP) signalling) generate intracellular signalling through a conserved family of proteins termed Smads. When a TGF- β or BMP ligand binds to a receptor, this phosphorylates an R-Smad protein (either Smad2/3 for TGF- β or Smad1/5/8 for BMP), which then bind to co-Smad (Smad4). The complex formed then translocates to the nucleus and activates gene transcription. The process during EMT in heart development is shown in Figure 2.27.

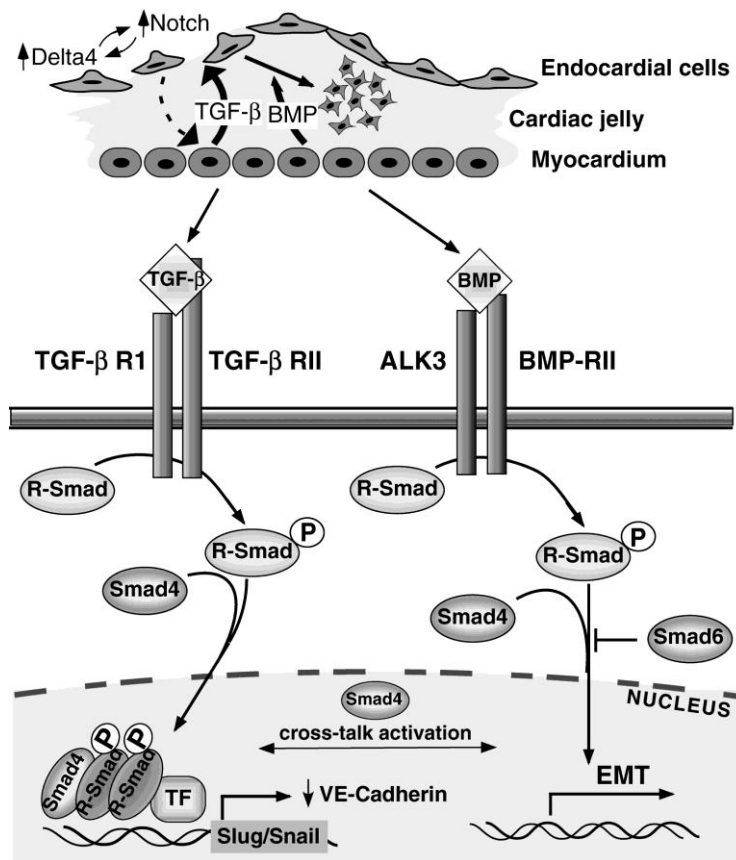


Figure 2.27 Notch, TGF-β and BMP activity in endocardial cushion formation (Wagner & Siddiqui 2007)

There have been several mathematical models of the TGF-β pathway. One model focuses on the receptor trafficking network, to demonstrate its potential for processing signals received from different TGF-β superfamily ligands (Vilar et al. 2006). Other models have focussed on the role of the shuttling of Smad proteins between nucleus and cytoplasm, and the phosphorylation of Smad in response to the TGF-β signal, without accounting for the receptor network (Clarke et al. 2006; Schmierer et al. 2008). The models of these different processes were combined using a constraint based method (Zi & Klipp 2007), which identified the importance of different mechanisms of endocytosis in regulating the signal response to TGF-β.

2.5.3. Wnt/ β -catenin and E-cadherin Signalling

It has been shown in numerous experiments that the loss of E-cadherin mediated cell adhesion leads to the release of β -catenin, which leads to nuclear localisation of β -catenin and the activation of Wnt/ β -catenin signalling. Conversely, many target genes of Wnt/ β -catenin signalling (including Twist and Snail2) lead to further E-cadherin degradation, and promote cell invasiveness characteristic of EMT (Heuberger & Birchmeier 2010). This positive feedback means that an EMT could be initiated either from canonical Wnt signalling, or from loss of E-cadherin adhesion by another pathway. β -catenin dynamics have been modelled as a system of differential equations, and implemented as a multiscale cell centre model of cancer cell EMT (Ramis-Conde et al. 2008). This model makes the assumption that E-cadherin and β -catenin will be stimulated to bind when a cell makes contact with another cell, and unbind when the cell detaches, and that this is proportional to the contact area. Furthermore, the model makes the assumption that when cytosolic β -catenin is above a certain threshold, there will be enough to translocate to the nucleus, activate gene transcription, and therefore induce migration. The model illustrates that when one cell receives a Wnt/ β -catenin pathway signal, the resulting invasion and rupturing of adhesive bonds with adjacent cells will in turn activate Wnt/ β -catenin signalling in those cells, inducing them to become invasive, and so on.

2.6. Challenges for Multiscale Modelling in Biomedicine

While it is widely accepted that appreciating the links between different levels of spatial and temporal scale is essential for integrated understanding of physiology, the majority of computational biology models are single-scale models. This is partly because different modelling techniques are suitable for representing the behaviour of biological systems at different levels of scale. Thus multiscale modelling usually requires the integration of two different approaches. For example: biochemical reactions are best represented with ordinary differential equations, or some form of network model; diffusion of chemical fields and cellular physiology are best modelled with partial differential equations; while cellular behaviour might be represented with agent based modelling. However, biological

systems are fundamentally multiscale. This may in part be a reflection of the evolution of complexity: from self-replicating macromolecules; to self-replicating units of interdependent molecules enclosed within phospholipid membranes (early cells); to cooperative dependence between groups of cells in the evolution of multicellular organisms; to the separation of physiological functions into interdependent organs and organ systems. The use of single scale models can thus only be taken so far, as it is necessary to understand most biological processes in terms of what is occurring at the levels of spatial and temporal scale above and/or below it.

As the use of computational models has grown within biological and biomedical research, along with the complexity and realism of such models, there has been an increasing realisation of the need for formality in the way that models are defined and developed. An increase in model complexity has also led to a growing desire for reuse of existing models as functional parts of new models. Multiscale modelling in particular has thrived on reuse. For example, a multiscale CompuCell3D model of somitogenesis (Hester et al. 2011) reused an existing segmentation clock model (Goldbeter & Pourquié 2008), which is available as CellML in the CellML models database, and also as SBML in the Biomodels database. Multiscale models of cardiac electrophysiology have reused models of calcium dynamics available in CellML (Hunter et al. 2008). There are a number of advantages to the reuse of published models in this way: as well as making the development of a complex model more manageable, it splits part of the tasks of verification and validation into component models. To some extent it also mirrors the reuse of components that occur in real biological systems: the same core molecular players, pathways and cellular behaviours are used in different developmental, physiological and disease processes, in different species. However, there are additional challenges of ensuring that models developed based on experiments in particular species will be valid in the context of a model derived from data in a different species. One means to address this is through semantically driven, machine readable annotation of models, using terms from biomedical ontologies. This at least makes it explicit what biological objects and properties the objects and parameters in the model represent; at whatever level of generality (e.g. “myocyte”, “mammalian cardiac ventricular myocyte”, or “rat cardiac ventricular myocyte”). There are now minimum information guidelines for the

annotation of models (Juty et al. 2012) and simulation experiments (Waltemath et al. 2011) that accord with similar guidelines for wet lab experimental research. Due to the logical rules expressed in ontologies, it is also possible to add a layer of reasoning, so that appropriate models and model connection points might be suggested when attempting to merge or couple models (Gennari et al. 2010).

The availability and adoption of standards is clearly an important aspect of reusability, as this enables the simulation of models encoded in the same language on different platforms. Quite different implementations are even possible for the same model. For example, an SBML model can be simulated as a deterministic system of ODEs using a large number of tools (e.g. JDesigner, Jarnac, CellDesigner, SBMLToolbox for MATLAB and Octave) or using a stochastic algorithm such as the Gillespie algorithm in, for example, COPASI or Dizzy. There have been many attempts to develop modelling standards in recent years. As one might expect, many have simply never taken off beyond the groups that developed them, while others have been extremely successful. Still others are highly successful, but limited by design to models within very specific domains. The latter include NeuroML (Gleeson et al. 2010) which now provides a common language for the handful of major neuronal simulators.

The success of eXtensible Markup Language (XML) specifications is partly due to this enforcing a declarative approach, that captures only the essential ‘information’ features of a model (objects, relationships and parameters), without any of the procedural code needed to run it. However, the standards with the greatest success have been for subcellular models: SBML and CellML. These are somewhat overlapping in terms of their domain, but both have been adopted as import and export formats by an impressive range of tools. The main difference between these languages is that CellML is more general, SBML being focussed mainly on pathway and reaction models. SBML data structures have names like “species” (e.g. molecular species) and “reaction”, while in CellML the biology is described only through metadata. CellML takes a modular approach while SBML is hierarchical. For example, while in SBML you have a list of molecular species, parameters, reactions, etc. in CellML there are modules, each of which may contain variables, objects, and other modules. This is intended to make reusing models

and parts of models (modules) easier. There is an import mechanism so that a module from one CellML file may be directly imported by another. CellML was developed by the Auckland Bioengineering Institute alongside another standard, FieldML. FieldML aims to provide a declarative language for models described by mathematical fields, such as geometrical models, with or without dynamic features. This can be used to create multiscale models of organs such as the heart by interpolating CellML models to discretised points on a FieldML field, and mapping CellML variables to a field degree of freedom (Popel & Hunter 2009).

2.6.1. Annotation of Multiscale Models

Another advantage of XML based languages is that it is possible to create an explicit link between concepts in a model and external web-accessible resources using the XLink mechanism (Hunter et al. 2006). This might be used to create a link to uniform resource identifiers for terms from biomedical ontologies. Ontologies are essentially collections of terms relevant to a particular domain, organised hierarchically, with logically defined relationships between terms. As such, they can be used for purposes such as classification, clinical decision support, making domain assumptions explicit, and semantic searching or querying of datasets.

Ontologies are particularly well developed within biomedicine. This is partly because it has become an extremely data intensive science, and partly because of the ‘messy’ nature of biology. Biology is riddled with exceptions, loose definitions, and confusing, co-existing, naming schemes; with dozens of synonyms for any given gene or protein. This makes formal semantic representation particularly important. Ontology development within biomedicine is being coordinated under the umbrella of the Open Biomedical Ontology (OBO) Foundry (Smith et al. 2007). These include both ‘reference ontologies’, which are considered to be the standard ontology for a particular domain, and ‘application ontologies’, which are designed for a particular purpose. Terms in application ontologies are often mapped back to terms from reference ontologies. The OBO Foundry provides reference ontologies with an increasingly good coverage of biomedical concepts at different levels of spatial and temporal scale. Initially, this grew from the coordination of heterogeneous databases that record the characteristics of gene products, primarily with the Gene Ontology (GO). Reference ontologies are now used for annotating a wide

variety of biomedical knowledge sources. These sources include images, database entries, publications, computational models and simulation results. By keeping reference ontologies well-bounded and essentially orthogonal the OBO Foundry minimizes logical inconsistencies and confusion over which ontology to use.

For many applications, there is a need to combine terms from multiple reference ontologies, in order to create a composite term suitable for a particular annotation. This can either be done by defining terms in application ontologies as equivalent to a composition of reference ontology terms (pre-composition); or through post-composition, whereby the annotator can compose terms ‘on the fly’, and add them to a repository of composite terms. While the former approach is less complex for the annotator, the latter approach is more flexible. An illustration of the post-composition approach is given in Figure 2.28.

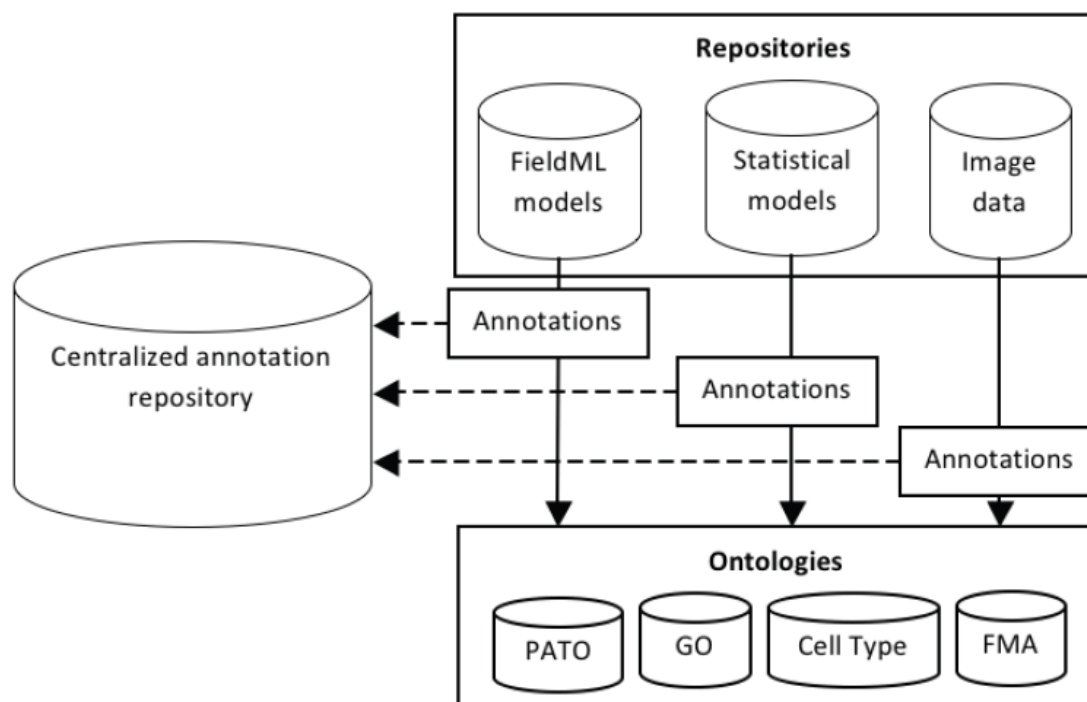


Figure 2.28 Relationship between models/data and annotations of models/data using ontologies. Models will be annotated to terms using a standard annotation scheme. These annotations will be stored in a central repository to enable fast querying (Baldock et al. 2010).

Multiscale modelling efforts have focused mainly on the physiology of adult organ systems. Post-composed annotation of models has so far been applied only to

physiological models with fairly simple physical properties (Neal et al. 2009). Embryonic heart morphogenesis involves complex cellular behaviour that is not well defined, and provides an interesting test case for the generality of the approach.

2.7. Existing Multiscale Simulation Platforms

2.7.1. VPH/Physiome Project

The VPH/Physiome project is an international public domain effort to provide a computational framework for understanding human physiology. Some of the initial outcomes include: specifications of standard languages for encoding biological models, web accessible databases of models encoded in these standard languages, development of ontologies for linking components between models and databases, and user interfaces for creating and running models; which facilitates linking between models at different levels of scale (Hunter et al. 2006).

Conventionally, when one bioscience research team wishes to make use of another team's model, it is necessary for them to entirely reconstruct the model from the published equations. This is cumbersome, time consuming and entails a number of steps in which mistakes are all too possible. Even when written source code of a model is made freely available, it is not easily integrated into models written in different languages, on different computer platforms. Overcoming these problems could greatly accelerate the pace of systems biology research. This is because, owing to the parsimonious nature of evolution, biological models are highly reusable for different contexts. Many signalling pathways, such as Notch signalling, are common across a multitude of diverse species, as are basic cellular mechanisms such as adhesion, migration, growth, division and death. It is for these very reasons that research on model organisms such as mice, yeast or fruit flies can yield results relevant to human physiology.

The Physiome project is addressing this challenge by developing a series of compatible eXtensible Markup Language (XML) based languages for encoding biological models, the

most developed of which is CellML. CellML makes use of another markup language, MathML, to express mathematical equations, and builds models as a network of interconnected components, so that an individual component of a model can easily be reused in a different model (Lloyd et al. 2004). There is an import mechanism that can be used for combining models, and grouping and containment can be used to organise model components in terms of their functions and physical locations. The model components normally represent components of a cell, though the language is not domain specific in this sense; and it could be used for modelling virtually any field. A related language, still in development, is FieldML. This is intended as a standard for modelling field descriptions, thus providing a standard interchange format for (potentially time-varying) geometrical models (Christie et al. 2009).

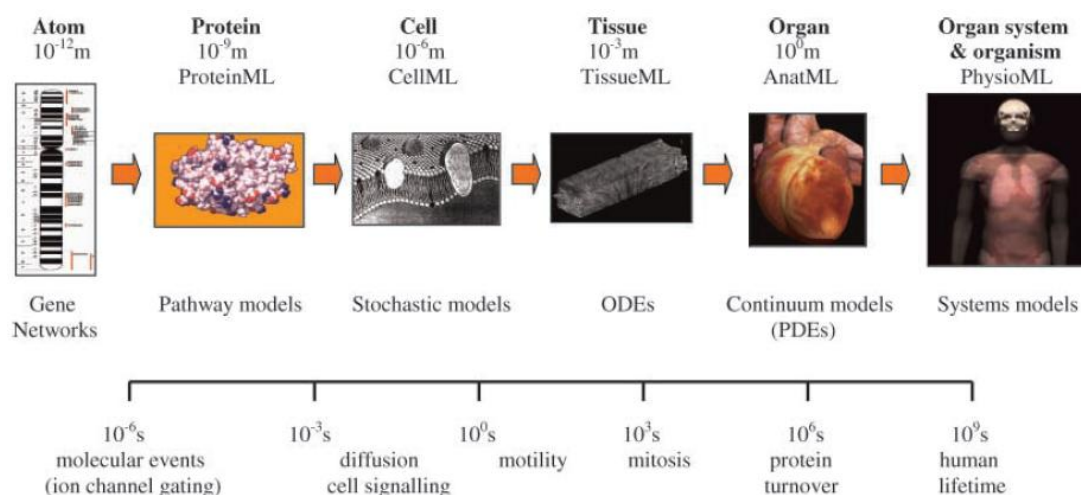


Figure 2.29 Spatial and temporal scales encompassed by the Human Physiome Project. Markup languages developed at Auckland Bioengineering Institute for different levels of scale are indicated, as well as the types of mathematical model appropriate to each level (Hunter et al. 2002).

The levels of scale encompassed by the Physiome project are illustrated in Figure 2.29. This is not to suggest that the entire spectrum of scales could be represented in a single model. Rather, depending on the question being addressed, different processes would be linked through a hierarchy of models; with different modelling approaches used at different scales. For example, ion channels are best represented by stochastic models, models of internal cellular dynamics by ordinary differential locations, the mechanics and

nutrient distribution of tissues and organs by partial differential equations and continuum models.

The first examples of Physiome projects focussed on cardiac physiology, coupling between models of excitation-contraction, cellular calcium dynamics, and cardiac electrophysiology, to produce geometrically accurate models of a beating heart (Crampin 2004). This has been achieved by coupling continuum models of excitation-contraction and heart electrophysiology, with ODE models of calcium dynamics. This approach couples between the single cell or subcellular level to the organ level, without explicit representation of cellular or tissue structure. An open source simulation environment, OpenCMISS (Bradley et al. 2011), has been developed at Auckland University for coupling of tissue or organ mesh models specified in FieldML with CellML models, to develop multiscale multiphysics models.

Modelling formalisms that treat cells as individual agents have tended to be more insular, with each model being encapsulated in a specific problem. While there is a library of models of biochemical reactions (Biomodels.net), and a library of models of cell physiology (CellML repository), and even the beginnings of a library of 3D geometrical models of human organs (FieldML repository); there is no comparable repository of models of cellular ‘social’ behaviour in tissues. It is precisely this social behaviour (changing shape, rearrangement, migration) that is the most fundamental aspect of modelling EMT. There are recent efforts to address the lack of cohesion in cellular agent based modelling, with the development of extendable simulation platforms, notably CompuCell3D, FLAME and Chaste. These are reviewed in the following sections. The question of whether declarative standards are possible for cellular agent based modelling is currently an issue of active discussion (Galdzicki et al. 2009; Osbourne 2012) but there are no such standards at the time of writing.

2.7.2. CompuCell3D

CompuCell3D is an open source platform for developing extended cellular Potts models, developed at Indiana University (Swat et al. 2009). It is built in C++ with a python wrapper. Models can be encoded in a combination of a native XML format (CC3DML) and Python; and the later allows dynamically changing parameters during a simulation run.

CompuCell3D provides a means for quick development of agent based cellular Potts models, circumventing the need to reinvent basic code for each investigation. While previously, every researcher developing cellular Potts models built their own platform, CompuCell3D allows this time to be saved for modelling and analysis; as well as allowing models to be more easily reproduced by other researchers. CompuCell3D boasts an interesting range of features, including a PDE solver which can be used for simulating the diffusion of chemical fields; either secreted by cells or as an intrinsic part of the environment. Cell types are specified with parameters for the target volume and surface area of a cell, constraints on these target values, response to chemical gradients, secretion of chemicals and surface energy between different cell types; which provides a representation of differential adhesion. Surface energy simulates cell-cell adhesion and repulsion, as cell types with low or negative surface energy between one another will tend to stick together, while those with high surface energy push apart. An intrinsic parameter for the speed at which all cells change shape is specified in the model, which is (misleadingly) termed “Temperature”. Simulations are normally run in a GUI that allows dynamic visualisation, but can also be run without the GUI. Parallel execution is supported through Open Message Passing. Multiscale models can be created by importing subcellular SBML models, and coupling between variables in the SBML and CompuCell3D models. As cellular Potts models were identified as the most appropriate technique for modelling EMT, CompuCell3D is used for simulations in this thesis, and the implementation is explained in more detail in the Methods chapter.

2.7.3. Chaste

Chaste (Cancer Heart and Soft Tissue Environment) is a general purpose simulation package developed at University of Oxford. It is aimed at multiscale and computationally demanding problems in physiology and biology, and as such is designed with the ability to run on clusters using Message Passing Interface. Chaste is developed in C++ using techniques adopted from the commercial sector. The team uses a variant of the agile development approach ‘eXtreme programming’, which includes test driven development, continuous integration and collective code ownership through programming in pairs (Pitt-Francis et al. 2009). This means a longer time is needed for adding new functionality initially, but this has rewards in the medium to long term by having a piece of software

that is readily understandable (through test suites and documentation). Due to the additional effort it requires to add new functionality, Chaste developers take the approach of adding new functionality as they require it, rather than predicting what they might need in the future (Osbourne 2012). This contrasts with the approach taken by the CompuCell3D team, of developing functionality so that it might be applied to a wide range of problems by a wide range of users; but at the expense of extensive testing and really thorough documentation. This difference reflects the fact that CompuCell3D is aimed at solving a much narrower set of problems: cellular Potts models with extensions, while Chaste aims to be a fully general simulation package. This means that Chaste needs to incorporate a much greater set of mathematical and computational techniques, with the consequence that a more rigorous approach is needed to ensure that these remain consistent when new modules are added.

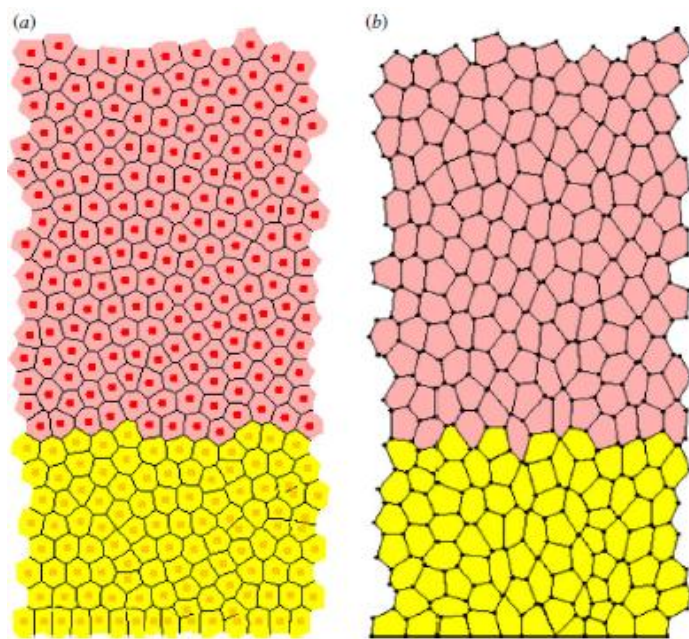


Figure 2.30 Example simulations of an intestinal crypt developed in Chaste using (a) the cell centre model and (b) the cell vertex model. Proliferating cells: yellow, non-proliferating cells: pink (Osbourne et al. 2010).

One of the advantages of having a range of techniques available within a unified simulation package is the ability to compare different modelling approaches for the same underlying biological model. This is shown for a comparison of the cell centre model with the cell vertex model in Figure 2.30. It would also be possible to do lattice based

simulations in Chaste, such as cellular automata or cellular Potts models, though this functionality has not been developed yet at the time of writing. The main applications of Chaste are currently simulations of cardiac electrophysiology and cell based simulations, particularly focussed on simulations of intestinal crypts, and the colorectal cancers that can appear within these. The cardiac simulations use a finite element mesh, and each point in the mesh is associated with a cardiac cell model that may or may not be stimulated. The cell based simulations also typically use a mesh to generate the geometry, and assign cell classes to that geometry; as well as forces to the simulation class that determine how the cells move (Pitt-Francis et al. 2009).

2.7.4. FLAME

FLAME (Flexible Large Scale Modelling Environment) is a generic agent based modelling platform developed at the University of Sheffield. FLAME has been used for modelling of cell populations as well as macroeconomic modelling, and modelling a signalling pathway, in which the agents represented molecules within a single cell (Holcombe et al. 2012). Individual agents are defined in a markup language (XMMML: X-Machine Markup Language) which is parsed to generate simulation code in C. FLAME can be used for multiscale agent based simulations of cells. This is achieved through the use of wrappers that allow agents to call either COPASI or JSim during a simulation. This means that models available in SBML, CellML or JSim's MML format can be coupled to FLAME agent based simulations, and solved deterministically or stochastically. This multiscale approach has been used to model the role of TGF- β 1 in epidermal wound healing (Sun et al. 2009).

2.7.5. SemGen

SemGen is a tool being developed at the University of Washington to automate the modular composition and decomposition of biological simulation models (Gennari et al. 2010). The approach taken is to first create lightweight semantic models (ontologies) representing each simulation model, and then attempt to merge those ontologies in order to suggest the points at which simulation models might be merged. This is achieved with composite annotations that leverage multiple existing biomedical reference ontologies. This allows the creation of complex definitions for the specific physical properties in particular models (Figure 2.31).

While SemGen is a work in progress, the approach suggested is an intriguing one. The demonstration used was to reproduce a previously hand-merged multiscale model, comprising of separate cardiovascular circulation, baroreceptor and vascular smooth muscle cell models (Gennari et al. 2010). The indication is that mapping between models at different levels of scale has the potential to be partially automated through the leveraging of semantic annotations. SemGen models are expressed in the SemSim language, and can be created from models in either of the major current declarative biomedical languages (SBML and CellML), but these need to be first imported and compiled in the JSim modelling tool. After merging the SemSim models, executable code can be exported. Export to JSim's MML is the only option currently available, but the developers intend to eventually offer translation to a range of languages. This approach has so far only been applied to models with well defined physical properties. It is not clear how well this might work for individual based cell level modelling, for example, where the physical properties that drive cell behaviour are not fully understood.

2.7.6. MoBi and PK-Sim

Developed by Bayer Technology Services, these two packages can be integrated to create multiscale models. PK-Sim provides a platform for whole body physiology based pharmacokinetics modelling (WB-PBPK), and MoBi provides a tool for protein and drug interactions. Applications of this include scaling models of drug responses between *in vivo* studies and human, or between adult and pediatric or elderly populations, by taking account of differences in organ sizes and composition (water, fat and muscle), as well as differences in physiology (Strougo et al. 2012). This is proprietary software, and with a different scale and application domain from that of this thesis. Nonetheless, this is an encouraging example of the use of multiscale modelling within the field of clinical drug development. There is likely much to be gained through greater collaboration between industry and academic research in the field of multiscale biomedical modelling.

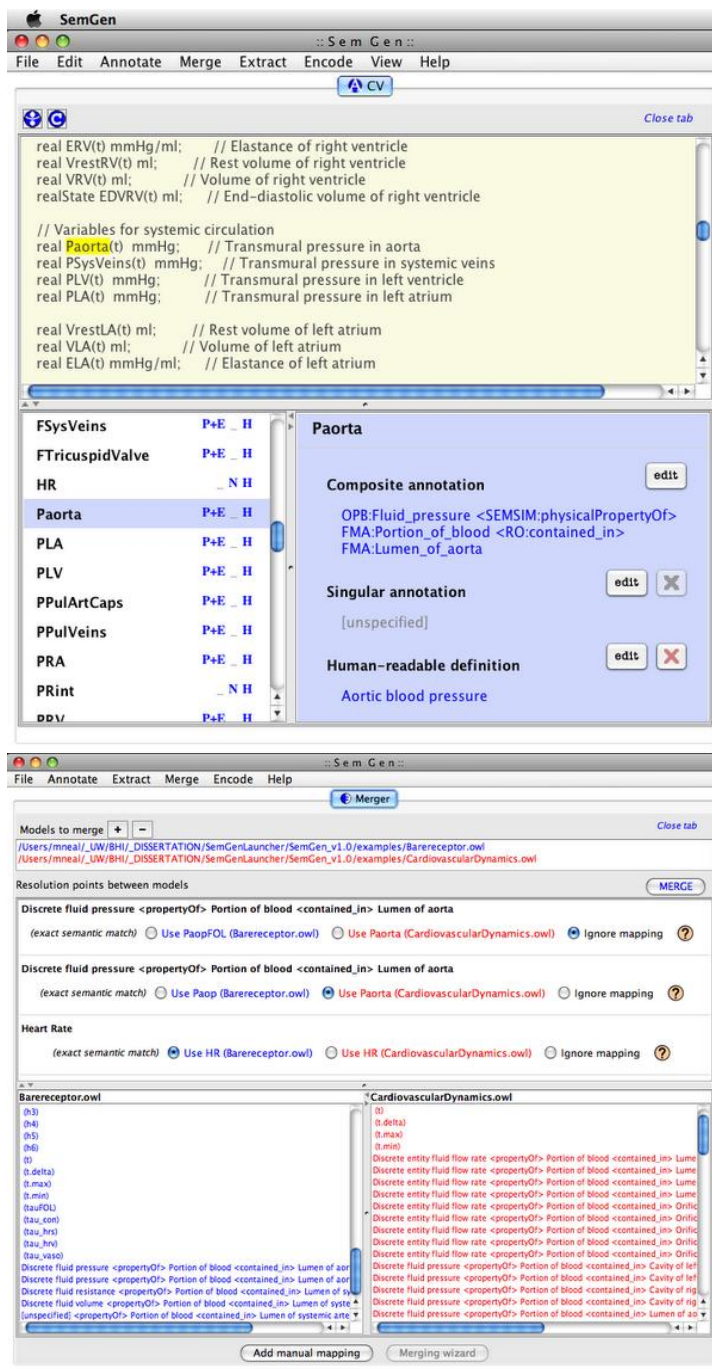


Figure 2.31 SemGen annotator tool (top) is used to create composite annotations for model variables, and the merger tool (bottom) is used to suggest resolution points between models (Neal 2010).

2.8. Summary

This chapter has provided a review of literature from several different disciplines to provide a solid foundation for the remaining chapters of the thesis. It has reviewed the anatomy and cell biology of EMT in heart development; the available cell-level modelling techniques and their suitability for modelling EMT; existing simulations of EMT; the challenges of multiscale modelling within biomedicine and existing platforms for multiscale simulation that address these challenges. The biological background of EMT and the technical background of modelling and simulation approaches will be built on in the following chapter; in which the methods used for modelling different features of EMT are described. The review of the approaches multiscale modelling in biomedicine is picked up in the Methods and Results chapters. In these following chapters, preliminary multiscale modelling of Notch signalling lateral inhibition is given, as well as an approach for multiscale annotation of for heart development processes with biomedical ontologies.

Chapter 3

3. Methods

3.1. Introduction

The Literature Review identified the cellular Potts model as having the most suitable range of features for representing EMT. This chapter begins with a description of the method selection, in terms of the types of cellular Potts models used to address different questions (Section 3.2.) The selection choices for different features of EMT (between 2D and 3D, single cell and multicell cellular Potts models, and whether or not to include subcellular reaction models) are explained.

In Section 3.3., the image processing methods used to extract cell shape parameters from the available imaging data are explained.

Section 3.4. explains the basis of the cellular Potts model, and the specific implementations used in this thesis. Further details are given in Appendix A. Section 3.4.1. details the constraints included in the 3D simulations and Section 3.4.2. details those included in the 2D simulations. The 2D simulations include more constraints than the 3D simulations, because cell morphology is investigated. The role of adhesion forces that bind across the cytoskeleton (in addition to those that operate at the surface) is investigated using the CompuCell3D FocalPointPlasticity plugin. Section 3.4.3. explains how the time and space scale of the simulations are matched to those of the real system, which is essential for comparing cell morphology and cell migration metrics. These are explained in Sections 3.4.4. and 3.4.5 respectively. Section 3.4.6 explains the multiscale implementation which uses the Bionetsolver library with CompuCell3D.

Section 3.5. provides details on the methods for model verification and validation.

3.2. Method Selection

Endocardial EMT is only one of the many complex developmental mechanisms that occur during heart development. However, it is a complex system in its own right, and comprises a large number signalling pathways and cellular changes. In order for computational models of endocardial EMT to be useful, they need focus on a small subset of the real interactions that take place. Rather than attempting to develop a unified model that integrates several aspects EMT regulation, this thesis takes the approach of providing a set of small models; each applied to a narrowly defined question or hypothesis. This means that a variety of models are presented, employing a range of methods. In order to assist the reader in understanding which methods are applied to which biological mechanism, these are summarised in Figure 3.1, along with the main justifications for the use of each method.

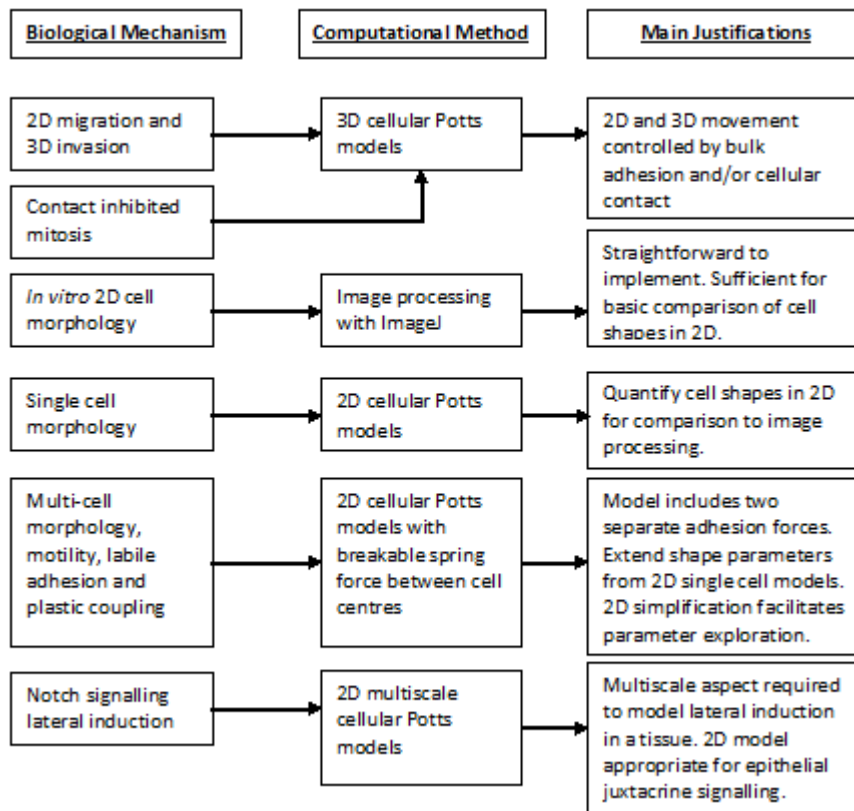


Figure 3.1 Overview of the computational methods applied to particular biological mechanisms, and the main justifications for the use of each method

Method selection proceeded on a number of criteria. In the case of quantifying *in vitro* cell morphology, image processing is really the only option, and it was a case of trying different techniques to find a method that worked and was easy to implement. Where the difference between 2D migration and 3D invasion needed to be captured in the model, modelling in 3D was a self-evident requirement. However, it would be complicated to quantify simulated cell shapes in 3D, and moreover the imaging data available was 2D. Therefore 2D simulations made more sense in all models that included changes in cell morphology as a feature (both the single cell and multi-cell models). In the case of the model of Notch signalling lateral induction, a 2D model was used. This makes the assumption that cells in an epithelial sheet are relatively thin in cross-section, and move around other epithelial cells, but not away from this thin layer. Under these assumptions, there would be no difference in the expression pattern produced by a 2D or 3D simulation.

One major point to note is that all the models are of either entirely abstract scenarios or represent endocardial cells *in vitro*. A model of *in vivo* EMT is not attempted. It is only possible to generate data on *in vivo* EMT through wildtype and mutant embryos, and tissue sectioning. While this can indicate something about the role of a particular gene in the entire process of EMT, it is insufficient for developing a multiscale or tissue-level model. This is due in part to the extremely large number of uncertainties and interactions present in a developing embryo. These interactions can be controlled to a much greater extent in an *in vitro* model of EMT, providing a means for model validation, and the potential for future feedback between modelling and wet-lab investigations. It is worth noting that, as indicated in the Literature Review, there is an existing 3D cellular Potts model that aims to represent *in vivo* endocardial cushion growth in terms of the qualitative functions of regionally restricted EMT, cell proliferation, differential adhesion and ECM production (Neagu et al. 2010). Whereas this thesis focuses on the cellular mechanics that occur during EMT, the model of Neagu et al. (2010) treats EMT as a lumped parameter. The models in this thesis agree qualitatively with those of Neagu et al. (2010). There may be potential for incorporating some of the more detailed mechanisms investigated in this thesis (such as changes in cell morphology and motility), within a functional representation of *in vivo* endocardial cushion growth. However it is not

immediately clear how this could be validated or would otherwise add value to the existing model of Neagu et al. (2010).

A recent *in vitro* investigation (Luna-Zurita et al. 2010) is used as the benchmark for all of the modelling in this thesis. This investigation provides, to our knowledge the greatest detail of any currently published *in vitro* investigation on the role of the main molecular players (e.g. Notch, TGF β , BMP2) in driving endocardial migration and invasion during EMT. 3D cellular Potts models were used to represent cell migration on the surface, and cell invasion into the gel, as a function of aggregate differential adhesion. High quality imaging data in the form of confocal stacks were provided by the authors of the investigation, and further analysis was conducted in the form of simple image processing. By this means, 2D cell shapes were extracted and quantified in the different experimental conditions. This was then used in order to fit 2D cellular Potts models that investigated the relationship between cell shape and cell motility. This was then extended to an investigation of the role of different adhesion forces in driving epithelial and mesenchymal morphology, by including a function for plastic coupling within multiple-cell 2D cellular Potts models.

More abstract simulations include the multiscale Notch lateral induction model. This is entirely generic to lateral induction in any epithelial tissue, though the implications for lateral induction in heart development in particular are considered. The 3D cellular Potts model of contact-inhibited mitosis is also highly abstract. It would require improvements to the model realism as well as further *in vitro* experimentation in order to validate this model. However it does at least give a visual and functional representation of one possible mechanism of the role of contact-inhibited mitosis in restraining cell migration and invasion during EMT.

3.3. Image Processing

In order to quantify the difference in shape between endocardial and mesenchymal cells, basic image processing analysis was performed on existing *in vitro* experimental results (Luna-Zurita et al. 2010) using ImageJ. Full resolution confocal stacks were provided by the authors of the *in vitro* study. Two conditions were investigated further for comparison: the wildtype condition and Notch1 activated (Tie2-Cre;N1ICD) cells. These conditions were chosen because the cells migrated mainly in 2D, rather than invading the collagen gel, allowing 2D comparisons of cell shape and migration to be made. This allowed investigation of the influence of Notch1 on cell shape and migration during EMT; as well as the interplay between cell shape, cell adhesion and cell migration.

As the wildtype condition consisted of a monolayer of endocardial cells, while the Notch1 activated cells were scattered, it was necessary to process them in different ways to extract individual cell shapes. For the wildtype condition, first a maximum intensity z-projection of the stack was produced, and the myocardium removed. The blue channel, which represents nuclei of the cells, was overlayed on the image in white. Then Find Maxima in ImageJ was used, which located these as local maxima (in terms of brightness) in the image. The output type Segmented Particles was used, which implements a watershed algorithm to extract an outline from the area surrounding the maxima. These outlines were then added to the ROI manager, allowing measurement of cell shape metrics to be performed. This workflow is shown in Figure 3.2. While it is impossible to strictly quantify the accuracy of this method without a ground truth segmentation with which to compare it to, a close inspection of the cell overlay indicated that the cell outlines were at least a reasonably good fit (Figure 3.3). As the method produced a lot small outlines, with very high circularity, cells with an area less than 100 μ m were excluded from the dataset.

For the Notch1 activated cells, slice number 15 was analysed, as this provided a representative image of cells moving on the surface of the collagen gel. The myocardium from the ventricular explants was removed from the image. Adjust Threshold in ImageJ was used to create a binary image of the cells. The Analyse Particles plugin was then

used to add the cell shapes to the ROI manager, which was used to measure the cell shape characteristics. This workflow is shown in Figure 3.4. Although there were some large shapes extracted, that clearly represent more than one cell, excluding these did not significantly affect the circularity or aspect ratio.

As the same shape metrics could also be calculated during a simulation run in CompuCell3D, this allowed a comparison to be made between the *in vitro* and simulation results. For the purpose of this comparison, the metrics of circularity ($4\pi \cdot \text{area} / \text{perimeter}^2$) and aspect ratio (length/width) were used. These metrics have the advantage that describe shape properties relevant to those taking place during EMT, and that they are widely used in other experimental work. Although a variety of names have been used for these metrics, e.g. 'form factor', 'cell shape index' or 'roundness' for circularity (Mendez et al. 2010; Malek & Izumo 1996; Gray et al. 2002), or 'length to width ratio' for aspect ratio (Davidson et al. 2010); their use in other research make the results of the image processing and simulations highly reusable. For the sake of consistency, the names for the metrics used in the ImageJ software are used to describe both the image processing and simulation results. The image processing results showed a significant difference in circularity between Wildtype and Notch activated endocardial cells. Full details are given in the Results chapter.

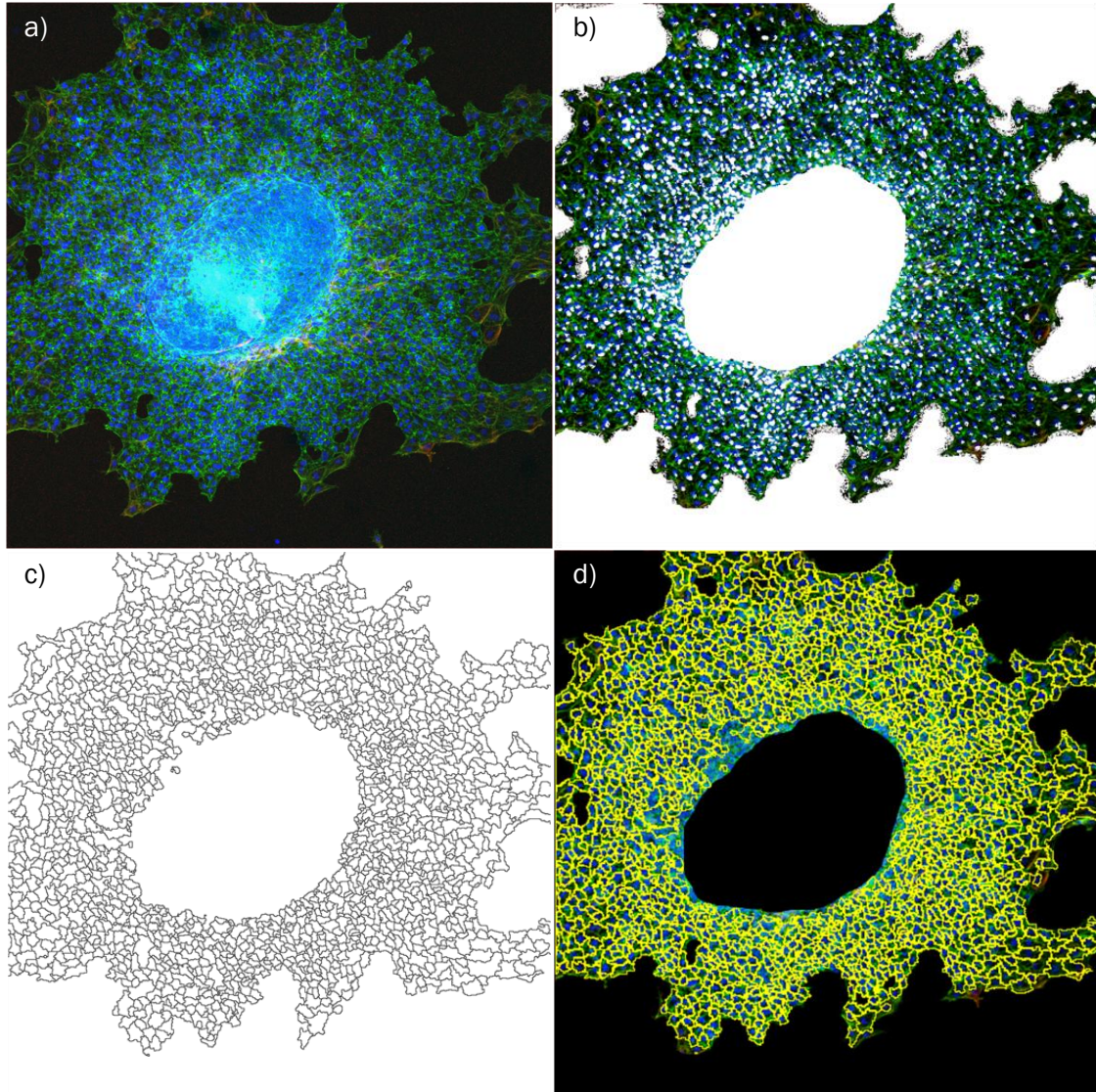


Figure 3.2 Image processing workflow in ImageJ for wildtype cells. a) Starting image b) Nuclei extracted and overlaid in white, background removed. c) Find maxima used to output segmented particles d) Outlines used to add cell shapes to ROI manager.

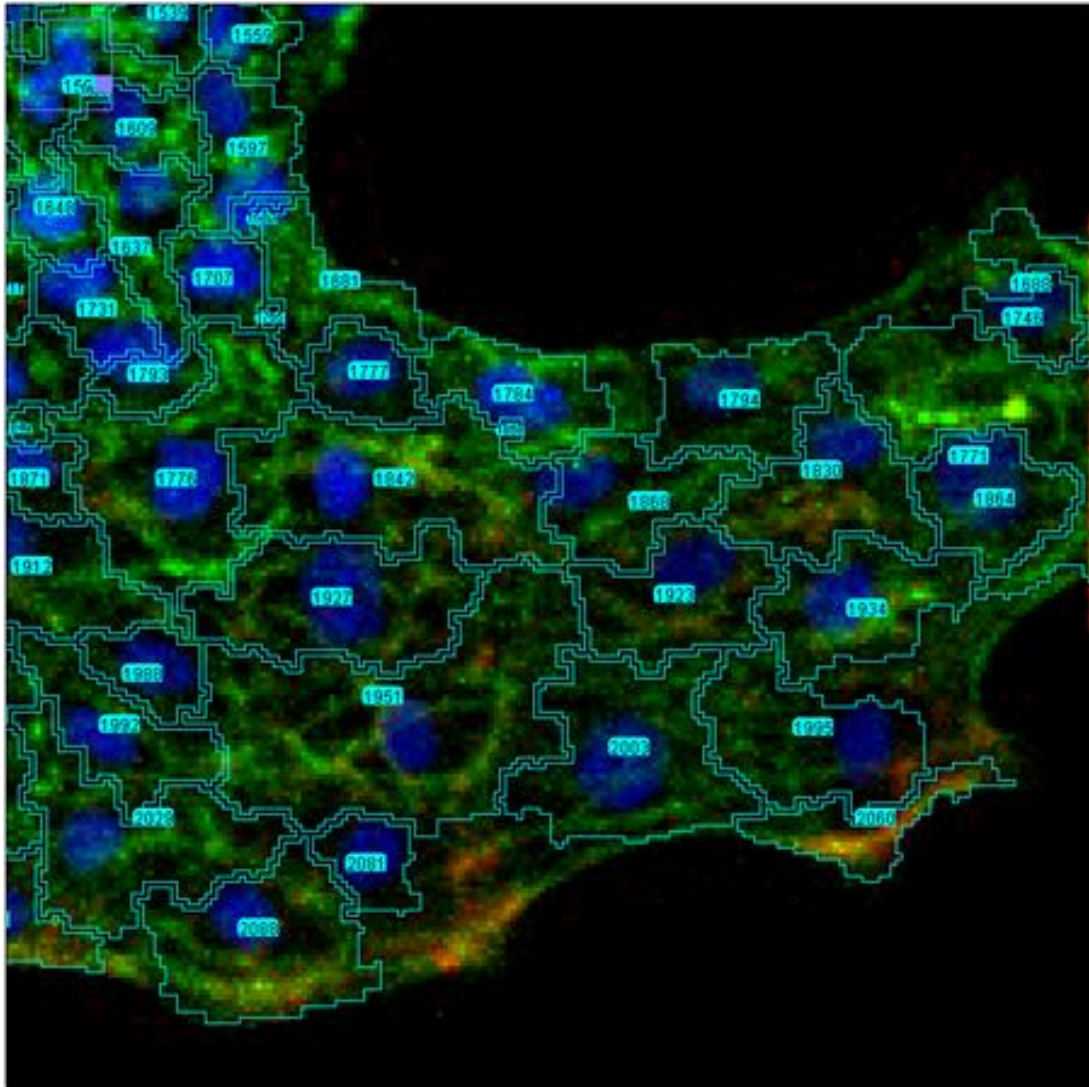


Figure 3.3 Closeup inspection of overlay for wildtype condition suggests that cell outlines provide a reasonably good fit.

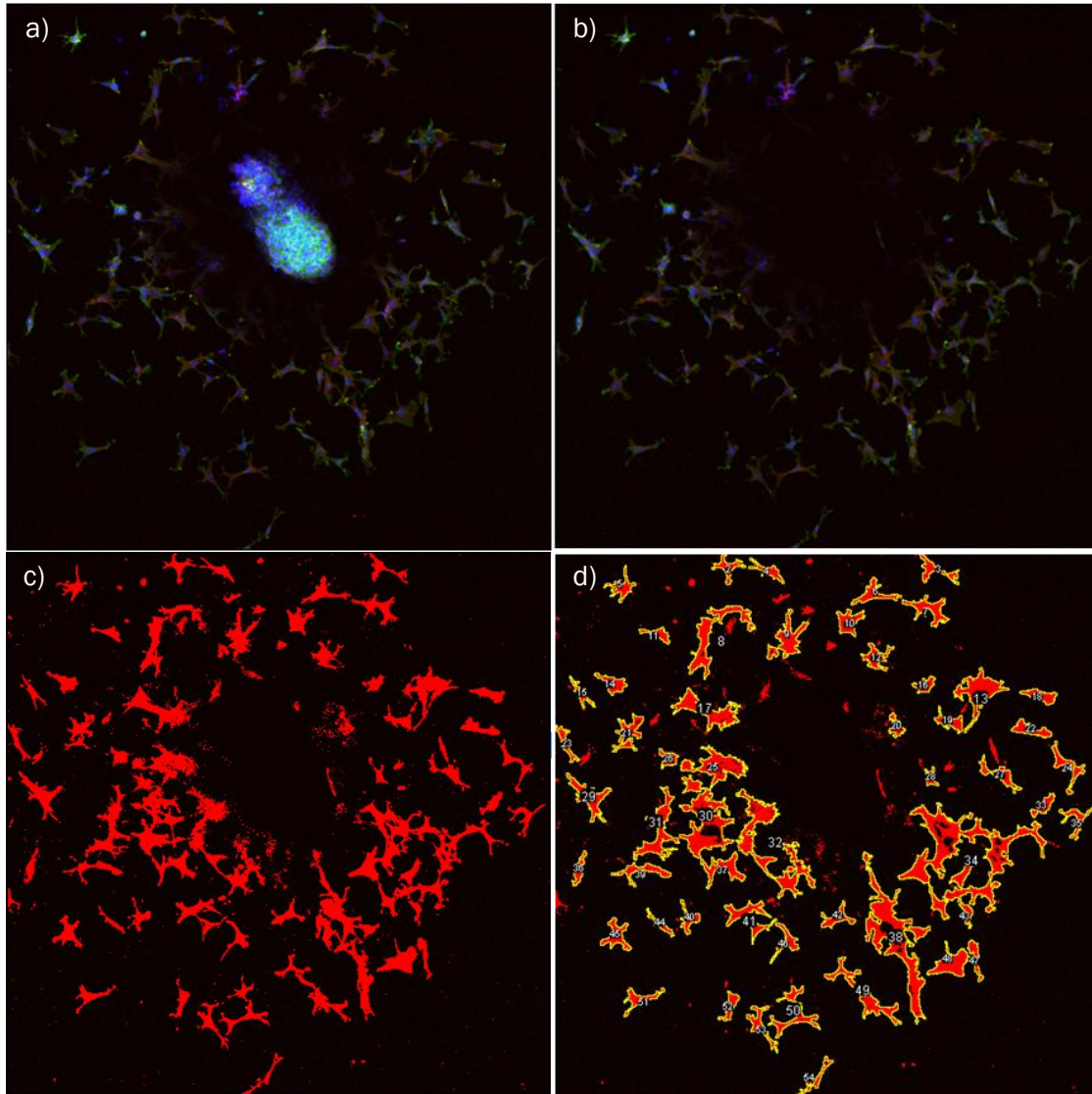


Figure 3.4 Image processing workflow in ImageJ for Notch1 activated cells undergoing EMT. a) Starting image b) Myocardium removed. c) Adjust Threshold to create binary image d) Analyse particles used to add cell shapes to ROI manager.

3.4. Cellular Potts Modelling

The simulations use the Cellular Potts model (CPM), also known as the Glazier-Graner-Hogeweg model (GGH). CPMs are lattice based simulations, with cells occupying multiple sites on the lattice. A key advantage of this is that it allows cell size and shape to be explicitly represented; thus the simulations have the potential to describe processes in which controlled cell shape plays an important role. CPM applications have included models of cancer cell invasion (Andasari et al. 2012) somitogenesis (Hester et al. 2011), bacterial biofilms (Popławski et al. 2008), slug formation in the cellular slime mould *Dictyostelium discoideum* (Savill 1997), developmental patterning of the chick limb bud (Cickovski et al. 2005), and the failure of Bruch's membrane in the eye by choroidal neovascularisation (Shirinifard et al. 2012).

In CPM, cell movements are described in terms of effective energies and constraints. According to the Differential Adhesion Hypothesis (DAH) morphogenetic changes are driven by cell displacements that lead to the lowest energy configurations, and thus the largest number of strong adhesive bonds. This can be simulated with CPMs using a Metropolis Monte Carlo algorithm. A Hamiltonian effective energy, H , is defined for the system. During each step in the simulation, a random copy attempt is made for each lattice site at a cell surface. For each copy attempt the resulting change in energy, ΔH , is calculated, and each copy attempt is accepted with a probability: $\min(1, e^{\Delta H/T})$; where T is used as an intrinsic measure of cell motility. This is because T scales the probability of any copy attempt being accepted, thus increasing T causes simulated cells to make more stochastic movements in a given time (hence move faster). There is a limit to increasing the T parameter, as under certain conditions this will lead to cell fragmentation and unrealistic results. In many cases, T can be increased within the range of realistic representation, to improve the speed and efficiency of simulations. Cell speeds and thus the lattice scale are then subsequently fitted to any available experimental data (Shirinifard 2012). In some of the simulations reported in this thesis however, the intrinsic motility of cells is itself a parameter that is to be investigated, as mesenchymal cells may well have a higher intrinsic motility than epithelial cells.

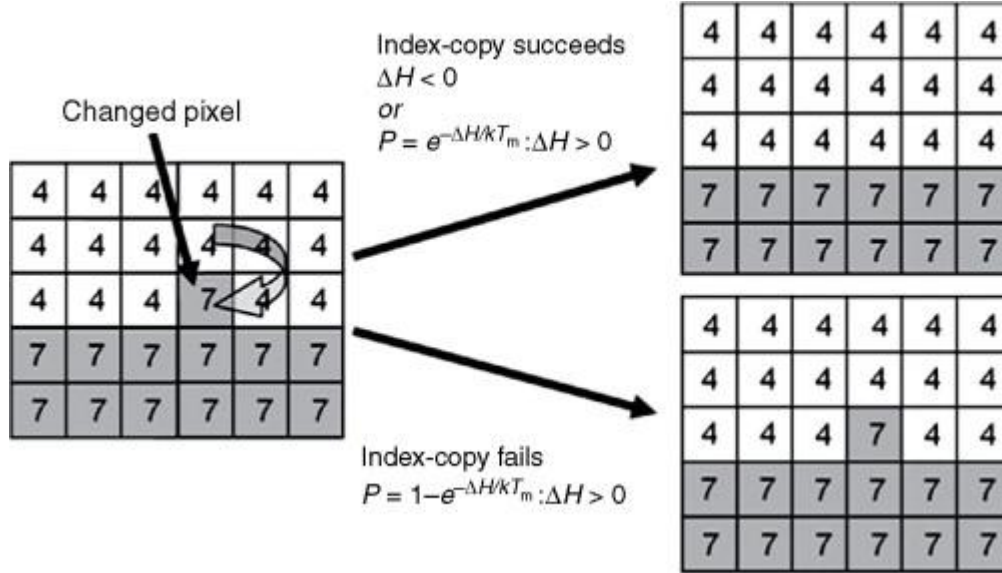


Figure 3.5 CPM representation of an index-copy attempt for two cells on a 2D square cell lattice – The “white” pixel (source) attempts to replace the “grey” pixel (target). The probability of accepting the index copy is given by $\min(1, e^{\Delta H/T})$. Thus any change reducing the entropy is accepted, while those increasing the entropy depend on the change in H and the T parameter (Swat et al. 2012).

In CPM, pixel copies (and cell movements) are biased towards creating configurations with a lower Hamiltonian energy H , so motile cells in a CPM will tend to move so as to reduce H . Thus over time cells move so as to reduce the entropy of the system. Typically the Hamiltonian equation includes terms for type dependent surface energies between each pair of different cell types. The basic effective energy in a simulation step is calculated by summing the surface energy across all cell boundaries. Thus cells with a lower surface energy (higher adhesion) will tend to move so as to come into contact, and stay in contact, while those with a high surface energy will tend to move apart. This contact energy parameter is usually treated as the aggregate surface tension between cell types, which is assumed to arise from the overall level of adhesion between cells.

It has been experimentally demonstrated that there is an almost perfect linear relationship between aggregate cell surface tension and the level of cadherin expression as shown in Figure 3.6. It is on this basis that the DAH and CPMs essentially treat cell aggregates as having the same properties as droplets of immiscible fluids (Foty & Steinberg 2005).

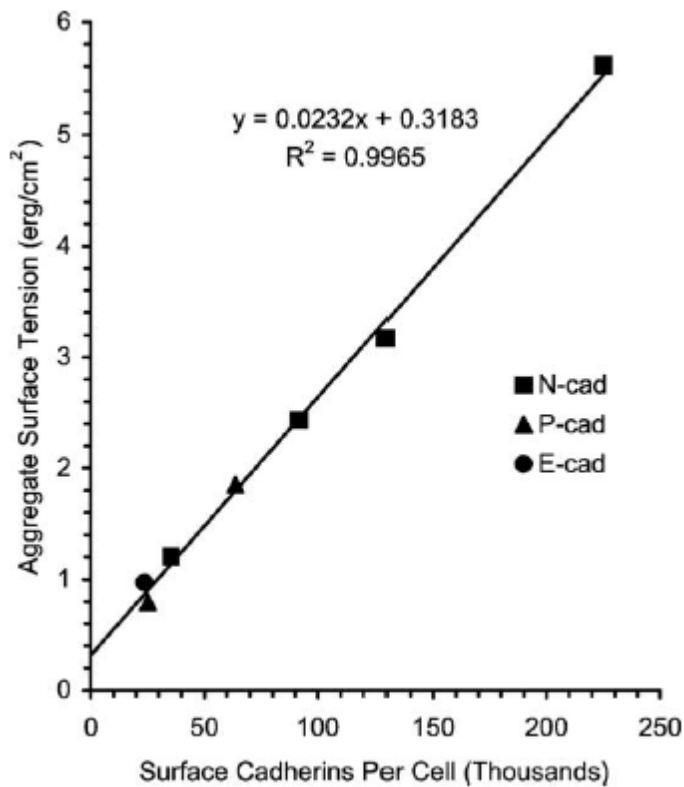


Figure 3.6 Data points expressing the relationship between cadherin expression level and aggregate surface tension fall almost exactly on a straight line ($R^2=0.9965$) that passes very close to the graph's origin, intersecting the Y axis (representing zero N-cad expression) at the very low surface tension value of 0.32 erg/cm² (Foty & Steinberg 2005).

Thus higher surface energy levels in CPMs represent a lower level of adhesion. The aggregate differences in adhesion levels between different cell types is sufficient to model of the kinds of cell sorting and patterning observed in many developmental processes. CPM can be extended to include terms for anything that can be calculated from the simulated cell attributes. For example, a type dependent target volume or target surface area can be included, with constraint values for the propensity of a cell to reach the target. These parameters directly constrain cell shape, and would be added to the effective energy equation so that they are accounted for in each pixel copy attempt. One of the simplest cellular Potts simulations represents an *in vitro* cell sorting experiment, where an initial mixed population of two or more cell types become sorted (Figure 3.7). The cells with higher preferential cohesion move to the centre of the cluster, while those with lower cohesion move to outer layers.

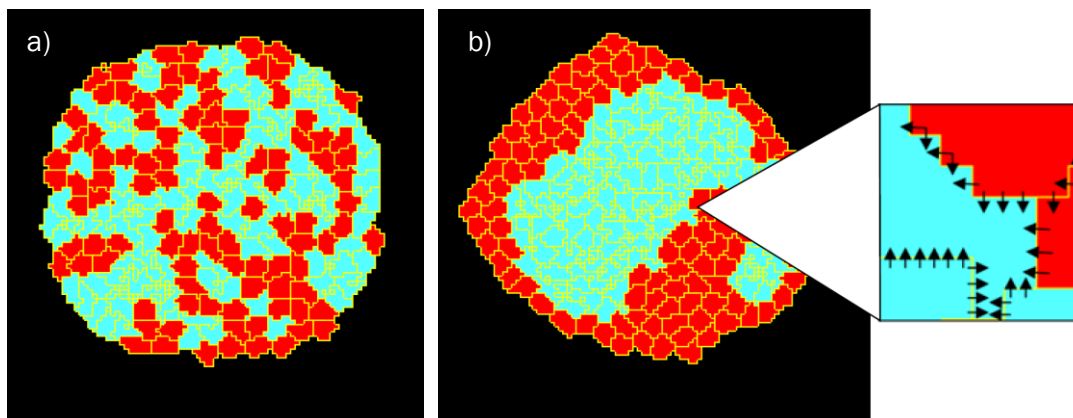


Figure 3.7 Cells interact at their surfaces (adjacent lattice sites) in cellular Potts models. a) At the start of the simulation, cell types are mixed. b) After 8000 Monte Carlo Steps (MCS), the cells are sorted, with the less cohesive type (red) forming an outer layer.

This provides a good qualitative representation of the behaviour a mixed population of *in vitro* cells that differ only in the expression of a single cadherin, as shown in Figure 3.8.

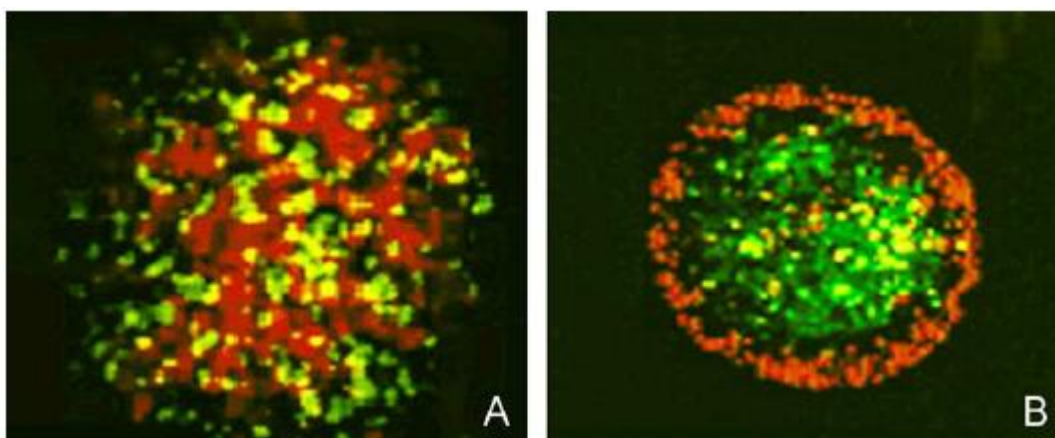


Figure 3.8 Sorting out of subclones differing only in expression level of a given cadherin. Two N-cad-transfected L cell subclones, expressing N-cad at their surfaces in the ratio of 2.4:1, were stained with red and green fluorescent membrane-intercalating dyes, mixed in equal proportions and cultured as hanging drops. (a) Confocal optical section through an aggregate after 4 hours of incubation, showing initial cell mixture. (b) Confocal optical section through another aggregate after 24 hours of incubation. As predicted by the DAH, the cell line expressing the lower level of N-cad (surface tension $\sigma = 2.4 \text{ erg/cm}^2$), labeled red, segregates from and envelops the cell line expressing higher amounts of N-cad (surface tension $\sigma = 5.6 \text{ erg/cm}^2$), labeled green. This demonstrates that cell sorting does not require (although it does, of course, permit) qualitative differences in cell-cell 'recognition specificity' (Steinberg 2007)

CPMs are able to simulate cell behaviour by representing any mechanism where cell rearrangement is determined principally by differences in adhesion. As the focus of CPM is cell reorganisation, they have been used mainly for modelling developmental

mechanisms. CompuCell3D (Swat et al. 2009) is the most widely used modelling environment for implementation of CPMs. It is open source software and very easy to extend, enabling flexibility, rapid model development and some degree of model reusability.

3.4.1. 3D Simulations

3D CompuCell3D simulations were used to represent the qualitative conditions of an *in vitro* study of EMT (Luna-Zurita et al. 2010) in terms of changes in adhesion, represented by changes in the contact energy parameters. The *in vitro* study demonstrated that 2D scattering of cells on the surface of collagen gel (partial EMT) could be induced separately from 3D invasion into the gel (full EMT). The models represent the collagen gel as one large generalised cell, and endocardial cells as initially forming a monolayer on the surface of the gel. Changes in surface energy parameters are sufficient to explain the behaviours of cell separation and migration on the surface, independently of cell invasion into the gel. These results are detailed fully in Section 4.1. 3D CompuCell3D simulations were also used to investigate the potential role of VEGF driven contact inhibited mitosis in maintaining the epithelial phenotype, and thus regulating EMT. The concept behind this is that, by preventing the formation of gaps between endocardial cells, contact inhibited mitosis may play a role in preventing the migration, invasion or transformation of cells. These are abstract models, and the results merely indicative. The results of the contact inhibited proliferation models are detailed fully in Section 4.2.

In addition to contact energy parameters, cells in the 3D simulations are also given a volume constraint in order to maintain them at an approximately constant size. Thus in these simulations, the Hamiltonian energy H is given by the following equation:

$$H_{3D} = E_{Contact} + E_{Volume} \quad \text{Equation 3.1}$$

Aggregate cell adhesion is represented by a contact energy parameter, J , between cells. The greater this number is, the lower the level of adhesion between cells:

$$E_{Contact} = \sum_{i,j - \text{pixel neighbours}} J[\tau_{\sigma(i)}, \tau_{\sigma(j)}][1 - \delta_{\sigma(i), \sigma(j)}] \quad \text{Equation 3.2}$$

Where i and j label two neighbouring lattice sites, σ 's denote cell ids, τ 's denote cell types, and the Kronecker delta is:

$$\delta_{x,y} = \{1, x = y; 0, x \neq y\} \quad \text{Equation 3.3}$$

The volume constraint term is given by:

$$E_{Volume} = \sum_{cells} \lambda_{volume} (v - V)^2 \quad \text{Equation 3.4}$$

Where λ_{Volume} denotes the strength of volume constraint, v denotes the current volume and V denotes the target volume.

3.4.2. 2D Simulations

2D CompuCell3D simulations of individual cells were used to simulate the changes in cell shape observed in EMT. By using simulations of a single cell, multiple parameters on the interplay of cell morphology and migration could be explored. To give control over cell shape in the model, terms for volume and surface constraints are added to the contact term in the Hamiltonian. An elongation term was used for simulations that investigate the effect of elongation, and a connectivity penalty to ensure that elongated and fibroblastic cells did not fragment. After achieving an understanding of the relationship between cell shape and motility in the CPM, part of the parameter space was implemented in 2D multicellular simulations of generic *in vitro* epithelial tissue.

The purpose of the 2D multicellular simulations was to investigate the roles of different types of adhesion in driving epithelial morphology. While the standard CPM includes adhesion only as a contact energy parameter, there are evidently a variety of forces acting between cells during different types and stages of cell adhesion. Some of these forces act at the surfaces of cells, while others involve the binding of cytoskeletal proteins across and between cells. Generally, the binding of cytoskeletal proteins is a stronger force that takes a longer period of time to form, while forces that act only at the surface are easier to form or break; and thus represent the initial stages of adhesion. A term for plastic coupling is added to the model, using the `FocalPointPlasticity` plugin in CompuCell3D. This implements plastic coupling as a breakable spring force

between cell center of masses. To distinguish this from contact adhesion in the model, the contact adhesion force will be referred to as labile adhesion, following (Shirinifard 2012). Thus in these simulations, the Hamiltonian energy H is given by the following equation:

$$H_{2D} = E_{Contact} + E_{Plastic} + E_{Surface} + E_{Volume} + E_{Length} + E_{Connectivity} \quad \text{Equation 3.5}$$

Labile adhesion is represented by a negative contact energy, J , between cells. The more negative this number is, the greater the level of labile adhesion between cells:

$$E_{Contact} = \sum_{i,j - \text{pixel neighbours}} J[\tau_{\sigma(i)}, \tau_{\sigma(j)}][1 - \delta_{\sigma(i), \sigma(j)}] \quad \text{Equation 3.6}$$

Where i and j label two neighbouring lattice sites, σ 's denote cell ids, τ 's denote cell types, and the Kronecker delta is $\delta_{x,y} = \{1, x = y; 0, x \neq y\}$.

Plastic coupling is represented by a breakable spring force between cell centres, and this term is added to the Hamiltonian equation governing the simulations.

$$E_{Plastic} = \sum_{i,j - \text{cell neighbours}} \lambda_{ij} (p_{ij} - P_{ij})^2 \quad \text{Equation 3.7}$$

Where p_{ij} is the distance between the centre of masses of cells i and j , and P_{ij} is the target distance corresponding to p_{ij} , and λ_{ij} is a constraint representing the strength of the plastic coupling. Additionally, a maximum distance is set, which determines the distance between cells' centre of masses when the links between them can break or form. The maximum number of links per cell is also defined. The minimum distance is set to be the same as the average cell radius, so that links tend to form when cells are in contact.

The volume and surface constraint terms are given by:

$$E_{Surface} = \sum_{cells} \lambda_{Surface} (s - S)^2 \quad \text{Equation 3.8}$$

$$E_{Volume} = \sum_{cells} \lambda_{volume} (v - V)^2 \quad \text{Equation 3.9}$$

$$E_{Elongation} = \sum_{cells} \lambda_{Length} (l - L)^2 \quad \text{Equation 3.10}$$

Where $\lambda_{Surface}$, λ_{Volume} and λ_{Length} denote the strength of volume, surface and length constraints, v , s and l denote the current volume, surface and major axis length, and V , S and L denote the target volume, surface and length. The connectivity constraint is a large energy penalty (10^7) for a pixel copy that would lead to cell fragmentation.

3.4.3. Fitting Space and Time Scale of Simulations

The 3D CompuCell3D models used initial 10x10x2 voxel cells, whereas 2D models are 15x15 pixels, because more detailed shape information was needed for the morphology investigations in the 2D simulations. While the 3D simulations attempt to produce only a generic, qualitative representation of the processes in terms of aggregate cell adhesion, the 2D models are directly fitted to the experimental data. For the 2D models, the length scale was set to 1 micron per pixel, and model cells given a width of 15 microns, based on the dimensions of cultured murine endocardial cells (Luna-Zurita et al. 2010). This choice reflects a prospective goal to include subcellular reaction kinetics; and thereby develop a multiscale model of the integrated roles of Notch and TGF- β signalling in endocardial to mesenchymal transition. Experimentally, *in vitro* epithelial cells such as MCF-7 cells move at a rate of about 0.28 $\mu\text{m}/\text{min}$, whereas mesenchymal cells move at about 0.4 $\mu\text{m}/\text{min}$ (Mendez et al. 2010). For typical parameter settings, simulated epithelial cells move at about 0.01 pixels/MCS (e.g. 0.01 $\mu\text{m}/\text{MCS}$). Equating the experimental and simulated cell speed implies 0.28 MCS = 0.01 min, or 1 MCS = 0.036 min (about 2 seconds). The 2D multicellular simulations were each run for 5×10^4 MCS, which equates to 30 hours. Cellular movement was quite stable after this time, and this provides a link with typical *in vitro* experiments, for which results are often given after 24, 48, or 72 hours, e.g. (Luna-Zurita et al. 2010).

3.4.4. Cell Morphology Metrics

There are a few measures that can be used for quantifying the shape of cells. The main criteria for selecting suitable metrics is that they should enable a comparison to be made between the *in vitro* results (Luna-Zurita et al. 2010) and the simulation results, that they should describe shape property relevant to the changes which take place during EMT, and preferably that they should be widely used in other experimental work, to allow for direct comparisons, making the results reusable. Based on these criteria, the two metrics selected were aspect ratio (length/width) and circularity ($4\pi \cdot \text{area} / \text{perimeter}^2$). Aspect ratio indicates the elongation of a cell. Circularity gives a number between 0 and 1, with 1 being a perfect circle and smaller values indicating a less rounded shape. These metrics could both be calculated during a simulation run, from basic cell attributes, and calculated from the confocal images of endocardial explants using ImageJ. This allowed a direct comparison to be made between simulated and *in vitro* cells.

3.4.5. Cell Migration Metrics

There are a plethora of metrics that are used to quantify cell migration. For example, *in vitro* assays sometimes report the average speed of cells (Gilles et al. 1999), and sometimes the average velocity (Mendez et al. 2010). Alternatively, the average displacement of cells over time can be measured (Rupp et al. 2008). When it comes to tissue simulations, there are even more possibilities, because the migration of an individual cell can also be quantified by its relation to other cells. For example, it is possible to measure the average contact area between cells, the average number of cell neighbours or the ratio of cells that remain adhered to other cells (de Rooij et al. 2005). Some investigations define their own metrics, for example, one study defined a 2D and 3D transformation index as the ratio of cells able to migrate on the surface, or invade the collagen gel (Luna-Zurita et al. 2010). It is even possible to analyse cell centre coordinates (provided by DAPI stained nuclei) by constructing a cell graph, and thus provide metrics such as the mean and standard deviation of edge lengths, a clustering coefficient, and the number of central, connected and isolated points (McKeen-Polizzotti et al. 2011). It is possible to implement all of these metrics in simulations. Indeed it is more straightforward to measure and analyse this information from simulation runs than from *in vitro* experiments. However, in order to provide validation, the average speed and

velocity remain the most suitable metrics of cell migration, as they are the most widely used in experimental work.

3.4.6. Compucell3D-Bionetsolver Implementation

Bionetsolver is a C++ programming library with a high level Python API that allows for easy definition of multiscale models by assigning reaction-kinetic models in SBML to cell objects in a Compucell3D simulation. Bionetsolver makes use of the SBML ODE Solver Library (SOSLib) for reading the SBML models and solving them as a system of ODEs. Typically the Bionetsolver API is imported and initialised in a Compucell3D Python steppables file, which makes the API available within the steppable. SBML models are loaded with a `loadSBMLModel` function, which uses a string argument to specify the directory of the model. The loaded SBML file can then be assigned to one or more cell types in the Compucell3D model with the function `addSBMLModeltoTemplateLibrary` by using the name of the cell type as the name of the template library. Additionally the function `setBionetworkInitialCondition` may be used to specify initial conditions for the parameters and variables in the assigned SBML file: thus different SBML initial conditions may be set for different cell types with the same SBML model. When the function `initializeBionetworks` is called this creates a separate bionetwork object for each cell of the given cell type using the specified initial conditions. Any parameters not set with `setBionetworkInitialCondition` are set by default to those given in the original SBML file.

The above functions are all called during the start function of the Compucell3D steppable. There are three additional functions that are called during the step function of the steppable. These are: `timeStepBionetworks` for time stepping the ODE integrator, `getBionetworkValue` for retrieving SBML parameters or variables from particular cells, and `setBioNetworkValue` for setting the SBML parameters of particular cells. In this way, SBML parameters can be set as a function of Compucell3D cell properties, and Compucell3D cell properties can be set as function of SBML variable values. Thus a true multiscale coupling can be achieved between Compucell3D and SBML dynamics, while maintaining the cell and subcellular levels as separate models, which

might well be developed independently in the first instance. A skeleton steppable that implements the Bionetwork API is given in Figure 3.9.

```
import bionetAPI # Import bionetAPI functions
class <someClass>(SteppableBasePy):
    def __init__(self, _simulator, _frequency=1):
        SteppableBasePy.__init__(self, _simulator, _frequency)
        bionetAPI.initializeBionetworkManager(self.simulator) # Initialize bionet inside class

    def start(self):
        # Load a specific subcellular SBML submodel
        ModelName = <sbmlModelName> # Name of the model
        ModelPath = <sbmlModelPath> # Path where the model is stored
        ModelKey = <modelKey> # Nickname of the model
        IntegrationStep = <timeStep> # Time step of integration
        bionetAPI.loadSBMLModel( ModelName, ModelPath, ModelKey, IntegrationStep )

        # Add SBML submodel to a group of cells or a single cell
        bionetAPI.addSBMLModelToTemplateLibrary(<sbmlModelName>, {<cellType> or <cellId>})

        ...

        # Modify the parameter value or molecular concentration of a cell (or group of cells)
        bionetAPI.setBionetworkValue(<molecule/parameter>, <value>, {<cellType> or <cellId>})

        ...

        # Initialize model
        bionetAPI.initializeBionetworks()

    def step(self, mcs):
        # Iterate the model (run it for the time step specified on the load command)
        bionetAPI.timestepBionetworks()

        ...

        # Get the parameter value or molecular concentration from a cell (or group of cells)
        <var>=bionetAPI.getBionetworkValue({<parameter> or <molecule>},{<cellType> or <cellId>})

        ...

        # Modify the parameter value or molecular concentration of a cell (or group of cells)
        bionetAPI.setBionetworkValue(<molecule/parameter>, <value>, {<cellType> or <cellId>})
```

Figure 3.9 Skeleton of Bionetwork API implementation within a CompuCell3D Python steppable

A multiscale model of Notch signalling lateral induction was implemented using the Bionetsolver API. This was achieved by translating an existing ODE model of lateral induction (Owen et al. 2000) into SBML using the Systems Biology Workbench (SBW). Multiscale models in CompuCell3D are typically implemented as four simulation files. For the sake of data provenance, CompuCell3D can be set to automatically copy these files into a new (output) directory, along with any screenshots or results recorded during a given simulation. This is a very useful feature of CompuCell3D, as this provides a record of the exact code used to generate a given set of results. Small iterative changes during model development are almost inevitable, and without this facility would be cumbersome and error-prone to record manually.

The four simulation files of a multiscale model generally consist of:

- A file in an XML format native to CompuCell3D (CC3DML), which sets the Potts dimensions, the CompuCell3D plugins to be loaded in the model (perhaps including the volume and surface constraints), the number of Monte Carlo steps to run the simulation for, the 'Temperature' parameter, the cell types and contact energies in the simulation, and a specification of the initial geometry as regions of cells within the lattice. The geometry may alternatively be set by linking to an additional Pixel Initialisation File (pif).
- A Python steppables file. This contains modules with code that is to be run at the start of the simulation, repeatedly (e.g. every step, or every 10 steps), and at the end of the simulation. Steppables are used for implementing dynamic features of the model. These include the rules governing mitosis, calling Bionetsolver functions (as shown above in Figure 3.9), doing calculations for cell migration and morphology metrics and printing these to a file, and creating dynamic visualisations such as graph plots or scalar or vector fields.
- The SBML file (a widely used XML format for biological reaction modelling). This provides details of the reaction model, including the protein species involved, the reactions, parameters and variables, as well as the initial values.
- The main Python simulation file. This adds the steppables to the model and initialises the simulation to be run.

The advantage of this approach to modelling is that it allows rapid development of models, and allows flexibility in terms of how a model is implemented. It hides much of the complexity of procedural code, used in every single Potts simulation, from the modeller. If the modeller wants to develop new C++ modules for their specific needs they can do so, as CompuCell3D is open source. While, in the recent past, every research group doing Potts modelling would have developed all of their own code, CompuCell3D is becoming an increasingly popular choice. This not only saves time, allowing more time for model implementation, it also facilitates the reuse and reproduction of published models. This verification through reproducibility is essential to the scientific process.

One disadvantage of the approach is that having several simulation files that specify parameters, plugins and simulation objects, means that it is quite easy to lose track of

exactly what has been specified where. It very easy, for example, to neglect to import the necessary plugin in the CC3D XML file in order for one of the Python steppables to work. The developers have begun to address this to some extent by including a program called Twedit++ with binary distributions of CompuCell3D. This is developed on top of the free source code editor Notepad++, and provides the option to create simulation files using forms to specify the behaviours included in the model. Boilerplate code snippets can also be selected from dropdown menus.

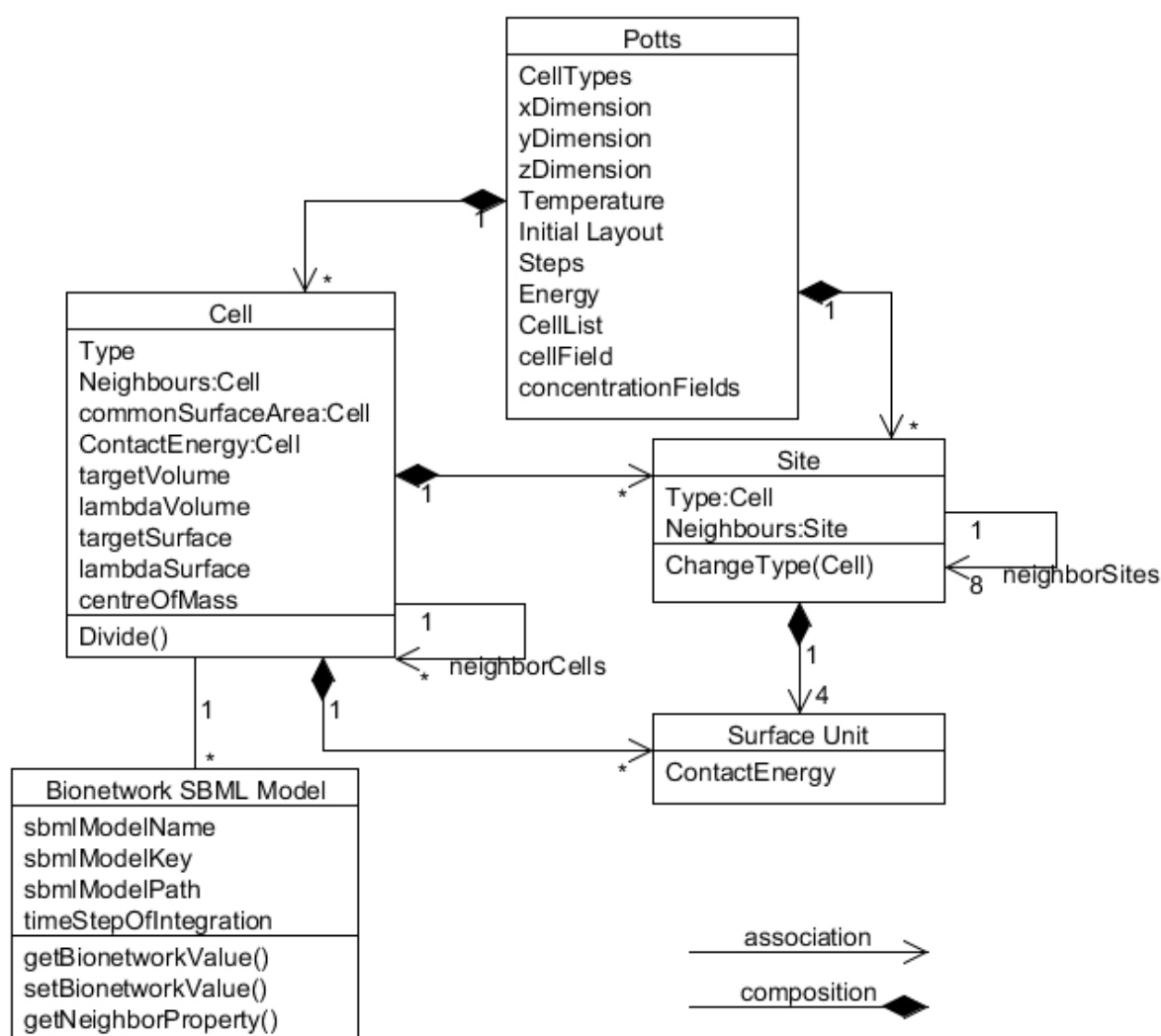


Figure 3.10 UML class diagram of multiscale cellular Potts model, as implemented in CompuCell3D

In order to give a clearer impression of the objects, attributes, methods and relationships in a typical CompuCell3D model, they are summarised as a class diagram in Figure 3.10. The case of a 2D model, on a square lattice, with a `NeighbourOrder` of 2 is illustrated; as each lattice site has eight neighbours, and four surfaces. As explained previously, each surface unit has a contact energy, which is the contact energy specified between the cells on either side of the surface, and the sum of these energies form the basis of the effective energy in the simulation; to which other terms are added. Each cell consists of many sites, and each site is associated with a given cell, at a particular point in time. Each Cell may be associated with multiple Bionetwork SBML models.

There is currently no common language for agent based multicell biological models. The developers of CompuCell3D, through a series of workshops, have begun to develop a Cell Behaviour Ontology, and associated markup language with the aim of creating such a language (Galdzicki et al. 2009). However, it seems unlikely that the rules governing behaviour of agents in a model would be straightforward to define using XML. An alternative approach might be to employ a graphical language such as UML (Bersini 2012), or a graphical language based on the same principle. As for existing language interchange formats (or for that that matter, the use of UML), different levels of adoption would coexist. Some platforms might develop the facilities to generate diagrams automatically from code, or to generate code from diagrams, while in other cases diagrams might be used to manually document code, or to interpret models generated on other platforms. In this way, cell centre models developed on FLAME (Holcombe et al. 2012), or cell vertex models developed on Chaste (Pitt-Francis et al. 2009), might be more easily re-implemented as Potts models in CompuCell3D, and vice versa. The testing of models using different formalisms provides a means of verification, and additionally highlights behaviours that arise as a result of artefacts of a particular method.

3.5. Verification and Validation

Verification and validation are independent procedures that are used in combination for checking that a product or system is fit for its intended application. While verification is generally concerned with internal accuracy and consistency (“Are we building the system right?”), validation addresses the question of fulfilling the specification and meeting the intended purpose (“Are we building the right system?”). In the context of simulation

models, verification is the process of ensuring that the model is a correct implementation of the conceptual model, and checking for errors. Validation is the process of checking the accuracy with which the model represents the real system.

3.5.1. Verification

Verification of models is addressed by checking every parameter three times. Parameters are first checked in the simulation code, and the files for executing each simulation are given names that reflect the parameters used. Each parameter is checked a second time in the graphical user interface at the beginning of each simulation run. Finally, the executed code for each simulation is copied into a new directory, along with results and visualisation snapshots for that simulation. Parameters are checked a third time in this copied code before aggregating the results. Verification that the simulation model accurately reflects the conceptual model is achieved by testing the models under different parameters and checking that the behaviour is as expected. E.g. that an increase in cell-cell contact energy leads to cell separation, reduction in cell-medium energy leads to cells migrating into the medium, and an increase in Temperature increases the speed of cell movement. A potential further method of verification, though not attempted in this thesis, would be to reproduce the models using a different formalism, such as cell centre or cell vertex modelling.

3.5.2. Validation

Model validation is addressed in the first instance by constructing models with high face validity. This is achieved by showing simulation results to domain experts, and then adopting their suggestions in order to improve the face validity of the models. The domain experts in this case are the cell biologists José Luis De la Pompa and Luis Luna-Zurita, who are experts on the role of Notch signalling in EMT in heart development. These experts were selected as they were authors of the *in vitro* investigation on which the simulations are principally based (Luna-Zurita et al. 2010). Therefore, they have the best available knowledge of the system being modelled. Feedback is used to improve the face validity of the model by adopting suggested features, such as shape changes during EMT. Simulations are also shown to clinical experts (Jean-Marc Schleich and Lucile Houyel) with specialist knowledge in heart development, in order to check whether the simulations appear to be an accurate representation of EMT.

The second validation test is checking cell shape parameters in the simulations against those derived from image processing results of data provided by the authors of (Luna-Zurita et al. 2010). In this way, the simulations are improved by adjusting parameters one at a time in order to calibrate the model with the cell shapes measured in the imaging data; and cell speeds recorded in mesenchymal cells.

While the models are fitted to the available experimental data, formal model validation is not possible at this time, as only one, historical, dataset is available for this purpose; which only provides static information. Further experimental imaging studies could be employed for additional validation by altering inputs in the real system that have a direct correspondence in the model. For example, increasing density of particular cell adhesion molecules in the collagen gel could be used to increase cell-matrix adhesion by a known amount. Input-output transformations could then be compared directly between the real system and the model, and the model iteratively improved. Additionally, recording image sequences of the evolution of the real system would enable checking of the dynamic behaviour of the model, e.g. the change in cell shape, motility and contact area between cells over time.

3.6. Summary

This chapter has explained the criteria for method selection, and has also explained each of the methods used in this thesis. These include image processing, and cell level modelling and simulation, as well as methods of verification and validation. The different implementations of cellular Potts models used have been explained, as well as the technique for fitting the time and space scales, and the quantitative cell metrics used for model validation. The method for implementing multiscale models with CompuCell3D and Bionetsolver was also explained. The next chapter reports the results that were derived using this set of methods.

Chapter 4

4. Results

Throughout this chapter, simulation objects are printed in bold (e.g. **endocardial cell**) in order to distinguish them from the physical objects that they represent (e.g. endocardial cell).

4.1. Introduction

This chapter details the results gained from implementing the methods described in the previous chapter. These results are based principally on simulation, rather than data from the real system. Section 4.2 describes the 3D simulation of *in vitro* EMT, and the qualitative results of cell behaviour under different simulated conditions. Section 4.3. describes a 3D simulation of contact inhibited mitosis, based on the conceptual model of how this process regulates EMT.

Section 4.4. details the results of 2D simulations that included cell morphology as a feature. First the image processing results are given in Section 4.4.1. Then results from simulations of single cells are given (Section 4.4.2.), which explore the relationship between motility and cell morphology in the cellular Potts model. Section 4.4.3 applies these results in multicell simulations, which were used to investigate the effect of **cell-cell** and **cell-Medium** adhesion on cell morphology.

Section 4.5. reports the results of a multiscale 2D model of Notch mediated lateral induction. This reuses a published subcellular model of lateral induction.

Section 4.6. describes an approach for composite annotation that can be used to establish a link between objects and parameters in biomedical simulations and concepts defined in reference ontologies. Example annotations are given for heart development.

4.2. 3D Simulation of *in vitro* EMT

A recent *in vitro* study of endocardial cells (EC) cultured on a collagen gel demonstrated that 2D migration of cells on the collagen surface could be induced independently of 3D invasion into the gel (Luna-Zurita et al. 2010). Thus an important distinction is made between cell migration (“2D transformation”) and cell invasion (“3D transformation”). These two properties were measured separately in the *in vitro* investigation by counting the proportion of cells able to migrate in 2D and/or invade in 3D into the gel.

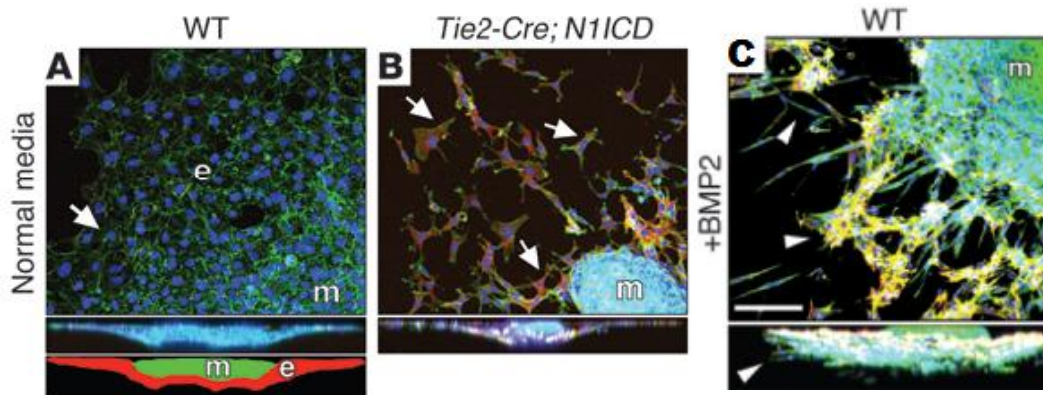


Figure 4.1 *In vitro* endocardial explants, comprising endocardium (e) and myocardium (m). a) Wildtype (WT) endocardial tissue remains in a monolayer. b) Notch activated (Tie2-Cre;N1ICD) endocardial cells migrate on the surface of the collagen gel. c) BMP2 treatment causes wildtype cells to both migrate on the surface and invade the gel (Luna-Zurita et al. 2010).

Wildtype ventricular endocardial explants from E9.5 mouse embryos remain in a monolayer when cultured on collagen gel (Figure 4.1a). This is because ventricular endocardial cells are not predisposed to EMT. 2D migration on the surface of the collagen gel could be induced by constitutively activating Notch1 in the cells, or by treating the media with TGF β 2 protein (Figure 4.1b). Anti-TGF β 2 both counteracted the effects of TGF β 2 and maintained the monolayer in Notch1 activated cells. Treatment with BMP2 induced both 2D migration and 3D invasiveness of wild type cells (Figure 4.1c). This suggests that both TGF β 2 and Notch1 in endocardium act to reduce endocardial cohesion, independently of factors that induce 3D invasion. BMP2, on the other hand, activates the mechanisms of 3D invasion, including increased endocardial-matrix adhesion and degradation of extracellular matrix proteins.

CompuCell3D simulations were created to represent these *in vitro* conditions. As the interest lay in the differences between 2D migration and 3D invasion, it was essential to

use a 3D model, to separate these two aspects of cell movement. For the initial conditions, a circular monolayer of **endocardial cells (EC)** was generated, lying on the surface of **collagen gel**, which was represented as one large homogenous cell. The default '**Medium**' cell type in the simulations is taken to be the air in the space above the cells and collagen gel, and is set to invisible for the sake of simulation renderings. The initial conditions were generated by using a simple Python script to create a simple 'cubic' arrangement, and saving this in the native CompuCell3D geometry format (pixel initialisation file: .pif). A simulation was then run for 1000 MonteCarlo steps (MCS) in CompuCell3D to allow this geometry to relax into a circular arrangement, in an energy-minimised (equilibrium) state. Under the assumption that cells behave like immiscible fluids, the cell shapes in this condition should be realistic. This 'realistic' layout was then saved to be used as the initial geometry for all further simulations. The 'cubic' layout consisted of 100 box-shaped **cells**, 10 x 10 x 2 voxels, e.g. with a volume of 200 voxels each. **ECs** were assigned a target volume of 400 voxels, and a fairly high volume constraint of 20 ($10^{-15}\text{kg}^1\text{s}^{-2}$), which ensured a consistent volume and rounded morphology typical of epithelial cells. The **gel** was given a target volume equal to its initial size ($140 \times 140 \times 20 = 392000$ voxels) and lower volume constraint value of 2 ($10^{-15}\text{kg}^1\text{s}^{-2}$), as this was found to be sufficient for it to maintain a constant volume.

As the **medium** represents the space above the cell culture, it has no intrinsic surface energy in these simulations. It is assumed that **EC-EC** adhesion is stronger in the wildtype situation than **EC-gel** adhesion, and that the latter is stronger than **gel-gel** adhesion. The contact energies with the surrounding air (**medium**) are taken to be higher than other surface energies, as there is no adhesion between the air and **cells** or **gel**; in fact there is a strong positive surface tension. It is assumed that the surface energy between **EC-medium** is higher than that between **gel-medium**. This is because cell membranes are less deformable than the surface of collagen gel, due to the comparatively rigid structure of cell walls and cytoskeletal proteins.

Therefore, to simulate wildtype **EC** on the surface of **gel**, the following energy hierarchy is assumed:

$$J_{\text{EC,medium}} > J_{\text{gel,medium}} > J_{\text{gel,gel}} > J_{\text{EC,gel}} > J_{\text{EC,EC}} > J_{\text{medium,medium}} = 0$$

Equation 4.1

For reasons explained in the methods chapter (and further explored in Appendix A) in the cellular Potts model it is the relative differences between these surface energies, rather than the absolute levels, that are important. From simulation it was found the corresponding parameters of set 1 (Table 4.1) give rise to an endocardial monolayer that does not invade the **gel**. Set 2 corresponds to a loss of endocardial cohesion: an increase in $J_{EC,EC}$ such that the surface energy between **endocardial cells** overcomes the intrinsic surface energy of the **gel**. Set 3 corresponds to a gain in **EC-gel** adhesion: a reduction in $J_{EC,gel}$ such that this is lower than the surface energy between **ECs**. Set 4 corresponds to both these effects simultaneously, thus re-ordering the energy hierarchy to be:

$$J_{EC,medium} > J_{gel,medium} > J_{EC,EC} > J_{gel,gel} > J_{EC,gel} > J_{medium,medium} = 0 \quad \text{Equation 4.2}$$

Table 4.1 Surface energy parameters J ($10^{-15} \text{Kg}^1 \text{s}^{-2}$) for *in vitro* EMT simulations.

Surface energy J	EC-Medium	Gel-Medium	Gel-Gel	EC-Gel	EC-EC	Medium-Medium
Set 1	16	14	8	4	2	0
Set 2	16	14	8	4	10	0
Set 3	16	14	8	1	2	0
Set 4	16	14	8	1	10	0

The base-case, or wildtype, scenario (Figure 4.2a) used the parameters from set 1 Table 4.1, as these were found through experimentation to produce a stable monolayer. The **cells** in the monolayer move around each other, but they do not migrate away from the monolayer, nor invade the **collagen gel**. This is consistent with the behaviour of the wildtype *in vitro* cells, demonstrating that the energy hierarchy assumptions are valid. The base-case was perturbed by adopting the parameters in Sets 2-4, and running the simulation for a further 1000 MCS in separate experiments.

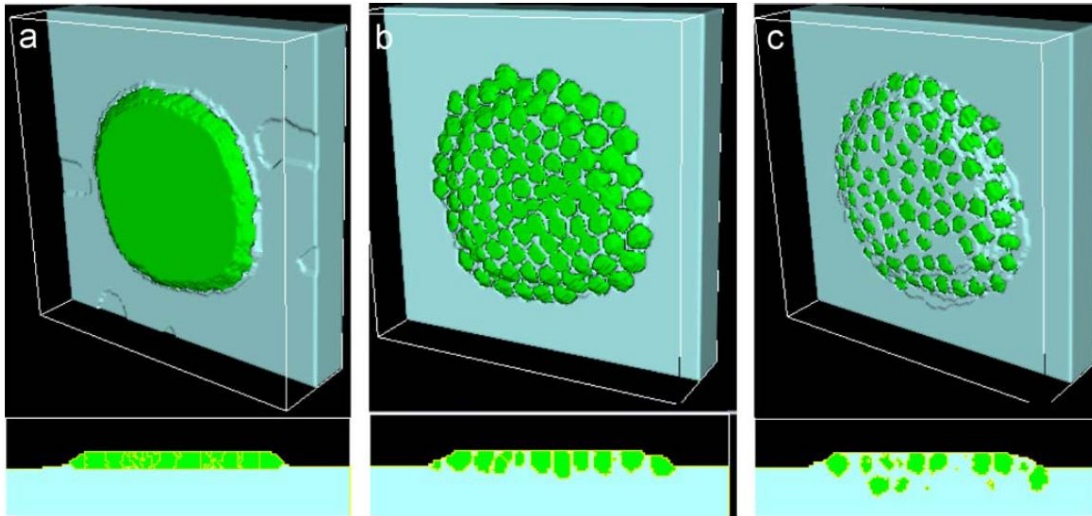


Figure 4.2 Cellular Potts simulations of *in vitro* EMT. a) Endothelial monolayer on the surface of collagen gel. b) With reduced EC-EC adhesion, cells migrate on the surface, but do not invade the gel (partial EMT). c) With reduced EC-EC adhesion *and* increased EC-Gel adhesion, some cells invade the gel in 3D (full EMT).

With the parameters in set 2, **ECs** scattered on the surface of the collagen gel without invading it (Figure 4.2b). This is consistent with the *in vitro* conditions of Notch activated cells, or a media treated with TGF β 2, suggesting that a loss of endocardial cohesion is a sufficient model of the behaviour exhibited in these examples. With the parameters in set 3, the **ECs** invaded the **collagen gel** but without delaminating from each other (results not shown). This behaviour was not exhibited by any of the *in vitro* conditions in this example (Luna-Zurita et al. 2010). However it is consistent with the behaviour of collective cell migration exhibited in other biological contexts (Ilina & Friedl 2009). With the parameters in set 4, all **ECs** delaminated from each other, and some of them invaded the **collagen gel** (Figure 4.2c). This is consistent with the 3D invasion of a proportion of cells observed in the *in vitro* explant with BMP2 treatment. This suggests that loss of endocardial cohesion and gain of EC-gel adhesion are sufficient conditions to model this invasive behaviour.

The simulations demonstrate correspondence with the *in vitro* experiments on which they are based. For example: in both the *in vitro* experiments and the simulations, it was possible to induce 2D scattering of ECs independently of 3D invasion into the collagen gel. In the *in vitro* experiments this was accomplished through Notch activation of the endocardial cells, or through TGF β 2 treatment. Alongside simulation results, this

supports the hypothesis that Notch primarily acts to reduce endocardial cohesion, without stimulating invasive mechanisms. In the simulation with the set 3 parameters, it was possible for the **EC** to invade the **gel**, but still remain attached together. This effect was not observed in these *in vitro* experiments (Luna-Zurita et al. 2010), however it is consistent with the collective cell migration as observed elsewhere. This could be because it was not practicable to isolate an increase in EC-gel adhesion from a decrease in EC-EC adhesion in real ECs, due to the nature of the signalling pathways involved. Notch is a downstream target of BMP signalling, and therefore inducing increased EC-ECM adhesion via treatment with BMP2 would have the additional effect of activating Notch signalling and reducing EC-EC Adhesion.

However, the simulations do not investigate many of the features clearly present in the *in vitro* experiments, such as the changes in cell shape from a rounded to fibroblastic morphology. These features are potentially influenced by so many different factors that it was intractable to explore the necessary parameter space in 3D simulations. Furthermore, it is hard to define appropriate metrics for quantifying 3D fibroblastic shapes, and the quality of the confocal scans was not high enough to extract 3D cell shapes, thus 3D simulations of this kind would have been impossible to validate with the available data. The role of cell morphology in 2D migration was investigated through 2D simulations, and this is reported later in this chapter.

Verification of the models in this section was addressed by checking each parameter three times: in the simulation code, in the graphical user interface and in the saved copy of the simulation code after execution. Verification that the simulation model sufficiently represented the conceptual model was achieved by testing the model under different parameter conditions, and checking that the behaviour was as expected. In this case, a reduction in cell-cell adhesion in the model resulted in cells separating on the surface of the collagen gel, without invading it. A simultaneous reduction in cell-cell adhesion and increase in cell-gel adhesion results in the behaviour of cell separation and invasion. These behaviours are consistent with those observed in the real system. Thus the model passed the verification test of representing expected behaviours of the system under different parameter conditions.

Validation of the models in this section was addressed by showing the simulation results to domain experts, to check the face validity of the models. In this case the domain experts were the cell biologists José Luis De la Pompa and Luis Luna-Zurita, who are leading researchers in the field of EMT in heart development. These subject matter experts were selected as the authors of the *in vitro* investigation upon which the simulations are principally based (Luna-Zurita et al. 2010). They therefore have the best technical knowledge available of the system being modelled. Simulation outputs were communicated in the form of screenshots under different parameter conditions and explanation of how the model functions. It was the domain experts' suggestion to include the feature of shape changes in EMT. This suggestion was adopted and explored as reported on in Section 4.4. Thus feedback from domain experts directly helped to improve the face validity of the model.

4.3. 3D Simulation of Contact Inhibited Proliferation

An abstract scenario was used for investigating the potential role of contact inhibited mitosis. In this case, the simulations are used to dynamically represent the conceptual model, and thereby improve understanding; but there is no available experimental data that can be used to directly validate the simulation models. The existing findings are that EMT in the endocardial cushions is tightly controlled by the expression of VEGF. One conceptual model of this is that insufficient VEGF leads to depletion of endothelial cells, and thus insufficient EMT; while too high an expression of VEGF leads to a sustained endothelial phenotype (Armstrong & Bischoff 2004). These simulations employ an abstract scenario in order to test the possibility that the epithelial phenotype can be maintained as a direct result of a high level of contact-inhibited proliferation. An **endocardial monolayer** of 100 10x10x2 **cells** was defined as occupying the entire midplane between two layers of default **medium**. Surface energy parameters were adapted from those in the previous section so that the **medium** would now represent collagen gel (

Table 4.2). The meaning of such parameters are explained in Appendix A. The **medium** was set as invisible for simulation renderings, to enable visualisation of **cell** migration in 2D and **cell** invasion in 3D.

Table 4.2 Surface energy parameters J ($10^{-15}\text{Kg}^1\text{s}^{-2}$) for contact-inhibited mitosis simulations.

Surface energy J	EC-Medium	EC-EC	Medium-Medium
Set A	16	2	0
Set B	4	10	0
Set C	2	2	0
Set D	2	10	0

The mechanisms by which epithelial cells in a monolayer regulate mitosis are not precisely known. For these simulations, it was assumed that mitosis is regulated by some form of contact inhibition. The CompuCell3D *NeighbourTracker* plugin and *Mitosis* Python steppable were adapted such that a simulated **EC** will undergo mitosis if it meets the condition that the surface area it shares with the **medium** is greater than the surface area it shares with other **ECs**. **Cells** were also required to have a volume greater than the initial volume of 200 voxels to undergo mitosis, in order to prevent excessive mitosis of small **cell** fragments. Simulated **ECs** were again assigned target volumes of 400 voxels, and a volume constraint of 20 ($10^{-15}\text{kg}^1\text{s}^{-2}$). This led to rapid cell growth until **cells** were either in contact with other **cells** (over 50% or more of their surface), or ready to divide again (200 voxels or more in volume). These rules are summarised in Figure 4.3.

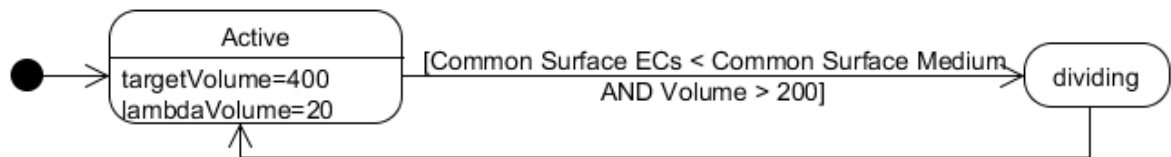


Figure 4.3 UML state machine diagram for contact inhibited mitosis model.

The mitosis scenarios produced results analogous to those of the 3D *in vitro* EMT simulations of the previous section. Set A parameters represent a high level of **endocardial** cohesion and a low level of **endocardial-gel** adhesion. With these parameters the **endothelial monolayer** was fully maintained with and without contact inhibited **mitosis** (results not shown). Set B parameters represent a loss of endocardial cohesion. The **ECs** delaminated from each other with these parameters (Figure 4.4a), but the inclusion of contact-inhibited **mitosis** caused the **endothelial monolayer** to prevail (Figure 4.4b). Set C parameters represent a gain in **EC-gel** adhesion, without a loss in **EC-EC** adhesion. With these parameters, the **ECs** invaded in 2D and 3D after 1000 MCS. The inclusion of contact inhibited **mitosis** caused the monolayer to prevail (Figure 4.4d). Set D parameters represent a gain in **EC-gel** adhesion and a loss in **EC-EC** adhesion. With these parameters, the **ECs** invaded in 2D and 3D after 1000 MCS. The inclusion of contact-inhibited **mitosis** led to the monolayer failing under the parameters of set D. Due to the conditions for **mitosis** specified, cells rapidly fill the entire lattice (results not shown). In this model, mitosis prevents breakdown of the monolayer under the conditions of reduced **EC-EC** adhesion or increased **EC-gel** adhesion, but not in combination.

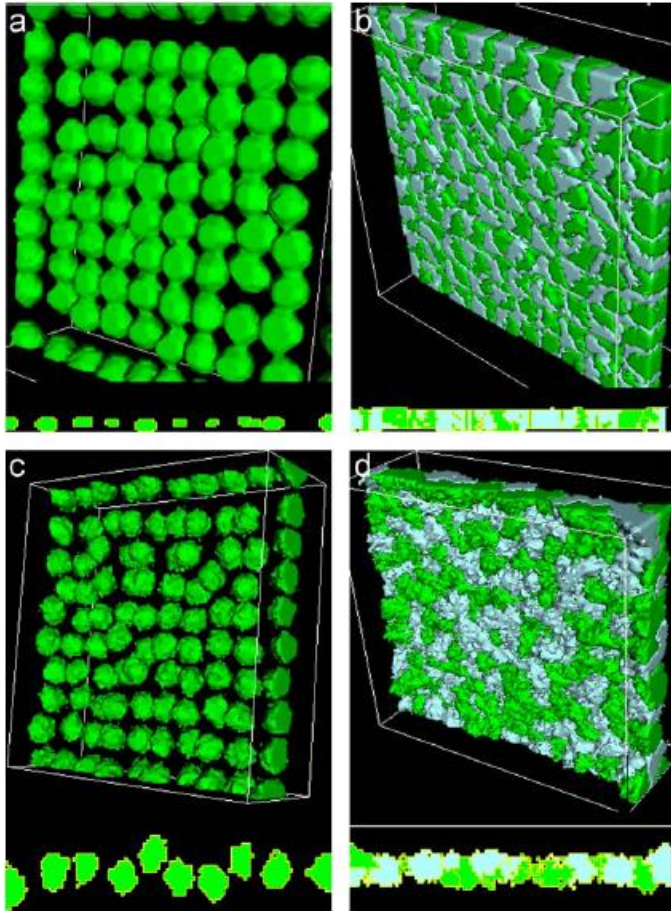


Figure 4.4 Mitosis simulations. a) Cells separate in 2D under set B parameters. b) monolayer prevails under set B if contact-inhibited mitosis is included. c) Cells migrate in 2D and 3D under set C parameters. d) Including contact-inhibited mitosis rescues monolayer integrity for set C. Daughter cells are illustrated in a different shade in order to highlight the effects of mitosis.

Although in this simulation mitosis was treated as a lumped variable that occurs instantaneously, the results demonstrate a plausible mechanism by which VEGF could control the level of EMT, by increasing the rate of contact-inhibited mitosis.

Verification of the models in this section was addressed by checking each parameter three times: in the simulation code, in the graphical user interface and in the saved copy of the simulation code after execution. Verification that the simulation model sufficiently represented the conceptual model was achieved by testing the model under different parameter conditions, and checking that the behaviour was as expected. In this case, the inclusion of contact inhibited mitosis led to behaviour of preservation of the monolayer in parameter conditions where this preservation would otherwise be lost. While this

provided agreement with the experimental hypothesis that contact-inhibited mitosis plays a role in preserving the endocardial monolayer; there is insufficient knowledge of how exactly this is achieved and implemented in the real system for further verification and validation of this model at the current time.

4.4. 2D Simulations of Cell Morphology

This section focuses on the same *in vitro* results of (Luna-Zurita et al. 2010), but narrows the analysis to 2D migration, in the absence of 3D invasion. The condition of cell invasion into the gel is not represented in this section, and the assumption is made that cells are relatively thin in cross-section, and move around, instead of over or under, each other. These assumptions allow for the simplification of 2D modelling to be used. This also allows for validation of the model from shape metrics derived from image processing of the *in vitro* imaging data. The simulations, and the images analysed, focus on the differences between wildtype and Notch activated *in vitro* endocardial cells, as these migrated only in 2D. As real endocardial cells are very thin in cross section, the simplification still allows for meaningful results. **Cells** are assumed to be uniformly 1 μ m (1 voxel) thick, allowing a direct representation of cell geometry, with cells approximately 15 μ m (15 voxels) wide. These dimensions are based on the sizes of cells measured in the confocal images of wildtype cells.

4.4.1. Image Processing Results

The image processing results indicated that there is a significant difference in shape in terms of circularity, but not aspect ratio, between wild type and Notch1 activated ventricular endocardial cells. As shown in Figure 4.5, the N1ICD cells had a significantly less circular morphology, but were not found to be significantly elongated compared to the wild type endocardial cells.

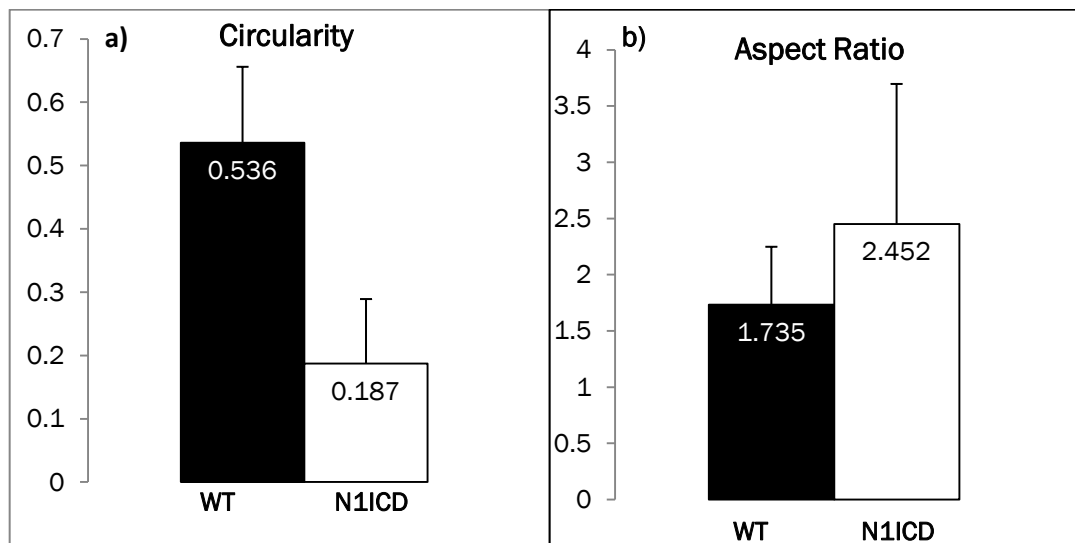


Figure 4.5 Shape metric comparison of wild type and N1ICD murine ventricular endocardial cells. a) Wildtype cells have a mean circularity of 0.536, and N1ICD cells 0.187 ($p < 0.001$). b) Wild type cells have a mean aspect ratio of 1.735 and N1ICD cells 2.452 however this difference is not significant. Error bars show standard deviation.

However, these results do not tell us anything about the mechanism by which Notch signalling induces a shape change in these cells. It is likely to be a combination of many factors. For example, Notch signalling induces a loss of VE-Cadherin protein, which means a loss of endocardial adherens junctions. A secondary effect of cells being able to migrate away from the epithelial layer is that they are no longer pulled into an approximately hexagonal shape by other epithelial cells. Thus loss of contact and adhesion with other endocardial cells might account for some of the loss of circularity. Additionally, there is significant crosstalk between the Notch and TGF- β signalling pathways (Fu et al. 2009), and TGF- β 2 expression is upregulated in Notch1-activated ventricular endocardial cells (Luna-Zurita et al. 2010). TGF- β signalling upregulates several integrin adhesion molecules, which bind to extracellular matrix components such as collagen, laminin, vimentin and fibronectin (Margadant & Sonnenberg 2010). This increased adhesion to the matrix might pull cells into a spindly morphology, as well as increasing their migrative capacity. Finally, Notch and TGF- β induce a relocation of E-cadherin and cortical actin from cell walls to the cytoskeleton, meaning a stronger cell, with more ability to survive under mechanical stress, and thus change shape (Zavadil et al. 2004). The potential role of these different effects was investigated through simulation models, as reported in the next section.

4.4.2. Single Cell Simulation Results

CompuCell3D simulations of individual **cells** were used to simulate changes in cell shape. Starting from a rounded morphology, the effects of an increased surface perimeter (relative to volume), and increased motility were simulated. Motility was increased by raising the 'Temperature' parameter in the simulations, which increases the number of pixel copies accepted, and thus the speed that simulated **cells** move. Changes in shape led to greater motility, while greater motility led to changes in cell shape. This indicates that cell shape and migration are highly interdependent in the model. In each case, the base case of a rounded morphology is used, before changing one parameter at a time. Base case simulations used the parameters: $\lambda_{\text{Volume}}=3.0$, $\lambda_{\text{Surface}}=3.0$, $\text{targetVolume}=225$, $\text{targetSurface}=60$, $\text{Temperature}=10$. These parameters maintain an approximately circular cell shape, with little deviation from this. Contact energy with the surrounding Medium is set to 0. **Cells** are initialised as a square shape of 15x15 pixels, which represents a 2D surface area (volume) of $225 \mu\text{m}^2$ per cell, assuming a constant thickness of $1 \mu\text{m}$.

Larger Surface Area Induces Greater Motility

An increase in surface area, relative to volume, can be induced by increasing the target surface parameter, while keeping other parameters constant. This constrains **cells** into adopting a fibroblastic (spindle-shaped) morphology. However this change in morphology is also accompanied by an increase in motility, as shown in Figure 4.6. This can be explained by the greater number of interactions between cell surfaces and medium leading to a greater number of pixel copies being attempted and accepted. This has the biological equivalent of a cell with a more fibroblastic morphology having a greater surface area over which to interact with and adhere to the matrix.

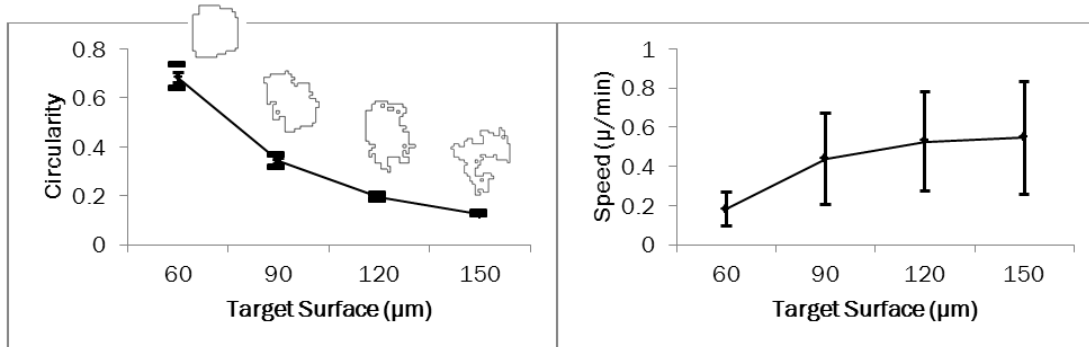


Figure 4.6 a) Circularity of simulated cells falls with increasing Target Surface. b) This is accompanied by increased speed (and hence motility). Error bars show standard deviation, and caps show the range, from 10 simulation replicas.

Greater Motility Induces Fibroblastic Morphology

An increase in **cell** motility can be directly induced by increasing the ‘Temperature’ parameter in the simulations, as a greater proportion of pixel copy attempts are accepted. However, this change in motility is accompanied by the adoption of a fibroblastic morphology. Furthermore, the magnitudes of reduced circularity associated with increased speeds are similar to those that result from increasing the target surface (compare Figure 4.6 with Figure 4.7). In both cases, migration speeds are consistent with those of *in vitro* epithelial and mesenchymal cells (Mendez et al. 2010).

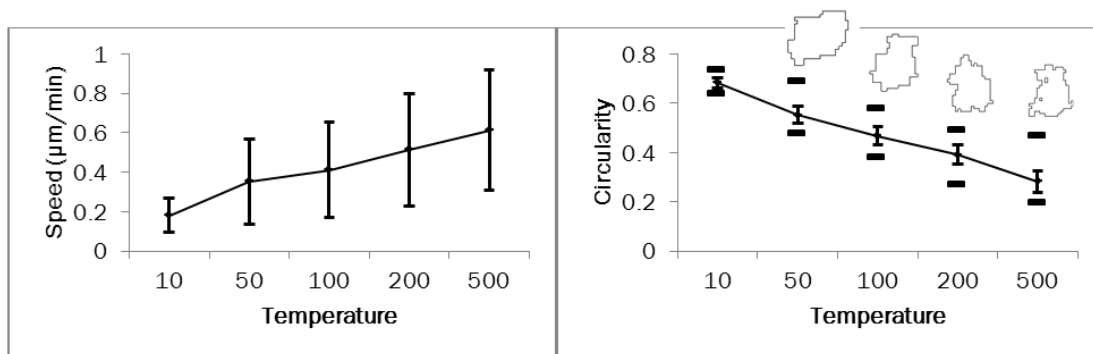


Figure 4.7 a) Speed of simulated cells increases as a result of increasing ‘Temperature’. b) This is accompanied by a reduction in circularity. Error bars show standard deviation, and caps show the range, from 10 simulation replicas.

There are two plausible explanations for this. Either the increase in Temperature has the side effect of giving the cells the flexibility to deviate from a rounded morphology; or the easiest (lowest entropy) way for a cell to be more motile is for it to adopt an

elongated or fibroblastic morphology. This has the biological equivalent of a cell changing shape in order to 'squeeze through' the matrix.

4.4.3. Multicell Simulation Results

Multicell simulations were performed, corresponding to the dimensions of the endocardial explants reported in (Luna-Zurita et al. 2010). Each simulation consisted of 316 cells, initially arranged in a filled circle. As with the single cell models, each **cell** has initial dimensions of 15x15 pixels (225 μm^2), with target volume parameters that maintain this size. In our multicell model we consider that there are four factors that might guide EMT: loss of epithelial labile adhesion, loss of epithelial junctional adhesion, increase in motility and increase in adhesion to the surrounding **medium**. For simplification, we assume that both types of epithelial adhesion would be lost during EMT, as a result of downregulation of VE Cadherin.

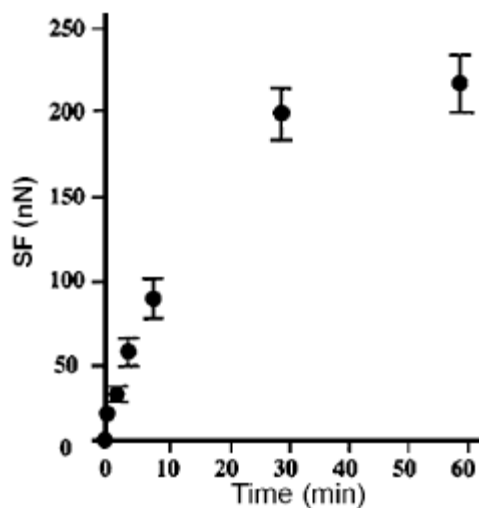


Figure 4.8 Separation force measurements for Ecad cells held in contact for 0.5-60min (Chu et al. 2004).

The adhesion force required to separate two E-cadherin expressing cells is initially of the order of a few nanonewtons (nN). However, this adhesion strength increases rapidly between 30 seconds and 30 minutes, and this is followed by a slower increase up to an hour, reaching a force over 200 nN. The profile of increasing adhesion strength is shown in Figure 4.8. The initial E-cadherin mediated contact adhesion doesn't require connection to the actin cytoskeleton. The stronger junctional adhesion forms over a

longer period of contact by a connecting the actin cytoskeleton between the two cells (Chu et al. 2004). For these simulations, it is assumed that the properties of VE-cadherin binding are qualitatively analogous to those of E-cadherin binding. This assumption is based on the fact that both form adherens junctions, and both bind to the same cytoskeletal proteins, including catenin and actin.

The two phases of adhesion (weak labile adhesion followed by strong junctional coupling) can be included in the model as two types of force, to investigate their individual roles in driving epithelial morphology. Labile adhesion is represented as a negative surface tension (contact energy) between **cells**, determining the likelihood of pixel copy attempts that separate or connect two adjacent **cell** boundaries. Junctional coupling is represented as a breakable spring force between **cell** centres, operating between **cells** whose centres are within a set maximum distance. In this way the model combines elements of the cellular Potts and cell centre models. In order to capture the difference in magnitude between adhesion phases, we generally set labile adhesion energy between 0 and -20 ($10^{-15}\text{Kg}^1\text{s}^{-2}$), and the strength of plastic coupling lambda between 0 and 200 ($10^{-15}\text{kg}^1\text{s}^{-2}$).

Labile Adhesion and Junctional Coupling

Multicell simulations incorporated terms for both plastic coupling and junctional adhesion in order to perform a basic parameter scan and investigate the role of these forces, and **cell** morphology, in maintaining an epithelial tissue. For **endothelial cells**, a target surface of 60 μm is used, while for **mesenchymal cells** a target surface of 120 μm is used, based on the results of the single cell simulations reported previously.

It was postulated that labile adhesion and junctional coupling might be able to rescue epithelial morphology, for **cells** that were given mesenchymal shape parameters. However, as shown in Figure 4.9, this was not the case. **Cells** with junctional coupling and/or labile adhesion neither maintained contact, nor formed a more rounded morphology. Higher strengths of plastic coupling produced incoherent results, and unrealistic **cell** shapes (not shown).

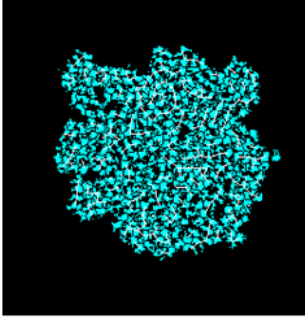
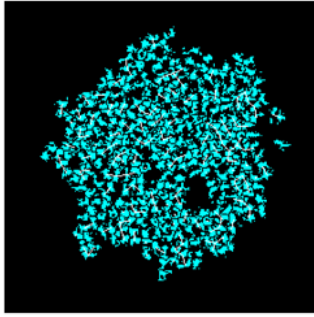
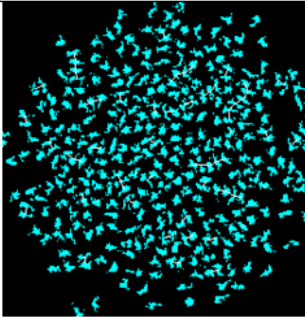
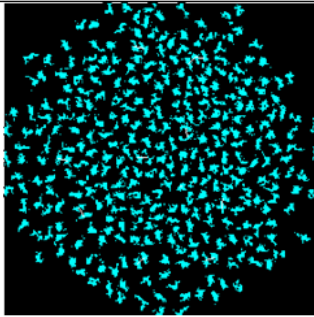
Plastic Coupling Lambda Cell-Cell Contact Energy	20	5
-15		
0		

Figure 4.9 Cells with fibroblastic morphology. Simulation snapshots at 50000 MCS, for different levels of plastic coupling lambda [20, 5] and cell-cell contact energy [-15, 0] (units are $10^{-15}\text{kg}^1\text{s}^{-2}$).

These results suggest that some inherent tendency towards a rounded morphology is required, in addition to the mechanisms of epithelial adhesion. It was also hypothesised that epithelial roundness could be shown to be an emergent property of cellular adhesion from simulation, in **cells** with epithelial shape parameters. However, neither junctional coupling nor labile adhesion significantly increased the roundness of these simulated **cells** (Figure 4.10).

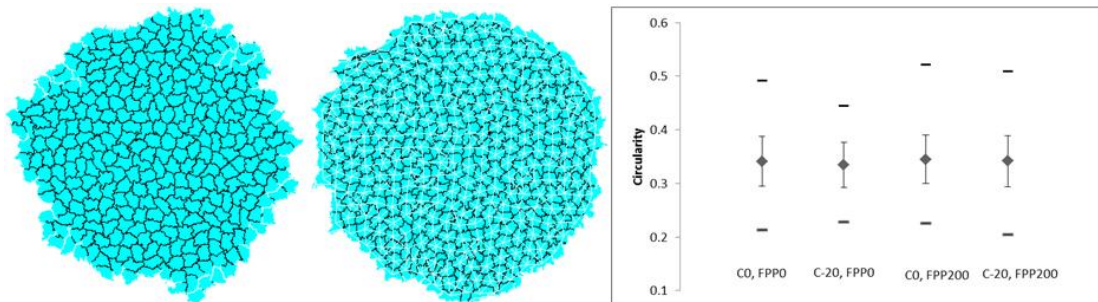


Figure 4.10 a) Monolayer with no epithelial adhesion and b) monolayer with both labile and junctional adhesion, after 50000 MCS. c) Labile adhesions, C, and junctional coupling, FPP, do not significantly affect average cell roundness ($p > 0.05$, $n=316$ in all cases).

This suggests that, while a loss of epithelial adhesion may be necessary for the cell migration and shape changes observed in EMT, it is not sufficient. Increased motility and/or increased **cell-medium** adhesion are also required.

Increase in Motility

Based on the results of the single cell simulations, it was postulated that an increase in cellular motility (Temperature parameter), coupled with a loss of epithelial adhesions, might be sufficient to induce an EMT process. In other words: that inherent shape changes might not be a necessary condition, provided there is both an increase in motility and a loss of epithelial adhesion. This was not the case, as the epithelial integrity and morphology was preserved even under large Temperature increases (from 20 to 500).

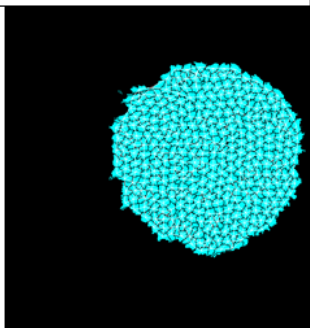
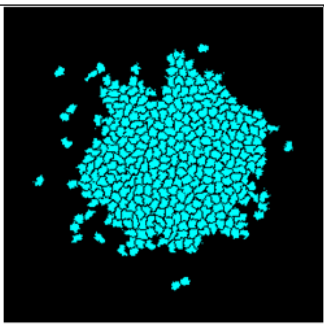
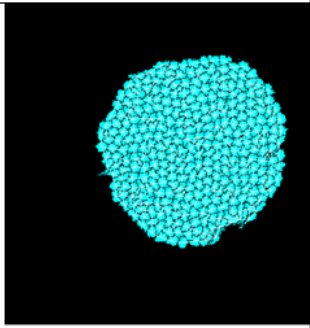
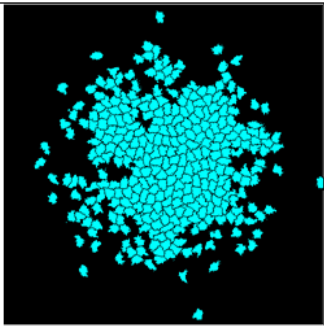
Plastic Coupling Lambda Cell-Cell Contact Energy	200	0
-20		
0		

Figure 4.11 Under conditions of increased motility (Temp=500) and moderate cell-medium adhesion (-20), cells scatter under conditions with a loss of junctional coupling, but not with a loss of labile adhesion alone. In all cases cells maintain a rounded morphology (units are $10^{-15}\text{kg}^1\text{s}^{-2}$).

However, when this motility increase was coupled with a moderate increase in **cell-medium** adhesion, an intermediate phenotype is observed, whereby a proportion of **cells** migrate, while maintaining a rounded morphology (Figure 4.11). The increase in **cell-medium** adhesion is modelled by a reduction in **cell-medium** contact energy from 0 to -20. The intermediate phenotype only occurred in the cases with a loss of junctional adhesion, illustrating that junctional adhesion has a greater effect in maintaining epithelial morphology in this case. Thus while plastic coupling was not found to influence cell shape, it did affect whether cells all stayed attached in a monolayer.

The rounded morphology observed in these migratory **cells** is consistent with an intermediate phenotype observed in endocardial cells for which some Notch target genes (Hey1 or Hey2) have been deleted (Fischer et al. 2007). This suggests that Notch signalling has multiple phenotypic effects in these cells. Loss of cell-cell adhesion is mediated by loss of VE-Cadherin (in endocardial explants), while increase in cell-gel adhesion is mediated by upregulation of vimentin and fibronectin. These different effects are induced to a greater or lesser extent by the different Snail genes; which are Notch targets (Barrallo-Gimeno & Nieto 2005). In the next section we turn to the effects of an increase in **cell-medium** adhesion.

Increase in Cell-Medium Adhesion

A strong **cell-medium** adhesion was added to the model by reducing the **cell-medium** contact energy parameter from 0 to -200 ($10^{-15}\text{kg}^1\text{s}^{-2}$). This case, with combined loss of **endocardial** adhesion, increase in motility and strong increase in **cell-medium** adhesion, leads to migration and **cell** morphology changes consistent with those induced in Notch activated mesenchymal cells (see Figure 4.12).

Simulated and measured values for aspect ratio and circularity were significantly different ($p < 0.01$), indicating that the model does not perfectly represent the morphological characteristics of Notch activated cells. There is not available data to make the comparison for migration speeds. While the values are not a perfect fit, they are consistent with the order of magnitude for the three EMT characteristics. The next section investigates the potential role for cell elongation to improve the model of EMT.

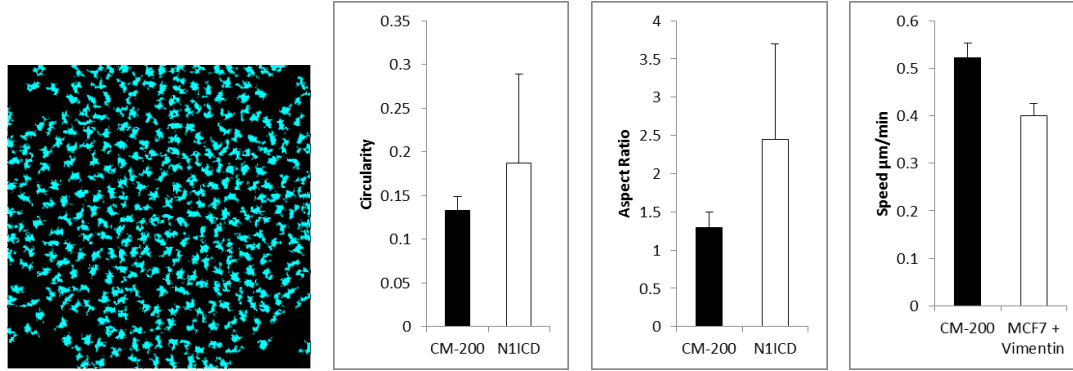


Figure 4.12 Combined loss of cell adhesion, gain of cell-medium adhesion and increased motility simulation a) Snapshot at 50000 MCS b,c,d) Mean and standard deviation of simulated (CM-200) cell characteristics at 50000 MCS (n=316) compared with those obtained from *in vitro* Notch activated endocardial cells (N1ICD, n=54) (Luna-Zurita et al. 2010) or MCF-7 cells transfected with vimentin (n=5) (Mendez et al. 2010).

Cell Elongation

In order to achieve the large aspect ratio observed in the N1ICD cells, a target length of 50 pixels (50µm) was set, with a constraint value of 5.0 ($10^{-15}\text{kg}^1\text{s}^{-2}$), for all cells in the simulation. Additionally, to increase circularity, while reducing speed, the **cell-medium** contact energy parameter was increased from -200 to -100 ($10^{-15}\text{kg}^1\text{s}^{-2}$). This produced cells which were not significantly different from the *in vitro* N1ICD cells (Figure 4.13).

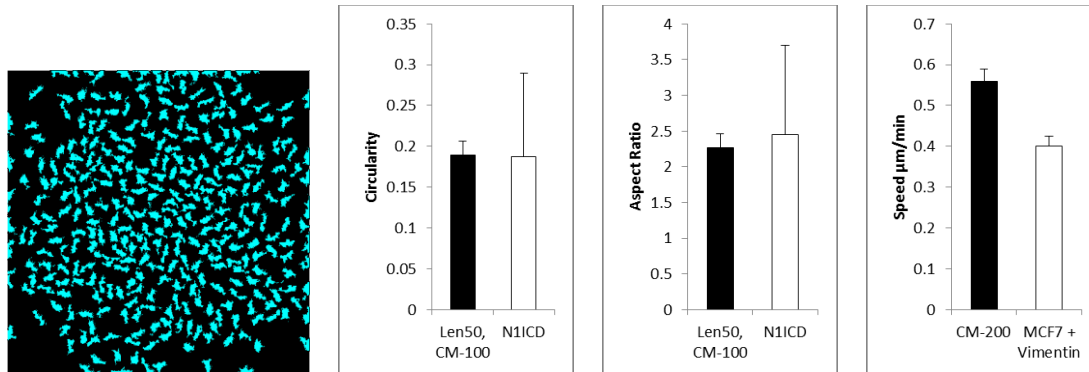


Figure 4.13 Addition of an elongation parameter to the model, and reduction of cell-medium adhesion from -200 to -100 ($10^{-15}\text{kg}^1\text{s}^{-2}$), provides a better fit, a) Snapshot at 50000 MCS b,c,d) Mean and standard deviation of simulated cell characteristics at 50000 MCS (n=316) compared with those obtained from *in vitro* notch activated endocardial cells (N1ICD, n=54) (Luna-Zurita et al. 2010) or MCF-7 cells transfected with vimentin (n=5) (Mendez et al. 2010).

Cell speed was not significantly affected by the reduction in **cell-medium** adhesion, though it did increase circularity, making the cells morphologically consistent with *in vitro* N1ICD cells ($p>0.05$). The elongation term increased the average aspect ratio of simulated cells, so that they were not significantly different from the *in vitro* model ($p>0.05$).

4.4.4. Verification and Validation

Verification of the models in this section was addressed by checking each parameter three times: in the simulation code, in the graphical user interface and in the saved copy of the simulation code after execution. Verification that the simulation model sufficiently represented the conceptual model was achieved by testing the model under different parameter conditions, and checking that the behaviour was as expected. In this case, in increase in the Temperature parameter increased cell motility, and also led to a more fibroblastic shape; and these same behaviours were observed with an increase in the Target Surface parameter. This interdependency of cell shape and cell motility is consistent with cell behaviour observed in the real system of EMT.

Validation of the models in this section was addressed by showing the simulation results to domain experts to check the face validity of the models. In this case the domain experts were the cell biologists José Luis De la Pompa and Luis Luna-Zurita; authors of the original *in vitro* investigation on which the simulations are based (Luna-Zurita et al. 2010). Model outputs were shared in the form of screenshots and quantitative results as presented here; along with explanation of how the simulations work. The domain experts indicated that the model provides a representation in good qualitative agreement with the cell changes that take place during EMT.

The second validation test was checking cell shape parameters in the simulations against those derived from image processing results of data provided by the domain experts from their previous investigations (Luna-Zurita et al. 2010). In this way, the simulations were improved by adjusting parameters one at a time in order to calibrate the model with the cell shapes measured in the imaging data; and cell speeds recorded in mesenchymal cells.

In the multi cell simulations, an increase in cell-medium adhesion was required in addition to loss of endocardial adhesion and increase in the Temperature parameter, to achieve cell shapes and motility consistent with those in notch-activated cells. This is consistent with the conceptual model of notch-activated EMT as comprising both loss of cell-cell adhesion and gain of cell-matrix adhesion. The addition of an elongation term was also required to achieve shape parameters consistent with those measured from the

in vitro results. Further experimental investigation could be directed at refining the conceptual model, to investigate whether directly driven shape changes such as elongation occur during EMT.

4.5. Multiscale Modelling of Notch Mediated Lateral Induction

Notch signalling patterns the vertebrate heart into cushion and non-cushion forming regions via the mechanism of lateral induction (Timmerman et al. 2004). The majority of wet-lab research and computer simulations of Notch signalling focus on the other well-known patterning mechanism: lateral inhibition (Collier et al. 1996; Podgorski et al. 2007; Sprinzak et al. 2011). The literature review of this thesis identified only one existing model of Notch lateral induction (Owen et al. 2000). This has also been re-used in a multiscale cellular Potts model to indicate that control of the size of epidermal stem cell clusters can be achieved by lateral-induction (Savill & Sherratt 2003). In this section, the Notch lateral induction model (Owen et al. 2000) is reimplemented as a multiscale model in CompuCell3D, in order to investigate some of the fundamental systemic aspects of lateral induction patterning within a field of tissue. The implications of these emergent properties for the regulation of cardiac morphogenesis are also assessed.

The Notch lateral induction pathway model was translated into Systems Biology Markup Language (SBML) using the JDesigner tool in Systems Biology Workbench. The model equations are as follows:

$$\frac{d}{dt}D = C_1 + C_2\{N\} - (1 - \langle N \rangle)D + C_3 \langle N \rangle - C_4D \quad \text{Equation 4.3}$$

$$\frac{d}{dt}N = (1 - N) \langle D \rangle - C_3N - C_5D \quad \text{Equation 4.4}$$

Where the variable D is the expression of Delta protein, N is the level of Notch activation in a cell (the amount bound with a ligand). Angle brackets: $\langle \rangle$ indicate the average level of a variable taken over all the *neighbouring cell* surfaces. Curly brackets: $\{ \}$ indicate a variable divided by the **cell** surface area. In this case this was implemented by a function

in CompuCell3D which iterates over all neighbour cell surfaces, adding the values of the variables together before dividing by the total surface area of the cell.

Thus the expression of free Delta depends on the background production of Delta in the **cell**, the production of Delta induced by (directly proportional to) the average level of Notch activation of neighbour **cells**, the proportion of unbound Notch in neighbour **cells**: $(1-\langle N \rangle)$ as the free Delta will bind to this Notch, the dissociation rate of Delta-Notch (as dissociation releases free Delta ligand), and the decay rate of Delta. The Notch activation (proportion of Notch binding) of a particular **cell** depends of the expression of free Delta in neighbour **cells** (as free Notch may bind to this), the dissociation rate of Notch and the Notch internalisation rate. The model constants C_1 to C_5 are given in Table 1. These use the dedimensionalised form following (Savill & Sherratt 2003) so that the expression of Notch and Delta in a given **cell** is a number between 0 and 1.

Table 4.3 Parameter values of the lateral induction model (N. J. Savill & J. A. Sherratt 2003)

Symbol	Meaning	Value
D	Delta Expression	Variable
N	Notch activation (bound Notch)	Variable
C_1	Background production of Delta	0.001
C_2	Delta production rate	1
C_3	Dissociation rate of Delta-Notch	1
C_4	Delta decay rate	1
C_5	Notch internalisation rate	0.1

The two equations of the lateral induction model were defined in the JDesigner tool through a combination of the graphical model editor and manually entering kinetic rate laws for each reaction, as shown in Figure 4.14. Fixed (boundary) species are used for the input and output of Notch and Delta, so that the two differential equations could be easily defined as a set of two reactions each.

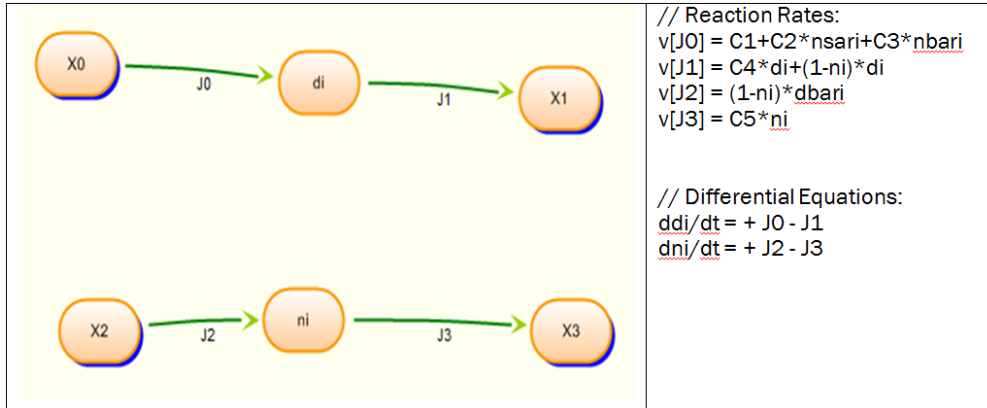


Figure 4.14 Notch Lateral induction model in JDesigner and reaction equations. ni: active Notch (NICD), di: free (unbound) Delta. Other variables are defined in the text.

Where it is necessary to calculate the average variable taken over cell surfaces, new parameters are defined: nsari, nbari and dbari. A Python steppable and the Bionetsolver API, are used to calculate these parameters at each integration of the SBML model. The parameter nsari represents the Notch activation in a given cell, divided by the surface area of that **cell**. The parameter nbari represents the sum of Notch activation for all *neighbouring cell* sites, divided by the surface area. The parameter dbari represents the sum of Delta for all neighbouring **cell** sites, divided by the surface area. This was achieved specifically with the following section of Python code (where the SBML model has previously been loaded and given the key DN):

```
for cell in self.cellList:
    weightedSumOfNeighborDeltaValues = 0.0
    weightedSumOfNeighborNotchValues = 0.0
    neighborContactAreas =
bionetAPI.getNeighborContactAreas(cell.id)
    neighborDeltaValues =
bionetAPI.getNeighborProperty("DN_di", cell.id)
    neighborNotchValues = bionetAPI.getNeighborProperty("DN_ni",
cell.id)
    nsari = bionetAPI.getBionetworkValue( "DN_ni", cell.id )
    for neighborID in neighborContactAreas.keys():
        weightedSumOfNeighborDeltaValues +=
(neighborContactAreas[neighborID] *
neighborDeltaValues[neighborID])
        weightedSumOfNeighborNotchValues +=
(neighborContactAreas[neighborID] *
neighborNotchValues[neighborID])
```

```

        bionetAPI.setBionetworkValue( "DN_dbari",
weightedSumOfNeighborDeltaValues/cell.surface, cell.id )
        bionetAPI.setBionetworkValue( "DN_nbari",
weightedSumOfNeighborNotchValues/cell.surface, cell.id )
        bionetAPI.setBionetworkValue("DN_nsari", nsari/cell.surface,
cell.id)

```

An SBML model is assigned to each cell in the simulation. Some SBML parameters are set as a function of **cell** level properties (e.g. surface area) and calculated at each integration of the SBML model in CompuCell3D. A scalar field was also defined so that the level of Notch activation across the **tissue** could be visualised. This was achieved calling the Bionetsolver API with the following code fragment:

```

for cell in self.cellList:
    fillScalarValueCellLevel(self.scalarField, cell,
bionetAPI.getBionetworkValue( "DN_ni", cell.id ))

```

The tissue level was modelled as a 2D circular tissue layer in CompuCell3D, with **cells** initialised with a 5x5 pixel volume. The simplification of modelling in 2D allowed the single mechanism of lateral induction to be investigated from a multiscale approach without introducing additional complexity. To represent the process of EMT, e.g. cells leaving the 2D endocardial layer as they invade the cardiac jelly the appropriate **cells** are set to disappear. The cells that undergo EMT in reality is likely a stochastic process, determined by thresholds in protein concentrations that a carefully controlled number of cells will cross. Parameters for this threshold were based on the outcome of the simulation, so that only a few **cells** would undergo **EMT** each time this function was called. Any **cell** that had a Notch activation above 0.88 was given a target volume of 0, so that it quickly disappeared. To replace migrated **cells** with new **cells**, any **cell** with a volume greater than 26 pixels was set to undergo **mitosis**, and daughter **cells** inherited a target volume of 25 pixels. Basic parameter adjustment was performed so that the number **cells** would stay approximately constant. Real endocardial cushions expand over time; however this feature is not represented in the current model.

Scale linking between the pathway and tissue level models is achieved by assigning a copy of the SBML pathway model to each cell in the tissue simulation, and integrated

several times at each timestep in the simulation. The pathway model therefore represents the overall level of Delta and Notch proteins in a given cell. The variables in the SBML models are accessible through the Bionetwork API in CompuCell3D, so it is possible to dynamically plot what is happening to the level of protein expression in each individual cell. It would also be possible to plot cell level properties (e.g. cell surface area, distance migrated) against protein level variables. The juxtacrine signaling between cells is captured by iterating over the surfaces that neighbor each cell, and thereby calculating the average density of Notch and Delta that each cell is exposed to at its membrane. For simplicity, it is assumed that proteins are equally distributed within each cell, as well as on their surfaces. To capture the magnitude of difference in speed between protein and cellular dynamics, the SBML models are integrated 10 times each timestep, while cell divisions and EMT only permitted every 30 timesteps.

As would be expected for lateral induction, the level of Notch activation is highly positively correlated with the level of Delta expression in each individual **cell** (Figure 4.15). Moreover the concentration of both Delta and Notch was higher towards the centre of the **tissue** field. As Notch is known to have a direct influence on the adhesive properties of cells, this suggests a simple mechanism by which the endocardial cushions could autonomously focus growth toward their centres.

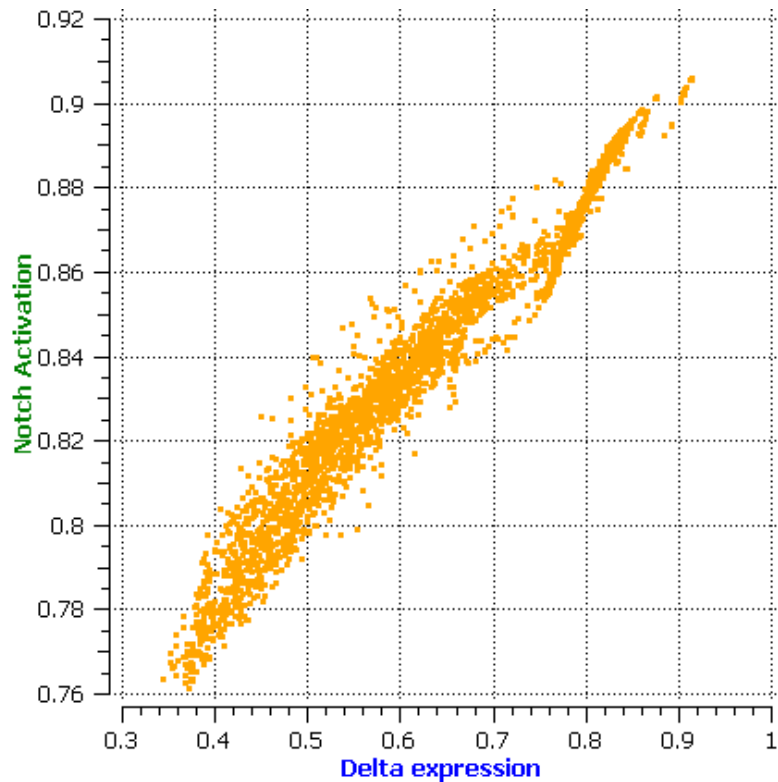


Figure 4.15 Notch activation is highly correlated with Delta expression in individual cells, as expected. Data taken from simulation as shown in Figure 4.16b over the first 800 MCS. Pearson correlation coefficient $r = 0.953$.

The reason why Notch activation is higher toward the centre of the **tissue** is because **cells** in the centre are surrounded on all their surfaces by other **cells**, and so receive a higher level of lateral induction. This also means that larger tissue areas will have a higher average Notch activation, due to the larger proportion of cells that are not close to the periphery.

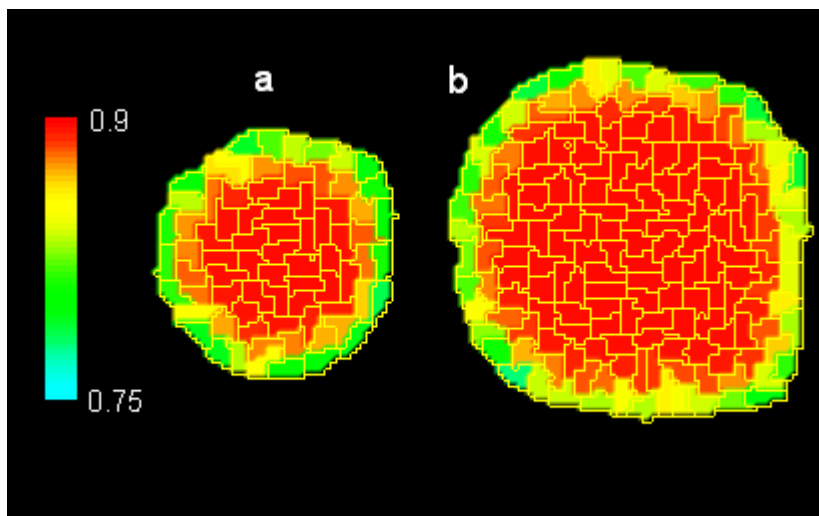


Figure 4.16 Notch activation after 800 MCS for (a) a tissue starting with a radius of 20 pixels and (b) a tissue starting with a radius of 30 pixels

This effect is illustrated by comparing two simulations which are identical except that they start with two different tissue sizes (Figure 4.16). The first has a radius of 20 pixels, and the second a radius of 30 pixels. The average Notch activation over all the cells was dynamically plotted for each (Figure 4.17).

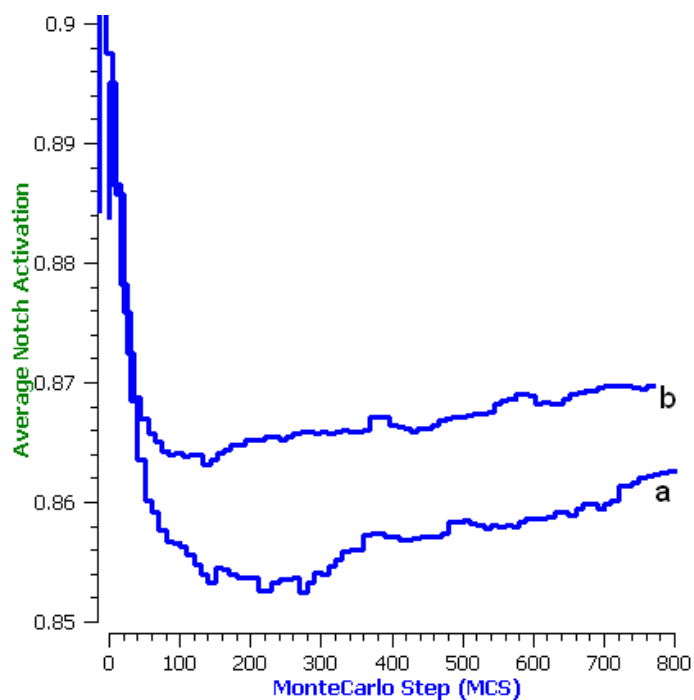


Figure 4.17 Average Notch activation over time in a tissue starting from (a) 20 pixel radius and (b) 30 pixel radius

Verification of the model in this section was addressed by checking each parameter three times, as in previous models. While the model is not validated with real biological data, the simulation results demonstrate the basic principle that larger fields of tissue will have a higher average level of Notch activation under a lateral induction mechanism. This effect would be significant in a developing embryo, as rapid growth in the endocardial cushions occurs. As shown in Figure 4.16, peripheral **cells** have lower expression of both Notch and Delta, and this could play a role in establishing tissue boundaries. Both the **tissues** grow slightly over time in the simulations, and this is because the effect of **cell** division slightly outweighs the loss of cells from the sheet as a result of **EMT**. This unintended feature is reflected in Figure 4.17, as the average Notch activation increases over time as both the fields of **tissue** grow.

The multiscale simulations demonstrate some interesting properties of the lateral induction signalling mechanism. There is a high correlation between Delta and Notch activation in all **cells** and both of these are higher towards the centre of a **tissue**. This has implications for the mechanisms of EMT in a growing endocardial cushion. It suggests that EMT would occur at a greater rate toward the centre of the endocardial cushion. It is plausible that this mechanism could be employed to autonomously regulate ‘bulging’ in the endocardial cushions.

4.6. Composite Annotation for Heart Development

As identified in the Literature Review of this thesis, heart development is a complex process that unfolds as a result of interactions between multiple levels of scale. The levels of temporal and spatial scale applicable to heart development, and methods of representation are illustrated in Figure 4.18. Computational modeling approaches that can be applied at different levels of scale are shown, as well as markup languages that enable a degree of model sharing between different platforms. The XML languages force a declarative expression of the components of a model, which allow it to be interpreted by different platforms. It is straightforward to annotate XML, and create an explicit link between entities in the model and external identifiers, that can be interpreted by software agents. In contrast, procedural code might only be annotated with in-line comments that need a human reader to interpret them.

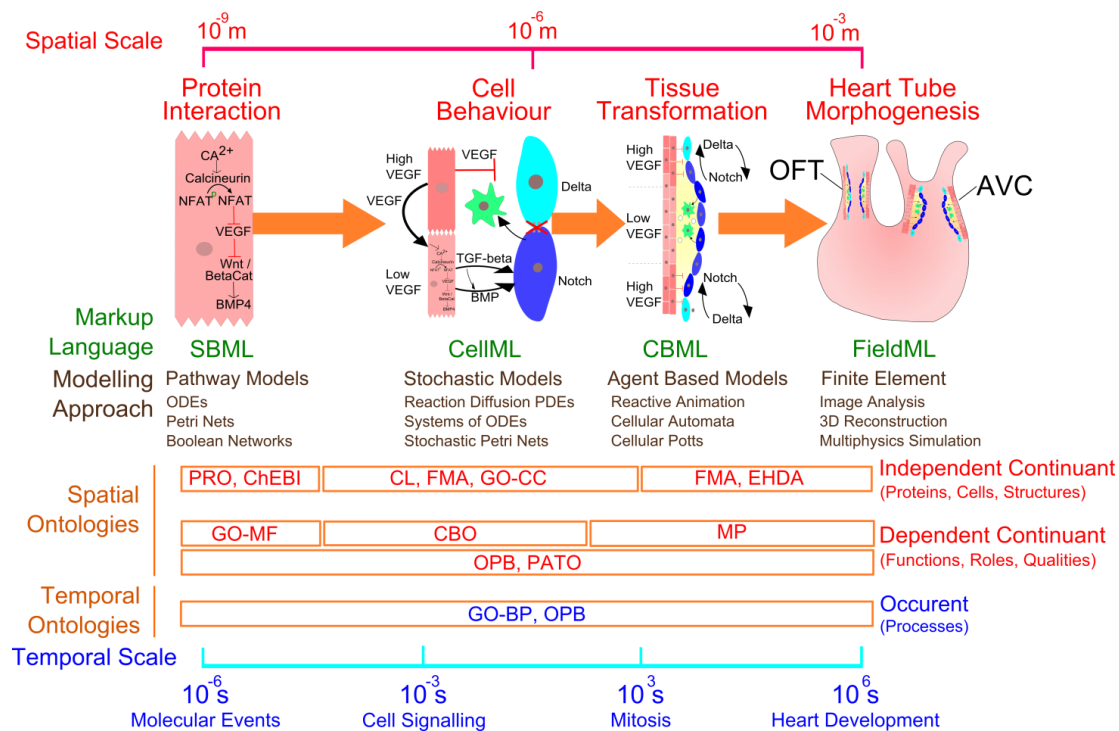


Figure 4.18 Spatial and temporal scales of heart morphogenesis modelling. The modelling framework encompasses spatial scales from 10^{-9} m (proteins) to 10^{-3} m (the primitive heart tube), and temporal scales from 10^{-6} s (molecular events) to 10^6 s (weeks of heart development). Spatial and Temporal ontologies that apply to different levels of scale are also shown. PRO: Protein Ontology, ChEBI: Chemical Entities of Biological Interest, CL: Cell Type Ontology, FMA: Foundational Model of Anatomy, GO-CC: Gene Ontology Cellular Component, EHDA: Edinburgh Human Developmental Anatomy, GO-MF: Gene Ontology Molecular Function, CBO: Cell Behaviour Ontology, MP: Mouse Phenotype Ontology, OPB: Ontology of Physics for Biology, PATO: Phenotype And Trait Ontology, GO-BP: Gene Ontology Biological Process

Along the bottom of Figure 4.18, the ontologies applicable to different levels scale are illustrated, which can be used for annotation of different model components. These ontologies are split between 'occurents', 'independent continuants' and 'dependent continuants', following BFO and OBO Foundry conventions. By making this high level distinction, the OBO community has created a clearly defined boundary between the spatial and temporal domains. As simulation models comprise both spatial and temporal domains, it is necessary to either combine terms in a post-composition approach, or make use of an application ontology for the annotation of a particular type of model.

The example process illustrated in Figure 4.18 is EMT in endocardial cushion growth. A signalling pathway within a single cell might be represented as an ODE within SBML. The interactions of cells and their chemical signalling might be represented with PDEs or

stochastic Petri nets. A simulation of a larger numbers of cells is likely to use some form of agent based modelling. Finally, at the level of the developing heart tube as an anatomical component, finite element and multiphysics simulation may be used, to understand the relationships between mechanical properties of the heart walls (affected by EMT), its function as a pump and its looping morphology.

Developmental biology is a well established field of quantitative analysis. New results emerge every day from *in vitro* and *in vivo* high-throughput analysis, and add to the growing knowledgebase of genotype-phenotype associations. Heart morphogenesis is an area of particularly intensive research, as heart defects are among the most common type of congenital disorder. This has led to a recent expansion of the gene ontology to include a much broader range of biological process terms for heart development (Khodiyar et al. 2011) and a corresponding initiative to increase the number of GO cardiovascular annotations. This represents a pre-composition approach, including creation of differentiation terms for 26 different cell types ('Endocardial cell differentiation', 'Pacemaker cell differentiation' etc.) Due to the logical structure of GO, these terms can be de-composed using cross-product extensions (Mungall et al. 2011).

Post-composition has been applied successfully for annotating phenotypic descriptions. This makes use of a particular type of ontology composition: the Entity Quality (EQ) formalism. This extends entity terms from reference ontologies by describing them as the intersection of the entity with a relationship to a quality term in PATO (Phenotype and Trait Ontology). The entities are most often from species specific anatomy or developmental anatomy ontologies, but may also be a cell type from CL; a biological process, molecular function or cellular component from GO; or a molecular level entity from PRO or ChEBI. The EQ formalism has been used for investigating the evolution of phenotypic traits (phylogenetics) (Balhoff et al. 2010) and in integrating phenotypic annotations from multiple species (Mungall et al. 2010), and in this way linking human diseases to mutant animal models (Washington et al. 2009). In contrast, the Mammalian Phenotype (MP) ontology takes a pre-composition approach, which aims to include terms sufficient for phenotypic description within a single ontology (Smith et al. 2005). This has been used successfully for the mouse and rat genome databases. The two approaches

are not mutually exclusive, as MP terms could be defined as equivalent to EQ terms, when appropriate.

The post-composition approach has also begun to be used for the annotation of biomedical simulation models. This is similar to the EQ formalism described above, but using the Ontology of Physics for Biology (OPB) rather than PATO. The OPB describes both physical properties and physical processes. This is because simulation models mainly represent the physical properties of biological entities. The SemGen tool enables modellers to annotate SBML or CellML code using OPB post-composition terms; although they must first be imported and compiled in the JSim modelling tool (Gennari et al. 2010). Once models are annotated in this way, a semantic comparison of several models can then be made through SemGen, automatically identifying entities that can be combined if models are merged. However, this approach to annotating models has only been applied to domains with well defined physical properties. It is not clear how well this would work for cell level modelling for example, where the physical properties that drive cell behaviour are not fully understood.

It is straightforward to adapt the EQ formalism for developmental phenotypes. The initial step is to select the relevant ontologies for the domain, as well as the types of sources that might be annotated. The process for the domain of heart development is illustrated in Figure 4.19.

PATO allows composite phenotype annotations such as ‘endocardial cushion with decreased concentration of SNAIL protein’, which are composed from the integration of multiple reference ontologies. OPB allows formalization of the physical properties of these composite annotations, such as the concentration of a particular protein in a particular endocardial cell, or the density of mesenchymal cells in an endocardial cushion. These terms can then be used to annotate variables in a computational model, or experimental data. PATO composites can also be mapped to disease classifications, such as OMIM.

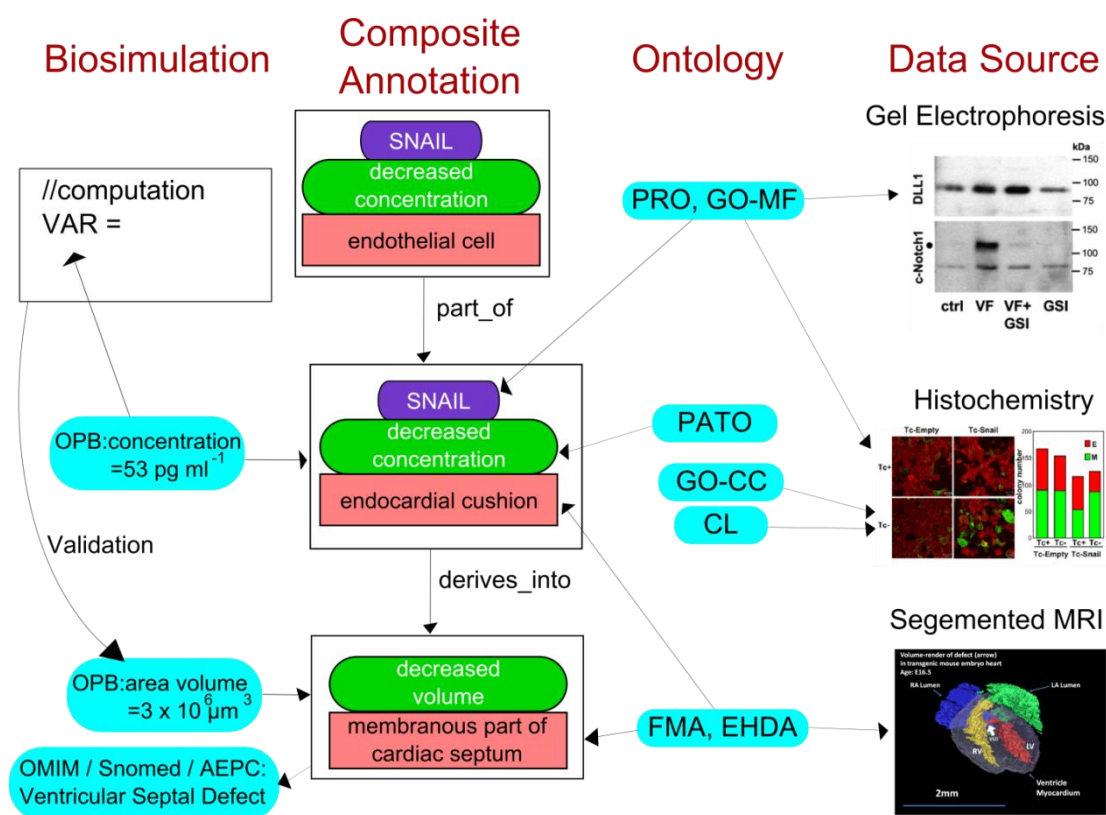


Figure 4.19 Schema for creating composite annotations from terms in multiple reference ontologies.

The topmost annotation shown in Figure 4.19 can be represented in OBO or OWL format as detailed in Table 4.4.

Table 4.4 OBO and OWL representation of composite annotation.

OBO	OWL
intersection_of:	EquivalentTo:
PATO:0001163 ! decreased	PATO:0001163
concentration	and (inheres_in some
intersection_of: inheres_in	PR:000015308)
PR:000015308 ! SNAIL	and (contained_in some
intersection_of: contained_in	CL:0002350)
CL:0002350 ! endocardial cell	

Example annotations will be represented using a simplified EQ syntax, using term labels rather than identification numbers. The annotation in Table 4.4 would be expressed as:

```
PATO:decreased concentration <inheres_in> PR: SNAIL
<contained_in> CL:endocardial cell
```

The OPB based composite annotation for the concentration parameter or measured value would be:

```
OPB:chemical concentration <property_of> OPB:portion of  
molecules <composed_of> PR:SNAI <contained_in> CL:endocardial  
cell
```

The same annotation could be used whether pointing to a model parameter, or an experimentally measured concentration. This suggests a method for leveraging the semantic relationships between very different types of information.

An EQ representation may be defined under a number of categories (Balhoff et al. 2010), with the examples below taken from the process of heart morphogenesis.

Monadic states are those that involve single entities or structures. For example, it has been previously shown that some congenital heart abnormalities are caused by an incorrect rotation of the outflow tract. This can be annotated in a general way as:

```
PATO:mislocalised_radially<inheres_in> EHDA:outflow_tract
```

Relational states are those that describe a phenotype that exists between two entities or structures. The first example in this section was relational, as describes a phenotype that exists between SNAI1 protein and endocardial cell:

```
PATO:decreased concentration <inheres_in> PR: SNAI1  
<contained_in> CL:endocardial cell
```

Quantitative states describe a measured value for a variable feature (e.g. size, area, count). For example, the volume of an endocardial cell would be annotated as:

```
OPB:volume region <inheres_in> CL:endocardial cell  
<has_magnitude> OPB:volume amount=3.2 <has_unit> UO:microliter
```

With post-composition, there is a lack of exact consistency in annotations between different annotators (Mungall et al. 2010). This is not always a major problem because, with sufficient guidelines, the differences are usually ones of specificity (e.g. did the annotator use the FMA term ‘endothelium’, ‘endothelium of endocardium’ or ‘endothelium of aortic valve’?). These annotations are still valid semantically, but where

a more coarse term is used there is a degree of information loss, to be avoided where possible. Restriction to terms of a specific domain and the use of customizable software tools for annotation improves consistency. For example, CompuCell3D and Chaste might allow you to tag cell types in a model as particular cell types from the CL ontology. An example of a customisable annotation tool is Phenote (Balhoff et al. 2010), an open source toolkit that facilitates annotation of biological data using OBO-format ontologies. However, it is still possible to have different perspectives on the same physiological phenomenon. For example, one decision might be whether the interest is in the decreased volume of the membranous septum, or the fact that the membranous septum is dysfunctional. From the perspective of exact volume quantification the actual size measurement is important, whereas in the more general disease classification the interest lies only in the fact that there is a dysfunction. There are often pre-composed terms in existing ontologies, which could also be made by post-composing terms from multiple ontologies. For example, in the MP ontology the term 'abnormal outflow tract development', could be composed as:

```
PATO:abnormal <inherits_in> GO:outflow_tract_morphogenesis
```

The degree of variability possible is a key advantage of post-composition: congenital heart diseases are a spectrum of overlapping phenotypes, and it is necessary to have flexibility in the way they are annotated. This accuracy in genotype-phenotype annotation, while arguably more complex, has greater benefits to wider biological research than mere coding and classification of defects. However, the strategies are not mutually exclusive. An intriguing possibility is to map anatomical measurements (such as those determined from the MRI of congenital heart disease specimens) to disease classifications.

The challenge of reasoning over multiple ontologies remains a considerable one. Nonetheless, it is much more feasible to achieve data integration in this way than in any existing alternative. In particular if new ontologies were constructed for each application, with no semantic links to existing reference ontologies, then ontologies would be of little use in integrating between applications.

4.7. Summary

This chapter has presented the simulations which comprise the main results of this thesis. First it was demonstrated that a 3D cellular Potts model can capture the qualitative behaviours of equilibrium, migration on the surface and invasion into collagen gel of an *in vitro* EMT. These behaviours can be observed as a result of changing contact energy parameters such that the hierarchy of contact energies is re-ordered; with different hierarchies producing the different observed behaviours. Next, a 3D simulation was used to represent qualitative behaviour in conditions of migratory and invasive cells with and without contact inhibited mitosis. It was demonstrated that contact inhibited mitosis potentially plays a role in preserving the endocardial phenotype, by preventing gaps from forming in the tissue layer. 2D simulations were then used to investigate the role of cell morphology in EMT, first with models of individual cells, and then in multicell simulations. It was found that there is a direct relationship between cell motility and cell morphology in the cellular Potts model. The multicell simulations applied this finding to investigate the relationship between Cell-Cell and Cell-Medium adhesion, and cell motility and morphology. The simulations failed to show that **Cell-Cell** adhesion can be used to increase the roundness of cells. It was also found that a moderate increase in **Cell-Medium** adhesion was necessary, in addition to loss of **cell-cell** adhesion and increase in motility, to produce migratory **cells** in the model. While a weak increase in **cell-medium** adhesion produced rounded migratory **cells**, a strong increase was sufficient to produce morphological changes. Adding elongation to this model was sufficient to produce **cells** with the same morphological and migratory characteristics of *in vitro* mesenchymal cells. The results of the multiscale simulation of Notch lateral induction were presented next, which indicated qualitative behaviour of the expression pattern in fields of tissue of different sizes. These suggest the potential for lateral induction to regulate bulging towards the centre of the endocardial cushions. However, further experimental validation would be needed to draw conclusions; especially with regards to the exact geometry of endocardial cushions and the areas of Notch expression during development. This chapter closed by presenting an approach used for composite annotation of concepts in heart development, which combine terms from multiple fields, and therefore multiple reference ontologies. The next chapter will discuss the results, the conclusions and contributions of this research, and opportunities for future investigation.

Chapter 5

5. Discussion and Conclusions

5.1. Review of the Aims and Objectives

The thesis has integrated existing knowledge of EMT in heart development at different levels of spatial scale **(A1)**. Integrating existing knowledge was achieved through extensive review of the experimental literature on EMT, heart development, and the major signalling pathways known to be involved in EMT. Through understanding EMT processes at various levels of spatial scale **(O1)**, existing knowledge was integrated and presented in a conceptual multiscale model **(O2)**. Key features identified in this conceptual model included:

- Lateral induction by Notch signalling
- Regulation of EMT by VEGF expression, and contact-inhibited proliferation
- Reduction in cell-cell adhesion and increase in cell-matrix adhesion by Notch and TGF- β signalling (via Snail transcription factors)
- Activation of an invasive phenotype by BMP2 signalling, which includes degradation of extracellular matrix proteins

The thesis increased understanding of these EMT processes by representing key features of EMT with computational simulation models **(A2)**. This was achieved by building computational cell and tissue simulation models of cells undergoing EMT **(O3)**. The simulation models included 3D models which represented conditions for 2D migration and 3D invasion of cultured *in vitro* endocardial cells. 2D models were used to represent the interplay of cell motility, morphology and adhesion as cellular changes that take place during EMT. The key EMT signalling pathway of Notch signalling lateral induction has been explored through protein level modelling **(O4)**. Simulations supported the existing hypothesis of the role of lateral induction in cardiac EMT; that it demarcates the regions of the endocardium where EMT takes place. Multiscale simulations of lateral induction showed that a regional pattern of Notch expression is generated in a tissue field through lateral induction signalling. Existing experimental results, in the form of

imaging data derived from an *in vitro* investigation of EMT were used for validation of cell and tissue level models (05). Cell outlines were extracted from images of wildtype and Notch activated *in vitro* endocardial cells. Shape metrics of aspect ratio and circularity were used to fit 2D models to the morphological characteristics of the two experimental conditions; thereby improving the validity of the model.

The thesis has also refined existing approaches for multiscale modelling of developmental processes (A3). Existing approaches to multiscale modelling were refined by applying the post-composition approach for ontological annotation of phenotypic descriptions to EMT in heart development, (06). This demonstrated the feasibility of using the post-composition approach to leverage terms from multiple existing biomedical reference ontologies for the annotation of multiscale simulation models. Multiscale modelling and simulation was investigated, using Delta-Notch lateral induction as an example, by re-implementing a published subcellular model of lateral induction as a multiscale cellular Potts model (07). This was implemented by assigning an ODE model of the subcellular reactions, encoded in SBML, to each cell in a CompuCell3D cellular Potts model.

5.2. Contributions to Original Knowledge

The thesis has made significant contributions to original knowledge in the field of computational biology. It has been demonstrated, through 3D simulations, that the representation of surface tension in the cellular Potts model is sufficient to capture the behaviours of cell migration on the surface of collagen gel, and invasion into the gel, independently, by taking account of reordering of contact energies (Section 4.2.). A reduction in cell-cell adhesion in the model was sufficient to represent the behaviour of 2D separation and migration of cells. A simultaneous reduction in cell-cell adhesion and increase in cell-matrix adhesion was sufficient to capture the behaviour of 3D invasion of cells into the collagen gel. Further experimental validation is needed to draw mechanistic conclusions about the true changes in adhesion that takes place, and their role in driving migration and invasion. Representing the system with a different modelling technique will provide validation against model artefacts (Section 5.3.1.). The 3D model does not capture some important features of EMT, such as changes in cell morphology. This was investigated in 2D simulations for the experimental conditions of Wildtype endocardial

cells (which remain rounded and in a monolayer) and Notch activated endocardial cells, which migrate on the surface of the gel and adopt mesenchymal morphology.

For the first time, a relationship has explicitly been shown between cell motility and cell shape in the cellular Potts model. Cells that are driven to be more motile (increase in T parameter) become less circular; where circularity is defined as $4\pi \cdot \text{area} / \text{perimeter}^2$. Cells that are driven to be less circular (increase in target surface area) become more motile as measured by average speed of the cell centre (Section 4.4.2.) Increasing the target surface to produce a reduction in circularity from ~ 0.7 to ~ 0.2 also caused an increase in motility from $\sim 0.2\mu\text{m}$ to $\sim 0.5\mu\text{m}$. Increasing the T parameter so as to increase motility from $\sim 0.2\mu\text{m}$ to $\sim 0.5\mu\text{m}$ also caused a reduction in circularity from ~ 0.7 to ~ 0.4 . This shows that there is an approximately inverse proportionality in the relationship between cell speed and cell circularity in the cellular Potts model. This result is interesting when considering that both changes occur simultaneously during EMT, and raises the question of what are the physical changes driving both motility and shape changes during EMT. Validation is needed to draw definitive mechanistic conclusions from these results, particularly through representing the same process using a different modelling formalism with equivalent parameters. Cell vertex model will be an appropriate technique for this (Section 5.3.1.).

This thesis is the first time, to the author's knowledge, that the features of cell morphology, adhesion and motility have been investigated together in a simulation model of EMT. As was shown in Section 4.4.3., in the cellular Potts model, shapes characteristic of mesenchymal cells can be achieved with high cell-ECM adhesion. However an elongation constraint was required to achieve aspect ratios consistent with those measured from images of *in vitro* mesenchymal cells. Furthermore, cell shape in the simulations was not found to be significantly affected by cell-cell adhesion. These findings show that the cellular Potts model requires inherent tendencies (shape driving parameters) towards rounded and mesenchymal cell morphologies that are not accounted for by cell-cell or cell-ECM adhesion alone. These findings are consistent with the experimental observation of intermediate phenotype cells that are migratory but remain rounded (Fischer et al. 2007). Such cells have clearly lost their adhesion to other

epithelial cells, but maintain a rounded epithelial morphology; as observed experimentally with deletion of the Notch target genes Hey1 or Hey2 (Fischer et al. 2007). A simulation that corresponds to this phenotype emerged from loss of epithelial adhesion, an increase in motility ('Temperature') and a moderate increase in cell-matrix adhesion. Validation involving a wider range of experimental parameters and measurements (Section 5.3.1.) will be needed to draw conclusions on the mechanisms sufficiently represented by the integrated cellular Potts models of cell adhesion, motility and morphology.

Methods for multiscale modelling have been investigated by re-implementing a published Delta-Notch lateral induction model as a multiscale cellular Potts model (Section 4.5.). While not entirely novel (Andasari et al. 2012), there are few examples of their use, and the multiscale model of Notch lateral induction provides an interesting case in which emergent expression patterns occur as a result of feedback between the models at two levels of scale. Cell level properties, such as adjacent surfaces between cells, were used to dynamically feedback with a subcellular model assigned to each cell, as described in Section 3.4.6. The multiscale model of Notch mediated lateral induction demonstrated some interesting generic properties of this signalling mechanism. Specifically, that there is a high correlation in cells at a given time between the expression of Delta ligand and the level of activated Notch (Pearson correlation coefficient $r = 0.953$). Additionally, that the expression of both Notch and Delta are higher towards the centre of a tissue in which lateral induction is operating. Cells at the outermost layer of a tissue have a reduced expression of Notch and Delta due to receiving Notch signal only from surfaces adjacent to surrounding cells. However, they rapidly adjust toward a maximal level of Notch and Delta expression as they move into the tissue, and become surrounded on all sides; as the protein signalling operates much more rapidly than cellular movement. This is captured in the model by integrating the subcellular models several times in each timestep, in which the cells can move. While these results are largely qualitative, they have implications for the mechanism of EMT in a growing endocardial cushion. The model shows that EMT occurs at a greater rate towards the centre of the endocardial cushion, due to the level of Notch activation being higher towards the centre of the region. Additionally, that as the cushion grows Notch signalling will induce the rate of

EMT to accelerate in the cushion, as a smaller proportion of the cells are at the tissue boundary. These implications need experimental validation by comparing the findings from the simulation with those from an appropriate *in vitro* study of the interaction between protein signalling and cell behaviour; and the threshold of Notch activation at which cells undergo EMT.

The thesis also made a significant contribution to methodological approaches for the annotation of multiscale biomedical models with terms from multiple biomedical reference ontologies. It has been demonstrated that a postcomposition annotation approach is applicable to multiscale processes, with heart development as an example. Using this approach, terms from existing reference ontologies for different domains and types of objects can be used to make it explicit what different objects in a model represent (e.g. cell types from CL ontology or protein species from PRO). Furthermore, lightweight ontologies have been used to provide descriptive terms that are a composition of terms from multiple ontologies, which has been used for semantic integration across different levels of spatial scale. For example, PATO is used for phenotypic annotations that may be spatial descriptions of anatomical parts, or qualitative observations of protein concentrations in particular cell types. OPB is used for physical descriptions, such as the volume parameter of a simulated cell, the size of an anatomical part, or a measured concentration level of a protein in a particular cell type. These approaches have the potential to leverage the semantics embodied in reference biomedical ontologies. The model annotation approach developed in this thesis has clearly shown (Section 4.6.) that it is now possible for multiple sets of data to be queried for the phenotypic effects associated with an increased concentration of a particular protein in a particular cell (narrower cell types or related protein species included).

5.3. Future Research Opportunities

Currently the models presented in this thesis are separated for the purpose of representing and understanding individual features of EMT. The models will be extended and integrated to provide a more holistic model of EMT. Future extensions to the models must include an integrated subcellular pathway for Notch, TGF- β and BMP signalling, and the improvement of the implementation of plastic coupling, by making it a direct function of contact area between cells over time. There is also the potential to explicitly represent

the independent effects of different molecular players on cell adhesion, cell motility and cell morphology. For a specific example, consider the secretion of TGF- β 2 protein by the myocardium underlying the endocardial cushions. The secreted TGF- β 2 protein both activates the Notch ligand Jagged1 and independently activates Snail1 and Snail2 (also termed Slug) which both act to represses VE-Cadherin (a major protein of endocardial cohesion). In contrast, only Snail2 is a direct target of Notch signalling. Simulations will be able to demonstrate the implications of these fine details in terms of genetic mutations that affect particular pathways; specifically how EMT is disrupted, and which signalling pathways are implicated in the myocardium or endocardium.

The hypothesis offered in Section 4.3. that VEGF regulates EMT via endocardial proliferation must be tested *in vitro*, by using markers for mitosis and altering the concentration of VEGF in endocardial explants. This will provide further refinement for the model, which will improve the model with a more realistic representation of endocardial proliferation.

The multiscale model of lateral induction must be improved through experimental validation involving imaging studies. Modelling the interactions of regulatory processes requires a greater understanding of the threshold of Notch activation at which cells undergo EMT, and greater geometrical accuracy in the model, regarding both cell size, cushion size and the areas of Notch expression. Including a realistic rate of cell proliferation, and Notch threshold for EMT would make it possible to study the relationships between growth rate, EMT and lateral induction. Both geometrical information on cushion architecture, and the threshold of Notch activation at which cells undergo EMT, could be derived from suitable imaging studies. A close approximation to the geometry of the endocardial cushions is important in this case. The larger the cushion region, relative to the size of the endocardial cells, the less pronounced will be the bias towards EMT at the centre of the tissue field. This is because it is the effect of a reduced number of contacts with Notch expressing cells at the periphery of a tissue in which Notch signalling lateral induction operates. This effect can only occur within the outermost layers of cells. As cellular movement is slow compared to protein signalling,

cells that return from the edge toward the centre rapidly adjust to a maximal level of Notch and Delta expression.

Further work must be done on the methods for annotating multiscale models with a postcomposition approach. EMT is a complex, multiscale process, which can only be understood through the integration of knowledge from multiple biomedical domains. Making this tractable requires abstraction, particularly through computational modelling, as has begun to be addressed by this thesis. The potential for sharing *in silico* research has been improved by common modelling standards, online databases of published models, and widespread adoption of biomedical ontologies. However, both ontologies and modelling standards tend to focus on particular levels of spatial and temporal scale, presenting a challenge for unified semantic representation of multiscale developmental processes. The methods for multiscale annotation of models described in Section 4.6. of this thesis begin to address this. As multiscale models develop in number and complexity, methods for composite annotation must also be applied and improved to meet the challenges of multiscale representation.

5.3.1. Verification and Validation

In-depth verification and validation are essential requirements to moving the modelling and simulation work in this thesis forward. Cellular Potts models are phenomenological (empirical or data-based) in that they can agree with observed biological behaviour in a statistical sense. This contrasts with mechanistic models of cell behaviour, which are based on rules that have been abstracted from an underlying biological process. Mechanistic and phenomenological models are useful for different purposes. Mechanistic models can be used to provide a satisfactory explanation of a proposed mechanism. Phenomenological models that have been sufficiently validated can be used to summarise or visualise data, to make predictions, or as an aid to designing experiments (Voss-Böhme 2012).

The 2D and 3D simulation models presented in this work have been empirically matched to the (limited) available experimental data on cell morphology, speed and migration behaviour in different experimental conditions of EMT. This is necessary for the models to be valid, but it is not sufficient. The mechanistic conclusions that may be drawn from

the models are thus of limited reliability without further validation. Without in-depth validation, it is not possible to distinguish whether agreement or disagreement with biological data is the result of the appropriateness of the model, or the validity of the hypothesised underlying biological mechanism. Future work appropriately validating the models must involve empirically matching the properties of the models under a wider range of parameter conditions. This requires further experimental work. In particular, measurement of the aggregate levels of cadherin expression under different experimental conditions, as well as the separation forces required between pairs of cells, and cells and matrix; and the average speed of individual cells. These measurements will distinguish between the mechanisms of cell-cell adhesion and cell-matrix adhesion in migration and invasion, and provide validation of the appropriateness of the cellular Potts models for representing controlled *in vitro* EMT. A wider range of experimental parameter conditions will be achieved by using cells with different expression levels of cell-cell and cell-matrix adhesion molecules. Fitting the models to the wider range of parameter conditions will test their validity empirically.

A further important method for validation will be constructing the same models using different modelling formalisms. As reviewed in Section 2.3., there are a few other techniques that are used for modelling the emergent behaviour of interacting cells. Each has different features and advantages. For example, coarse grained techniques such cell centre modelling are mechanistically better understood, and thus open to more formal analysis. However this technique is not able to represent features at the sub-cellular level (e.g. cell shape) and therefore not appropriate for the 2D models presented in this thesis; in which cell shape is a critical feature. The 2D models will be validated by constructing equivalent cell vertex models, while the 3D models will be validated by constructing equivalent cell centre models. The use of different formalisms for representing the same process will elucidate between behaviours that are truly captured by the model, and those that are an artefact of the formalism being used. For example, constructing cell vertex models in which cells are driven to be more or less circular in shape, will validate whether the greater motility of less circular cells is a true physical property, or an artefact that exists in the cellular Potts model. A high resolution cell vertex model will be required for, to avoid shape artefacts caused by the coarseness of the

vertices. Cell perimeter, and thus circularity, will be derived from the total length of the surface vertices of a cell. Cell-cell adhesion will be represented in an equivalent way to the cellular Potts model (as an energy parameter between adjacent vertices, proportional to the contact area between cells). This will provide validation of the tentative mechanistic interpretations of the roles of cell-cell and cell-matrix adhesion in driving EMT that have been offered in this thesis.

5.3.2. Potential Model Applications

EMTs are diverse, context dependent, and regulated by different protein signalling pathways in different contexts. Thus different roles may be played by cell morphology, motility and adhesion in each case. The computational models presented in this thesis provide the flexibility to simulate cells with different combinations of physical parameters. This allows investigation of the conditions under which cells remain epithelial, undergo EMT or exhibit an intermediate phenotype. The cell morphology and migration metrics described in Sections 3.4.4. and 3.4.5. provide a means to match the model to a particular example of EMT. In combination with *in vitro* analyses that investigate the effect of individual proteins, this has the potential for quantifying the cell-level effects of gene expression changes.

While the simulation models developed in this thesis are aimed at representation for the sake of understanding, with appropriate validation it would be possible to extend them into a number of applied directions. It is also worthwhile discussing these as an indication of some of the soon-to-be-realised applications of multiscale biomedical modelling.

In Silico Screening for Toxicity

Drug development is a long and expensive process. The most recent formal estimate for the average cost of development put the cost in 2003 at \$802 million (DiMasi et al. 2003). Many candidate drugs fail in the late stages of development during preclinical and clinical trials. Identifying the candidate drugs that have a higher risk of toxicity earlier on in the development cycle has the potential to create savings, and to focus resources on compounds with a higher chance of success (Figure 5.1).

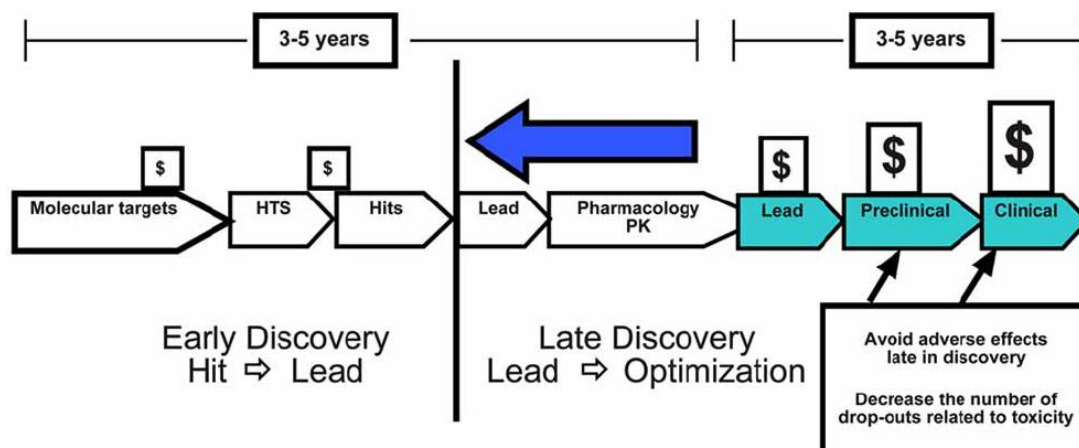


Figure 5.1 Major steps in the drug discovery process. A typical path for drug discovery is presented. Compounds in a chemical library are screened to identify those molecules that interact with the intended target. Molecules that are positive in this assay “Hits” begin the process of lead identification (Hit-to-Lead) and Lead optimization. *In vitro* toxicity screening as well as screens designed to identify ADME, genotoxicity, and cardiac toxicity should be done early in this process in order to identify high risk molecules early (McKim 2010).

This has led to increasing interest in the use of multiscale modelling for *in silico* toxicology studies, to augment the existing *in vitro* and *in vivo* protocols. This notably includes the United States Environmental Protection Agency’s Virtual Embryo project (Shah & Wambaugh 2010). Currently this involves CompuCell3D models of vascular development, limb development and eye development. With improved representation of the pathways involved, a model of EMT could be applied to *in silico* toxicology. EMT would be a good candidate due to the ubiquity of EMT throughout embryonic development, the central role of EMT in congenital heart malformations, and the fact that there have been a lot of experimental studies on this system. Another example is Germany’s Virtual Liver network, which has developed a network model for a single hepatocyte cell, and is now developing multiscale agent based models of liver lobules for *in silico* toxicology (Abbott 2010). The multiscale aspect of the models is crucial to their application in this example. We may know which proteins in which cell types a drug interacts with, but we can’t predict the effects without a model of interacting pathways and interacting cells.

Predicting the Success of Cancer Treatments

The processes that cancer cells undergo during the stages of invasion are phenotypically similar to developmental EMT. Current cancer therapies include killing cancer cells, interfering with their ability to follow chemotactic gradients, and inhibiting the ability of tumors to create vascular networks that allow them to collect the nutrients they require for growth and metastasis. There is potential for developing cancer therapies that target pathways controlling the adhesiveness of cancer cells (van Nimwegen & van de Water 2007). Such approaches might aim at strengthening cell-cell adhesion while inhibiting cell-ECM adhesion, in order to prevent metastasis.

Tissue Engineering

As EMT is crucial to developing complex tissues, understanding how cellular transitions are regulated in different tissue types, in different contexts, has potential applications in tissue engineering and regenerative medicine. Modelling and simulation could be used to suggest cell seeding strategies, environmental conditions, or the potential success of different scaffolding structures. Heart valve tissue engineering is an active area of research currently, due to the limitations of mechanical and bioprosthetic replacement valves (Weber et al. 2011; Sewell-Loftin et al. 2011; Schmidt et al. 2010). A potential future application for simulation models of EMT is to inform strategies for heart valve tissue engineering which use endocardial EMT under controlled conditions.

References

- Abbott, A., 2010. Germans cook up liver project. *Nature*, 468(7326), p.879. Available at: <http://www.nature.com/news/2010/101215/full/468879a.html> [Accessed August 17, 2012].
- Adra, S., Sun, T., MacNeil, S., Holcombe, M. & Smallwood, R., 2010. Development of a three dimensional multiscale computational model of the human epidermis. P. R. Lowenstein, ed. *PLoS ONE*, 5(1), p.e8511. Available at: <http://dx.plos.org/10.1371/journal.pone.0008511> [Accessed July 22, 2010].
- Agrawal, S., Archer, C. & Schaffer, D. V., 2009. Computational models of the Notch network elucidate mechanisms of context-dependent signaling. *PLoS Computational Biology*, 5(5), p.e1000390. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2680760&tool=pmcentrez&rendertype=abstract> [Accessed August 12, 2010].
- Andasari, V., Roper, R.T., Swat, M.H. & Chaplain, M. A. J., 2012. Integrating intracellular dynamics using CompuCell3D and Bionetsolver: applications to multiscale modelling of cancer cell growth and invasion. S. S. Dadras, ed. *PLoS ONE*, 7(3), p.e33726. Available at: <http://dx.plos.org/10.1371/journal.pone.0033726> [Accessed March 26, 2012].
- Armstrong, E.J. & Bischoff, J., 2004. Heart valve development: endothelial cell signaling and differentiation. *Circulation research*, 95(5), pp.459–70. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15345668>.
- Artavanis-Tsakonas, S., Matsuno, K. & Fortini, M.E., 1995. Notch signaling. *Science (New York, N.Y.)*, 268(1994), pp.225–232. Available at: <http://ukpmc.ac.uk/abstract/MED/7716513> [Accessed October 2, 2012].
- Bajolle, F., Zaffran, S. & Bonnet, D., 2009. Genetics and embryological mechanisms of congenital heart diseases. *Archives of cardiovascular diseases*, 102(1), pp.59–63. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19233110>.
- Baldock, R., Bisbal, J., Bono, B. De, Davidson, D., Hunter, P.J., Sukno, F. & Wimalaratne, S., 2010. Researching interoperability using core reference datasets and ontologies for the virtual physiological human, Available at: <http://vph-ricordo.eu>.
- Balhoff, J.P., Dahdul, W.M., Kothari, C.R., Lapp, H., Lundberg, J.G., Mabee, P., Midford, P.E., Westerfield, M. & Vision, T.J., 2010. Phenex: ontological annotation of phenotypic diversity A. F. Y. Poon, ed. *PLoS ONE*, 5(5), p.e10500. Available at: <http://dx.plos.org/10.1371/journal.pone.0010500>.

- Barrallo-Gimeno, A. & Nieto, M.A., 2005. The Snail genes as inducers of cell movement and survival: implications in development and cancer. *Development* (Cambridge, England), 132(14), pp.3151–61. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15983400> [Accessed June 15, 2011].
- Bentley, K., Gerhardt, H. & Bates, P.A., 2008. Agent-based simulation of notch-mediated tip cell selection in angiogenic sprout initialisation. *Journal of Theoretical Biology*, 250(1), pp.25–36. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18028963>.
- Bernabeu, M.O., Bordas, R., Pathmanathan, P., Pitt-Francis, J., Cooper, J., Garny, A., Gavaghan, D.J., Rodriguez, B., Southern, J. A. & Whiteley, J.P., 2009. CHASTE: incorporating a novel multi-scale spatial and temporal algorithm into a large-scale open source library. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 367(1895), pp.1907–30. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19380318> [Accessed July 15, 2010].
- Bersini, H., 2012. UML for ABM. *Journal of Artificial Societies and Social Simulation*, 15, pp.1–15. Available at: <http://jasss.soc.surrey.ac.uk/15/1/9.html> [Accessed September 6, 2012].
- Blokzijl, A., Dahlqvist, C., Reissmann, E., Falk, A., Moliner, A., Lendahl, U. & Ibáñez, C.F., 2003. Cross-talk between the Notch and TGF-beta signaling pathways mediated by interaction of the Notch intracellular domain with Smad3. *The Journal of Cell Biology*, 163(4), pp.723–8. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2173673&tool=pmcentrez&rendertype=abstract> [Accessed June 13, 2011].
- Bolós, V., Grego-Bessa, J. & De la Pompa, J.L., 2007. Notch signaling in development and cancer. *Endocrine reviews*, 28(3), pp.339–63. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17409286>.
- Borggreve, T. & Oswald, F., 2009. The Notch signaling pathway: transcriptional regulation at Notch target genes. *Cellular and Molecular Life Sciences: CMLS*, 66(10), pp.1631–46. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19165418> [Accessed July 31, 2011].
- Bradley, C., Bowery, A., Britten, R., Budelmann, V., Camara, O., Christie, R., Cookson, A., Frangi, A.F., Gamage, T.B., Heidlauf, T., Krittian, S., Ladd, D., Little, C., Mithraratne, K., Nash, M., Nickerson, D., Nielsen, P., Nordbø, O., Omholt, S., Pashaei, A., Paterson, D., Rajagopal, V., Reeve, A., Röhrle, O., Safaei, S., Sebastián, R., Steghöfer, M., Wu, T., Yu, T., Zhang, H. & Hunter, P., 2011. OpenCMISS: a multi-physics & multi-scale computational infrastructure for the VPH/Physiome project. *Progress in Biophysics and Molecular Biology*, 107(1), pp.32–47. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21762717> [Accessed September 7, 2012].

- Bray, S., 1998. Notch signalling in *Drosophila*: three ways to use a pathway. *Seminars in Cell & Developmental Biology*, 9(6), pp.591–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10075488>.
- Butcher, J.T., Mahler, G.J. & Hockaday, L. A., 2011. Aortic valve disease and treatment: The need for naturally engineered solutions. *Advanced Drug Delivery Reviews*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21281685> [Accessed March 10, 2011].
- Cau, E. & Blader, P., 2009. Notch activity in the nervous system: to switch or not switch? *Neural development*, 4(36). Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2761386&tool=pmcentrez&rendertype=abstract> [Accessed June 24, 2010].
- Chang, C.-P., Neilson, J.R., Bayle, J.H., Gestwicki, J.E., Kuo, A., Stankunas, K., Graef, I. A. & Crabtree, G.R., 2004. A field of myocardial-endocardial NFAT signaling underlies heart valve morphogenesis. *Cell*, 118(5), pp.649–63. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15339668>.
- Chaplain, M.A.J. & Lolas, G., 2006. Mathematical modelling of cancer invasion of tissue: dynamic heterogeneity. *Net. Hetero. Med*, 1(3), pp.399–439. Available at: http://www.maths.dundee.ac.uk/mbg/public/2006_Chaplain_Lolas_NHM.pdf [Accessed October 2, 2012].
- Christie, G.R., Nielsen, P.M.F., Blackett, S. A., Bradley, C.P. & Hunter, P.J., 2009. FieldML: concepts and implementation. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 367(1895), pp.1869–84. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19380316>.
- Christoffels, V.M., Smits, G.J., Kispert, A. & Moorman, A.F.M., 2010. Development of the pacemaker tissues of the heart. *Circulation Research*, 106(2), pp.240–54. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20133910>.
- Chu, Y.-S., Thomas, W. A., Eder, O., Pincet, F., Perez, E., Thiery, J.P. & Dufour, S., 2004. Force measurements in E-cadherin-mediated cell doublets reveal rapid adhesion strengthened by actin cytoskeleton remodeling through Rac and Cdc42. *The Journal of Cell Biology*, 167(6), pp.1183–94. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2172605&tool=pmcentrez&rendertype=abstract> [Accessed March 1, 2012].
- Cickovski, T.M., Aras, K., Alber, M.S., Izaguirre, J.A., Swat, M.H., Glazier, J.A., Merks, R.M.H., Glimm, T., Hentschel, H.G.E. & Newman, S.A., 2007. From genes to organisms via the cell. *Computing in Science and Engineering*, 9(4), pp.50–60.
- Cickovski, T.M., Huang, C., Chaturvedi, R., Glimm, T., Hentschel, H., George E, Alber, M.S., Glazier, J.A., Newman, S.A. & Izaguirre, J.A., 2005. A framework for three-

- dimensional simulation of morphogenesis. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2(4), pp.273–88. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17044166>.
- Clarke, D., Betterton, M. & Liu, X., 2006. Systems theory of Smad signalling. *IEE Proc.-Syst. Biol.*, 153(6), pp.412–24. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4027957 [Accessed October 2, 2012].
- Cohen, M., Baum, B. & Miodownik, M., 2010. The importance of structured noise in the generation of self-organizing tissue patterns through contact-mediated cell-cell signalling. *Journal of The Royal Society Interface*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21084342> [Accessed November 23, 2010].
- Collier, J.R., Monk, N.A., Maini, P.K. & Lewis, J.H., 1996. Pattern formation by lateral inhibition with feedback: a mathematical model of delta-notch intercellular signalling. *Journal of theoretical biology*, 183(4), pp.429–46. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9015458>.
- Cook, D.L., Mejino, J.L.V., Neal, M.L. & Gennari, J.H., 2009. Composite annotations: requirements for mapping multiscale data and models to biomedical ontologies. In *31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. pp. 2791–4. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2797541&tool=pmcentrez&rendertype=abstract>.
- Crabtree, G.R. & Olson, E.N., 2002. NFAT signaling: choreographing the social lives of cells. *Cell*, 109 Suppl, pp.S67–79. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11983154>.
- Crampin, E.J., 2004. Computational physiology and the physiome project. *Experimental Physiology*, 89(1), pp.1–26. Available at: <http://ep.physoc.org/cgi/doi/10.1113/expphysiol.2003.026740>.
- Daudet, N. & Lewis, J., 2005. Two contrasting roles for Notch activity in chick inner ear development: specification of prosensory patches and lateral inhibition of hair-cell differentiation. *Development (Cambridge, England)*, 132(3), pp.541–51. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15634704> [Accessed October 8, 2010].
- Davidson, L.A., Joshi, S.D., Kim, H.Y., Zhang, L. & Zhou, J., 2010. Emergent morphogenesis: elastic mechanics of a self-deforming tissue. *Journal of Biomechanics*, 43, pp.63–70.
- DiMasi, J.A., Hansen, R.W. & Grabowski, H.G., 2003. The price of innovation: new estimates of drug development costs. *Journal of Health Economics*, 22(2), pp.151–

85. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12606142> [Accessed July 12, 2012].
- Donnelly, L.F. & Higgins, C.B., 1996. MR Imaging of Conotruncal Abnormalities. *AJR*, 166, pp.925–928.
- Fischer, A., Steidl, C., Wagner, T.U., Lang, E., Jakob, P.M., Friedl, P., Knobloch, K.-P. & Gessler, M., 2007. Combined loss of Hey1 and HeyL causes congenital heart defects because of impaired epithelial to mesenchymal transition. *Circulation Research*, 100(6), pp.856–63. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17303760> [Accessed March 14, 2012].
- Fitsialos, G., Chassot, A.-A., Turchi, L., Dayem, M.A., LeBrigand, K., Moreilhon, C., Meneguzzi, G., Buscà, R., Mari, B., Barbry, P. & Ponzio, G., 2007. Transcriptional signature of epidermal keratinocytes subjected to *in vitro* scratch wounding reveals selective roles for ERK1/2, p38, and phosphatidylinositol 3-kinase signaling pathways. *The Journal of Biological Chemistry*, 282(20), pp.15090–102. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17363378> [Accessed July 14, 2012].
- Foty, R.A. & Steinberg, M.S., 2005. The differential adhesion hypothesis: a direct evaluation. *Developmental Biology*, 278(1), pp.255–63. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15649477>.
- Fu, Y., Chang, A., Chang, L., Niessen, K., Eapen, S., Setiadi, A. & Karsan, A., 2009. Differential regulation of transforming growth factor beta signaling pathways by Notch in human endothelial cells. *The Journal of Biological Chemistry*, 284(29), pp.19452–62. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2740571&tool=pmc.ncbi.res&rendertype=abstract> [Accessed August 3, 2011].
- Galdzicki, M., Cook, D.L., Zaitlan, B.L. & Roper, R.T., 2009. Building the Cell Behavior Ontology (CBO): an illustrated guide. In *Cell Behaviour Workshop 2, Biocomplexity X*.
- Gallant, N.D., Michael, K.E. & García, J., 2005. Cell adhesion strengthening: contributions of adhesive area, integrin binding, and focal adhesion assembly. *Molecular Biology of the Cell*, 16(September), pp.4329–40.
- Galle, J., Loeffler, M. & Drasdo, D., 2005. Modeling the effect of deregulated proliferation and apoptosis on the growth dynamics of epithelial cell populations *in vitro*. *Biophysical Journal*, 88(1), pp.62–75. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15475585>.
- Garny, A., Cooper, J. & Hunter, P.J., 2010. Toward a VPH/Physiome ToolKit. *WIREs Syst Biol Med*, 2, pp.134–147. Available at: <http://doi.wiley.com/10.1002/wsbm.63>.

- Garny, A., Nickerson, D.P., Cooper, J., Weber dos Santos, R., Miller, A.K., McKeever, S., Nielsen, P.M.F. & Hunter, P.J., 2008. CellML and associated tools and techniques. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 366(1878), pp.3017–43. Available at: <http://rsta.royalsocietypublishing.org/cgi/content/abstract/366/1878/3017> [Accessed September 5, 2010].
- Gennari, J.H., Neal, M.L., Galdzicki, M. & Cook, D.L., 2010. Multiple ontologies in action: Composite annotations for biosimulation models. *Journal of Biomedical Informatics*, 44(1), pp.146–154. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2989341&tool=pmcentrez&rendertype=abstract> [Accessed February 23, 2011].
- Gilles, C., Polette, M., Zahm, J.M., Tournier, J.M., Volders, L., Foidart, J.M. & Birembaut, P., 1999. Vimentin contributes to human mammary epithelial cell migration. *Journal of Cell Science*, 112 (Pt 2), pp.4615–25. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10574710>.
- Giudicelli, F., Ozbudak, E.M., Wright, G.J. & Lewis, J., 2007. Setting the tempo in development: an investigation of the zebrafish somite clock mechanism. *PLoS Biology*, 5(6), p.e150. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1877819&tool=pmcentrez&rendertype=abstract> [Accessed July 15, 2012].
- Gleeson, P., Crook, S., Cannon, R.C., Hines, M.L., Billings, G.O., Farinella, M., Morse, T.M., Davison, A.P., Ray, S., Bhalla, U.S., Barnes, S.R., Dimitrova, Y.D. & Silver, R.A., 2010. NeuroML: a language for describing data driven models of neurons and networks with a high degree of biological detail. *PLoS Computational Biology*, 6(6), p.e1000815. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2887454&tool=pmcentrez&rendertype=abstract> [Accessed July 13, 2012].
- Goldbeter, A. & Pourquié, O., 2008. Modeling the segmentation clock as a network of coupled oscillations in the Notch, Wnt and FGF signaling pathways. *Journal of Theoretical Biology*, 252(3), pp.574–85. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18308339>.
- Gray, B.L., Lieu, D.K., Collins, S.D., Smith, R.L. & Barakat, A.I., 2002. Microchannel platform for the study of endothelial cell shape and function. *Computer Engineering*, pp.9–16.
- Heitzler, P., 2010. Biodiversity and noncanonical Notch signaling. *Current Topics in Developmental Biology*, 92, pp.457–81. Available at: [http://dx.doi.org/10.1016/S0070-2153\(10\)92014-0](http://dx.doi.org/10.1016/S0070-2153(10)92014-0) [Accessed September 13, 2012].

- Hester, S.D., Belmonte, J.M., Gens, J.S., Clendenon, S.G. & Glazier, J.A., 2011. A multi-cell, multi-scale model of vertebrate segmentation and somite formation E. J. Crampin, ed. PLoS Computational Biology, 7(10), p.e1002155. Available at: <http://dx.plos.org/10.1371/journal.pcbi.1002155> [Accessed October 11, 2011].
- Heuberger, J. & Birchmeier, W., 2010. Interplay of cadherin-mediated cell adhesion and canonical Wnt signaling. Cold Spring Harbor Perspectives in Biology, 2(2), p.a002915. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2828280&tool=pmcentrez&rendertype=abstract> [Accessed July 18, 2011].
- High, F.A. & Epstein, J.A., 2008. The multifaceted role of Notch in cardiac development and disease. Nature Reviews Genetics, 9(1), pp.49–61. Available at: <http://dx.doi.org/10.1038/nrg2279>.
- Hill, M., 2011. UNSW Embryology. Available at: <http://embryology.med.unsw.edu.au/embryo.htm> [Accessed May 29, 2012].
- Hoffman, J.I.E. & Kaplan, S., 2002. The incidence of congenital heart disease. Journal of the American College of Cardiology, 39(12), pp.1890–900. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12084585>.
- Hoglund, V.J. & Majesky, M.W., 2012. Patterning the artery wall by lateral induction of Notch signaling. Circulation, 125(2), pp.212–5. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22147908> [Accessed April 13, 2012].
- Holcombe, M., Adra, S., Bicak, M., Chin, S., Coakley, S., Graham, A.I., Green, J., Greenough, C., Jackson, D., Kiran, M., MacNeil, S., Maleki-Dizaji, A., McMinn, P., Pogson, M., Poole, R., Qwarnstrom, E., Ratnieks, F., Rolfe, M.D., Smallwood, R., Sun, T. & Worth, D. 2012. Modelling complex biological systems using an agent-based approach. Integrative biology, 4(1), pp.53–64. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22052476> [Accessed July 13, 2012].
- Honda, H., Tanemura, M. & Nagai, T., 2004. A three-dimensional vertex dynamics cell model of space-filling polyhedra simulating cell behavior in a cell aggregate. Journal of Theoretical Biology, 226(4), pp.439–53. Available at: <http://dx.doi.org/10.1016/j.jtbi.2003.10.001> [Accessed July 23, 2012].
- Hunter, P.J., Crampin, E.J. & Nielsen, P.M.F., 2008. Bioinformatics, multiscale modeling and the IUPS Physiome Project. Briefings in bioinformatics, 9(4), pp.333–43. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18477639>.
- Hunter, P.J., Li, W., McCulloch, A. & Noble, D., 2006. Multiscale modeling: physiome project standards, tools, and databases. Computer, 39(11), pp.48–54. Available at: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4014765>.

- Hunter, P.J., Robbins, P. & Noble, D., 2002. The IUPS human Physiome Project. *European journal of physiology*, 445(1), pp.1–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12397380>.
- Hutson, M.R. & Kirby, M.L., 2007. Model systems for the study of heart development and disease. Cardiac neural crest and conotruncal malformations. *Seminars in Cell & Developmental Biology*, 18(1), pp.101–10. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17224285>.
- Ilina, O. & Friedl, P., 2009. Mechanisms of collective cell migration at a glance. *Journal of Cell Science*, 122(Pt 18), pp.3203–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19726629>.
- Juty, N., Le Novère, N. & Laibe, C., 2012. Identifiers.org and MIRIAM Registry: community resources to provide persistent identification. *Nucleic acids research*, 40(Database issue), pp.D580–6. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3245029&tool=pmcentrez&rendertype=abstract> [Accessed August 24, 2012].
- Khodiyar, V.K., Hill, D.P., Howe, D., Berardini, T.Z., Tweedie, S., Talmud, P.J., Breckenridge, R., Bhattacharya, S., Riley, P., Scambler, P. & Lovering, R.C., 2011. The representation of heart development in the gene ontology. *Developmental biology*, 354(1), pp.9–17. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21419760> [Accessed May 17, 2011].
- Kirby, M.L., 2007. *Cardiac Development*, New York: OUP.
- Kitano, H., 2002. Systems biology: a brief overview. *Science (New York, N.Y.)*, 295(5560), pp.1662–4. Available at: <http://www.sciencemag.org/content/295/5560/1662.abstract> [Accessed July 13, 2012].
- Lai, E.C., 2004. Notch signaling: Control of cell communication and cell fate. *Development*, 131, pp.965–973.
- Lambrechts, D. & Carmeliet, P., 2004. Sculpting heart valves with NFATc and VEGF. *Cell*, 118(5), pp.532–4. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15339657>.
- Lanford, P.J., Lan, Y., Jiang, R., Lindsell, C., Weinmaster, G., Gridley, T. & Kelley, M.W., 1999. Notch signalling pathway mediates hair cell development in mammalian cochlea. *Nature genetics*, 21(3), pp.289–92. Available at: <http://dx.doi.org/10.1038/6804> [Accessed August 13, 2012].
- Lewis, J., 1998. Notch signalling and the control of cell fate choices in vertebrates. *Seminars in Cell & Developmental Biology*, 9(6), pp.583–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9892564>.

- Li, C., Courtot, M., Le Novère, N. & Laipe, C., 2010. BioModels.net Web Services, a free and integrated toolkit for computational modelling software. *Briefings in bioinformatics*, 11(3), pp.270–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19939940>.
- Lloyd, C.M., Halstead, M.D.B. & Nielsen, P.F., 2004. CellML: its future, present and past. *Progress in biophysics and molecular biology*, 85(2-3), pp.433–50. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15142756> [Accessed August 22, 2012].
- Luna-Zurita, L., Prados, B., Grego-bessa, J., Luxán, G., Monte, G., Benguría, A., Adams, R.H., Pérez-pomares, J.M. & De la Pompa, J.L., 2010. Integration of a Notch-dependent mesenchymal gene program and Bmp2-driven cell invasiveness regulates murine cardiac valve formation. *The Journal of Clinical Investigation*, 120(10), pp.3493–507.
- Malek, a M. & Izumo, S., 1996. Mechanism of endothelial cell shape change and cytoskeletal remodeling in response to fluid shear stress. *Journal of Cell Science*, 109 (Pt 4, pp.713–26. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8718663>.
- Mansi, T., Voigt, I., Leonardi, B., Pennec, X., Durrleman, S., Sermesant, M., Delingette, H., Taylor, a M., Boudjemline, Y., Pongiglione, G. & Ayache, N., 2011. A statistical model for quantification and prediction of cardiac remodelling: application to tetralogy of Fallot. *IEEE transactions on medical imaging*, 30(9), pp.1605–16. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21880565>.
- Margadant, C. & Sonnenberg, A., 2010. Integrin-TGF-beta crosstalk in fibrosis, cancer and wound healing. *EMBO reports*, 11(2), pp.97–105. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20075988>.
- McDaniell, R., Warthen, D.M., Sanchez-Lara, P. A., Pai, A., Krantz, I.D., Piccoli, D.A. & Spinner, N.B., 2006. NOTCH2 mutations cause Alagille syndrome, a heterogeneous disorder of the notch signaling pathway. *American journal of human genetics*, 79(1), pp.169–73. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1474136&tool=pmcentrez&rendertype=abstract>.
- McKeen-Polizzotti, L., Henderson, K.M., Oztan, B., Bilgin, C.C., Yener, B. & Plopper, G.E., 2011. Quantitative metric profiles capture three-dimensional temporospatial architecture to discriminate cellular functional states. *BMC medical imaging*, 11(1), p.11. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3125246&tool=pmcentrez&rendertype=abstract> [Accessed March 16, 2012].

- McKim, J.M., 2010. Building a tiered approach to *in vitro* predictive toxicity screening: a focus on assays with *in vivo* relevance. *Combinatorial chemistry & high throughput screening*, 13(2), pp.188–206.
- Meineke, F.A., Potten, C.S. & Loeffler, M., 2001. Cell migration and organization in the intestinal crypt using a lattice-free model. *Cell proliferation*, 34(4), pp.253–66. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11529883>.
- Mendez, M.G., Kojima, S.-I. & Goldman, R.D., 2010. Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 24(6), pp.1838–51. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2874471&tool=pmcentrez&rendertype=abstract> [Accessed March 28, 2012].
- Merks, R. & Glazier, J.A., 2005. A cell-centered approach to developmental biology. *Physica A*, 352, pp.113–130. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0378437104016188> [Accessed July 28, 2010].
- Mungall, C.J., Bada, M., Berardini, T.Z., Deegan, J., Ireland, A., Harris, M.A., Hill, D.P. & Lomax, J., 2011. Cross-product extensions of the Gene Ontology. *Journal of Biomedical Informatics*, 44(1), pp.80–6. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2910209&tool=pmcentrez&rendertype=abstract>.
- Mungall, C.J., Gkoutos, G. V, Smith, C.L., Haendel, M.A., Lewis, S.E. & Ashburner, M., 2010. Integrating phenotype ontologies across multiple species. *Genome biology*, 11(1), p.R2. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20064205>.
- Nagai, T. & Honda, H., 2006. Wound Healing Mechanism in Epithelial Tissues Cell Adhesion to Basal Lamina. In pp. 111–116.
- Neagu, A., Mironov, V., Kosztin, I., Barz, B., Neagu, M., Moreno-Rodriguez, R.A., Markwald, R. & Forgacs, G., 2010. Computational modeling of epithelial-mesenchymal transformations. *Biosystems*, 100(1), pp.23–30. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0303264709002068>.
- Neal, M.L., 2010. Modular, semantics-based composition of biosimulation models. University of Washington.
- Neal, M.L., Gennari, J.H., Arts, T. & Cook, D.L., 2009. Advances in semantic representation for multiscale biosimulation: In *Pacific Symposium on Biocomputing* 14. pp. 304–315.

- Newman, T.J., 2005. Modeling Multicellular Systems Using Subcellular Elements. *Mathematical Biosciences and Engineering*, 2(3), pp.613–624. Available at: <http://www.aims sciences.org/journals/displayArticles.jsp?paperID=1173>.
- Niessen, K., Fu, Y., Chang, L., Hoodless, P.A., McFadden, D. & Karsan, A., 2008. Slug is a direct Notch target required for initiation of cardiac cushion cellularization. *The Journal of Cell Biology*, 182(2), pp.315–25. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18663143>.
- Van Nimwegen, M.J. & Van de Water, B., 2007. Focal adhesion kinase: a potential target in cancer therapy. *Biochemical pharmacology*, 73(5), pp.597–609. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16997283> [Accessed July 13, 2012].
- Nomura-Kitabayashi, A., Phoon, C.K.L., Kishigami, S., Rosenthal, J., Yamauchi, Y., Abe, K., Yamamura, K., Samtani, R., Lo, C.W. & Mishina, Y., 2009. Outflow tract cushions perform a critical valve-like function in the early embryonic heart requiring BMPRIA-mediated signaling in cardiac neural crest. *American journal of physiology. Heart and circulatory physiology*, 297(5), pp.H1617–28. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19717734>.
- Osborne, J M, Walter, a, Kershaw, S.K., Mirams, G R, Fletcher, a G., Pathmanathan, P, Gavaghan, D., Jensen, O.E., Maini, P K & Byrne, H.M., 2010. A hybrid approach to multi-scale modelling of cancer. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 368(1930), pp.5013–28. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20921009> [Accessed July 20, 2011].
- Osbourne, J., 2012. Multiscale modelling of biological systems: the Chaste framework. In *INCF Multiscale Modeling Program Workshop: From cellular/network models to tissue simulation*. Stockholm, Sweden. Available at: <http://www.youtube.com/watch?v=TH-mN-2v6R0>.
- Otto, C.M., 2002. Calcification of bicuspid aortic valves. *Heart (British Cardiac Society)*, 88(4), pp.321–2. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1767390&tool=pmcentrez&rendertype=abstract>.
- Owen, M.R., Sherratt, J.A. & Wearing, H.J., 2000. Lateral induction by juxtacrine signaling is a new mechanism for pattern formation. *Developmental biology*, 217(1), pp.54–61. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10625535>.
- O'Dea, R.D. & King, J.R., 2011. Multiscale analysis of pattern formation via intercellular signalling. *Mathematical biosciences*, 231(2), pp.172–85. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21385590> [Accessed August 15, 2012].
- Painter, K.J., Armstrong, N.J. & Sherratt, Jonathan A., 2010. The impact of adhesion on cellular invasion processes in cancer and development. *Journal of Theoretical*

Biology, 264(3), pp.1057–67. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/20346958> [Accessed July 18, 2012].

- Pierpont, M.E., Basson, C.T., Benson, D.W., Gelb, B.D., Giglia, T.M., Goldmuntz, E., McGee, G., Sable, C.A., Srivastava, D. & Webb, C.L., 2007. Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. *Circulation*, 115(23), pp.3015–38. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/17519398>.
- Pitt-Francis, J., Pathmanathan, P., Bernabeu, M.O., Bordas, R., Cooper, J., Fletcher, A.G., Mirams, G.R., Murray, P., Osborne, J.M. & Walter, A., 2009. Chaste: A test-driven approach to software development for biological modelling. *Computer Physics Communications*, 180(12), pp.2452–2471. Available at:
<http://linkinghub.elsevier.com/retrieve/pii/S0010465509002604> [Accessed September 27, 2010].
- Podgorski, G.J., Bansal, M. & Flann, N.S., 2007. Regular mosaic pattern development: a study of the interplay between lateral inhibition, apoptosis and differential adhesion. *Theoretical biology & medical modelling*, 4, p.43. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/17974031>.
- Popel, A.S. & Hunter, P.J., 2009. Systems biology and physiome projects. *WIREs Syst Biol Med*, 1, pp.153–158. Available at:
<http://onlinelibrary.wiley.com/doi/10.1002/wsbm.67/full> [Accessed May 13, 2012].
- Popławski, N.J., Shirinifard, A., Swat, M. & Glazier, J.A., 2008. Simulation of single-species bacterial-biofilm growth using the Glazier-Graner-Hogeweg model and the CompuCell3D modeling environment. *Mathematical biosciences and engineering : MBE*, 5(2), pp.355–88. Available at:
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2547990&tool=pmcentrez&rendertype=abstract>.
- Ramis-Conde, I., Chaplain, M.A.J., Anderson, A.R.A. & Drasdo, D., 2009. Multi-scale modelling of cancer cell intravasation: the role of cadherins in metastasis. *Physical biology*, 6(1), p.016008. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/19321920> [Accessed January 21, 2012].
- Ramis-Conde, I., Drasdo, D., Anderson, A.R.A. & Chaplain, M.A.J., 2008. Modeling the influence of the E-cadherin-beta-catenin pathway in cancer cell invasion: a multiscale approach. *Biophysical Journal*, 95(1), pp.155–65. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/18339758>.

- De Rooij, J., Kerstens, A., Danuser, G., Schwartz, M.A. & Waterman-Storer, C.M., 2005. Integrin-dependent actomyosin contraction regulates epithelial cell scattering. *The Journal of Cell Biology*, 171(1), pp.153–64. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2171213&tool=pmcentrez&rendertype=abstract> [Accessed March 1, 2012].
- Rubenstein, B.M. & Kaufman, L.J., 2008. The role of extracellular matrix in glioma invasion: a cellular Potts model approach. *Biophysical Journal*, 95(12), pp.5661–80. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2599859&tool=pmcentrez&rendertype=abstract> [Accessed June 11, 2011].
- Rupp, P., Visconti, R., Czirók, A., Cheres, D.A. & Little, C.D., 2008. Matrix Metalloproteinase 2-Integrin $\alpha\beta 3$ Binding Is Required for Mesenchymal Cell Invasive Activity but Not Epithelial Locomotion: A Computational Time-Lapse Study. *Molecular Biology of the Cell*, 19, pp.5529–5540. Available at: <http://www.molbiolcell.org/content/19/12/5529.short> [Accessed May 16, 2012].
- Rutenberg, J.B., Fischer, A., Jia, H., Gessler, M., Zhong, T.P. & Mercola, M., 2006. Developmental patterning of the cardiac atrioventricular canal by Notch and Hairy-related transcription factors. *Development (Cambridge, England)*, 133(21), pp.4381–90. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17021042>.
- Saravanamuthu, S.S., Gao, C.Y. & Zelenka, P.S., 2009. Notch signaling is required for lateral induction of Jagged1 during FGF-induced lens fiber differentiation. *Developmental biology*, 332(1), pp.166–76. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2730671&tool=pmcentrez&rendertype=abstract> [Accessed June 28, 2010].
- Savagner, P., 2010. The epithelial-mesenchymal transition (EMT) phenomenon. *Annals of Oncology*, 21(Supplement 7), pp.89–92. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20943648> [Accessed August 12, 2011].
- Savill, N., 1997. Modelling Morphogenesis: From Single Cells to Crawling Slugs. *Journal of Theoretical Biology*, 184(3), pp.229–235. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0022519396902374>.
- Savill, N.J. & Sherratt, J.A., 2003. Control of epidermal stem cell clusters by Notch-mediated lateral induction. *Developmental Biology*, 258(1), pp.141–153. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0012160603001076> [Accessed February 28, 2011].
- Schmidt, D., Dijkman, P.E., Driessen-Mol, A., Stenger, R., Mariani, C., Puolakka, A., Rissanen, M., Deichmann, T., Odermatt, B., Weber, B., Emmert, M.Y., Zund, G., Baaijens, F.P.T. & Hoerstrup, S.P., 2010. Minimally-invasive implantation of living tissue engineered heart valves: a comprehensive approach from autologous

- vascular cells to stem cells. *Journal of the American College of Cardiology*, 56(6), pp.510–20. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20670763> [Accessed July 29, 2011].
- Schmierer, B., Tournier, A.L., Bates, P.A. & Hill, C.S., 2008. Mathematical modeling identifies Smad nucleocytoplasmic shuttling as a dynamic signal-interpreting system. *Proceedings of the National Academy of Sciences of the United States of America*, 105(18), pp.6608–13. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2373357&tool=pmcentrez&rendertype=abstract>.
- Sewell-Loftin, M.K., Chun, Y.W., Khademhosseini, A. & Merryman, W.D., 2011. EMT-Inducing Biomaterials for Heart Valve Engineering: Taking Cues from Developmental Biology. *Journal of cardiovascular translational research*, pp.658–671. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21751069> [Accessed August 15, 2011].
- Shah, I. & Wambaugh, J., 2010. Virtual tissues in toxicology. *Journal of toxicology and environmental health. Part B, Critical reviews*, 13(2-4), pp.314–28. Available at: <http://dx.doi.org/10.1080/10937404.2010.483948> [Accessed October 1, 2012].
- Shirinifard, A., 2012. Vascular Patterning and Its Application in Cancer and Choroidal Neovascularization. Indiana University.
- Shirinifard, A., Glazier, J.A., Swat, Maciej, Gens, J.S., Family, F., Jiang, Y. & Grossniklaus, H.E., 2012. Adhesion Failures Determine the Pattern of Choroidal Neovascularization in the Eye: A Computer Simulation Study E. J. Crampin, ed. *PLoS Computational Biology*, 8(5), p.e1002440. Available at: <http://dx.plos.org/10.1371/journal.pcbi.1002440> [Accessed May 4, 2012].
- Sluka, J.P., Glazier, J.A. & Swat, M., 2010. Frameworks for Shareable Multiscale Modeling. Biocomplexity Institute, Indiana University, pp.1–25. Available at: http://www.imagwiki.nibib.nih.gov/mediawiki/images/2/2b/MultiScale_Modeling_NIH_Oct_21_2010.pdf.
- Smallwood, R., 2006. The epitheliome project: multiscale agent-based modeling of epithelial cells. 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro., pp.816–819. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1625043 [Accessed May 13, 2012].
- Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., Goldberg, L.J., Eilbeck, K., Ireland, A., Mungall, C.J., Leontis, N., Rocca-Serra, P., Ruttenberg, A., Sansone, S.-A., Scheuermann, R.H., Shah, N., Whetzel, P.L. & Lewis, S., 2007. The OBO Foundry: coordinated evolution of ontologies to support biomedical data

- integration. *Nature biotechnology*, 25(11), pp.1251–5. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17989687>.
- Smith, C.L., Goldsmith, C.-A.W. & Eppig, J.T., 2005. The Mammalian Phenotype Ontology as a tool for annotating, analyzing and comparing phenotypic information. *Genome biology*, 6(1), p.R7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15642099>.
- Sprinzak, D., Lakhanpal, A., Lebon, L., Garcia-Ojalvo, J. & Elowitz, M.B., 2011. Mutual inactivation of notch receptors and ligands facilitates developmental patterning. *PLoS Computational Biology*, 7(6), p.e1002069. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3111533&tool=pmcentrez&rendertype=abstract> [Accessed August 10, 2011].
- Srivastava, D & Olson, E N, 2000. A genetic blueprint for cardiac development. *Nature*, 407(6801), pp.221–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11001064>.
- Steinberg, M.S., 2007. Differential adhesion in morphogenesis: a modern view. *Current opinion in genetics & development*, 17(4), pp.281–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17624758> [Accessed July 19, 2010].
- Strougo, A., Eissing, T., Yassen, A., Willmann, S., Danhof, M. & Freijer, J., 2012. First dose in children: physiological insights into pharmacokinetic scaling approaches and their implications in paediatric drug development. *Journal of pharmacokinetics and pharmacodynamics*, 39(2), pp.195–203. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22311388> [Accessed March 20, 2012].
- Sun, T., Adra, S., Smallwood, R., Holcombe, M. & MacNeil, S., 2009. Exploring hypotheses of the actions of TGF-beta1 in epidermal wound healing using a 3D computational multiscale model of the human epidermis. P. R. Lowenstein, ed. *PLoS ONE*, 4(12), p.e8515. Available at: <http://dx.plos.org/10.1371/journal.pone.0008515> [Accessed July 22, 2010].
- Sun, Y., Lowther, W., Kato, K., Bianco, C., Kenney, N., Strizzi, L., Raafat, D., Hirota, M., Khan, N.I., Bargo, S., Jones, B., Salomon, D. & Callahan, R., 2005. Notch4 intracellular domain binding to Smad3 and inhibition of the TGF-beta signaling. *Oncogene*, 24(34), pp.5365–74. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16007227> [Accessed July 25, 2012].
- Swat, M.H., Hester, S.D., Heiland, R.W., Zaitlen, B.L. & Glazier, J.A., 2009. Multi-Cell Simulations of Development and Disease Using the CompuCell3D Simulation Environment I. V. Maly, ed. *Methods Mol Biol.*, 500, pp.361–428. Available at: <http://www.springerlink.com/index/10.1007/978-1-59745-525-1>.
- Swat, M.H., Thomas, G.L., Belmonte, J.M., Shirinifard, A., Hmeljak, D. & Glazier, J.A., 2012. Multi-scale modeling of tissues using CompuCell3D. *Methods in cell biology*,

- 110, pp.325–66. Available at: <http://dx.doi.org/10.1016/B978-0-12-388403-9.00013-8> [Accessed September 13, 2012].
- Szabó, A., Varga, K., Garay, T., Hegedus, B. & Czirók, A., 2012. Invasion from a cell aggregate—the roles of active cell motion and mechanical equilibrium. *Physical biology*, 9(1), p.016010. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22313673> [Accessed March 14, 2012].
- Thiery, J.P. & Sleeman, J.P., 2006. Complex networks orchestrate epithelial-mesenchymal transitions. *Nature Reviews Molecular Cell Biology*, 7(2), pp.131–42. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16493418>.
- Timmerman, L.A., Grego-Bessa, J., Raya, A., Bertrán, E., Pérez-Pomares, J.M., Díez, J., Aranda, S., Palomo, S., McCormick, F., Izpisua-Belmonte, J.C. & De la Pompa, J.L., 2004. Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. *Genes & development*, 18(1), pp.99–115. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=314285&tool=pmcentrez&rendertype=abstract> [Accessed August 1, 2011].
- Trichas, G., Smith, A.M., White, N., Wilkins, V., Watanabe, T., Moore, A., Joyce, B., Sugnaseelan, J., Rodriguez, T.A., Kay, D., Baker, R.E., Maini, P.K. & Srinivas, S., 2012. Multi-cellular rosettes in the mouse visceral endoderm facilitate the ordered migration of anterior visceral endoderm cells. *PLoS biology*, 10(2), p.e1001256. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3274502&tool=pmcentrez&rendertype=abstract> [Accessed May 14, 2012].
- Turner, S. & Sherratt, J., 2002. Intercellular adhesion and cancer invasion: a discrete simulation using the extended Potts model. *Journal of Theoretical Biology*, pp.85–100. Available at: <http://www.sciencedirect.com/science/article/pii/S0022519301925226> [Accessed September 29, 2012].
- Vilar, J.M.G., Jansen, R. & Sander, C., 2006. Signal Processing in the TGF- β Superfamily Ligand-Receptor Network. *PLoS Computational Biology*, 2(1), pp.36–45.
- Voss-Böhme, A., 2012. Multi-Scale Modeling in Morphogenesis: A Critical Analysis of the Cellular Potts Model. *PLoS ONE*, 7(9), p.e42852.
- Wagner, M. & Siddiqui, M.A., 2007. Signal Transduction in Early Heart Development (II): Ventricular Chamber Specification, Trabeculation, and Heart Valve Formation. *Experimental Biology and Medicine*, 232, pp.866–880.
- Walker, D.C., Georgopoulos, N.T. & Southgate, J., 2010. Anti-social cells: predicting the influence of E-cadherin loss on the growth of epithelial cell populations. *Journal of*

- Theoretical Biology, 262(3), pp.425–40. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/19852973> [Accessed July 22, 2010].
- Waltemath, D., Adams, R., Beard, D.A., Bergmann, F.T., Bhalla, U.S., Britten, R., Chelliah, V., Cooling, M.T., Cooper, J., Crampin, E.J., Garny, A., Hoops, S., Hucka, M., Hunter, P., Klipp, E., Laibe, C., Miller, A.K., Moraru, I., Nickerson, D., Nielsen, P., Nikolski, M., Sahle, S., Sauro, H. M., Schmidt, H., Snoep, J. L., Tolle, D., Wolkenhauer, O. & Le Novère, N., 2011. Minimum Information About a Simulation Experiment (MIASE). *PLoS Computational Biology*, 7(4), p.e1001122.
- Walter, A.C., 2009. A Comparison of Continuum and Cell-based Models of Colorectal Cancer. The University of Nottingham. Available at:
http://etheses.nottingham.ac.uk/763/1/thesisAll_colour.pdf.
- Washington, N.L., Haendel, M. A., Mungall, C.J., Ashburner, M., Westerfield, M. & Lewis, S.E., 2009. Linking human diseases to animal models using ontology-based phenotype annotation. *PLoS biology*, 7(11), p.e1000247. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/19956802>.
- Weber, B., Scherman, J., Emmert, M.Y., Gruenenfelder, J., Verbeek, R., Bracher, M., Black, M., Kortsmit, J., Franz, T., Schoenauer, R., Baumgartner, L., Brokopp, C., Agarkova, I., Wolint, P., Zund, G., Falk, V., Zilla, P. & Hoerstrup, S.P., 2011. Injectable living marrow stromal cell-based autologous tissue engineered heart valves: first experiences with a one-step intervention in primates. *European heart journal*, 32(22), pp.2830–40. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/21415068> [Accessed December 28, 2011].
- Weliky, M. & Oster, G., 1990. The mechanical basis of cell rearrangement. *Development*, 386, pp.373–386.
- Wolf, K., Wu, Y.I., Liu, Y., Geiger, J., Tam, E., Overall, C., Stack, M.S. & Friedl, P., 2007. Multi-step pericellular proteolysis controls the transition from individual to collective cancer cell invasion. *Nature cell biology*, 9(8), pp.893–904. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/17618273> [Accessed July 15, 2012].
- Yi, J., Chen, M., Wu, X., Yang, X., Xu, T., Zhuang, Y., Han, M. & Xu, R., 2010. Endothelial SUR-8 acts in an ERK-independent pathway during atrioventricular cushion development. *Developmental dynamics*, 239(7), pp.2005–13. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/20549726> [Accessed July 14, 2010].
- Zavadil, J., Cermak, L., Soto-Nieves, N. & Böttinger, E.P., 2004. Integration of TGF-beta/Smad and Jagged1/Notch signalling in epithelial-to-mesenchymal transition. *The EMBO journal*, 23(5), pp.1155–65.
- Zi, Z. & Klipp, E., 2007. Constraint-based modeling and kinetic analysis of the Smad dependent TGF-beta signaling pathway. *PLoS ONE*, 2(9), p.e936.

Appendix A.

Modelling and Simulation Methods in Further Detail

Readers unfamiliar with cellular Potts modelling may be left wondering how parameters such as the contact energies and Temperature are derived, and what these figures might represent in the real system. Furthermore, the common practice of using convenient integer values for parameters can leave an unfortunate impression that simulations are ad-hoc setups.

The first important point to make is that the absolute values for contact energy parameters have no direct meaning whatsoever. What is important is the hierarchy of energies, and the relative differences between them. However, simply scaling the contact energies (multiplying them by 10 say) will not produce exactly equivalent simulations. This is because, with higher contact energies, there will be a greater change in the system energy at each simulation step. So for a given Temperature value, a model with higher contact energies will have a lower tendency to change.

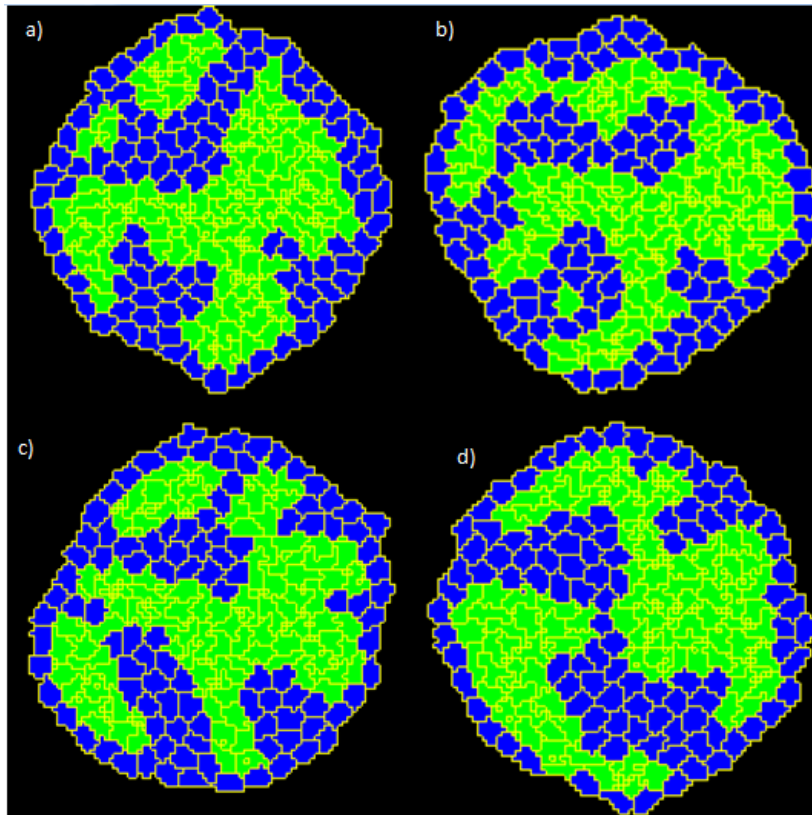
Recall from Section 3.3. that the acceptance probability function used for each pixel copy in a Potts simulation is: $\min(1, e^{\Delta H/T})$. Therefore, to achieve an equivalent set of parameters, it is necessary to scale the Temperature by the same factor. Any other constraints (e.g. the volume constraint) need to also be scaled by the same factor, in order that they have the same effect.

In other words, the four sets of parameters given in the table below are all exactly equivalent. In each case, for example, the contact energy between A cells (A-A) is eight times as high as that between B cells (B-B).

Table_Apx A.1 Four equivalent parameter sets can be produced by scaling the contact energies, the Temperature, and Lambda Volume.

Surface energy J	A-A	A-Medium	B-B	B-Medium	A-B	Medium-Medium	T	Target Volume	Lambda Volume
Set 1	16	16	2	16	11	0	10	25	2
Set 2	160	160	20	160	110	0	100	25	20
Set 3	1600	1600	200	1600	1100	0	1000	25	200
Set 4	4	4	0.5	4	2.75	0	2.5	25	0.5

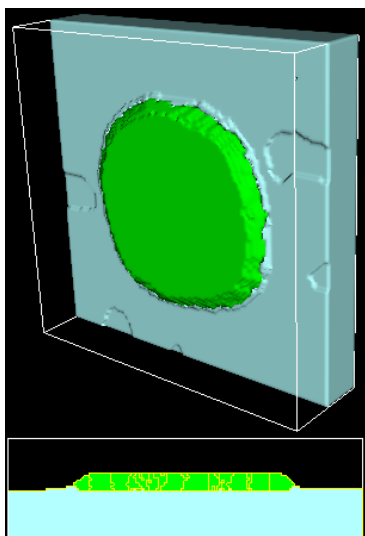
By way of illustrating that these parameter sets are exactly equivalent, the figure below shows a snapshot at 2000 MCS from simulations with each of the parameter sets 1-4. While, of course, the snapshots are not exactly the same, due to the stochastic nature of the simulations, in each case the type A (blue) cells surround the type B (green) cells. Furthermore, due to the (relatively) low contact energy (high adhesion) between type B cells, they tend to be more condensed than type A cells.



Figure_Apx A.1 Simulations with equivalent parameter sets at 2000 Monte Carlo steps: a) Set 1 b) Set 2 c) Set 3 d) Set 4. Green: cell type A, Blue: cell type B

Therefore, these input parameters do not have a direct biological meaning (apart from the Target Volume, which can be treated as the average volume of the cells). Nor do they have any meaning at all outside the context of the other parameters. However the behaviours observed in the simulations do have biological meaning. In each case, type B cells tend to envelop type A cells. This is due to the contact energy A-A being 8 times as high as that of B-B. Due to the linear relationship between the expression level of cadherins and the aggregate surface tension of cells (Foty & Steinberg 2005), it is reasonable to interpret this as: “in the model, type B cells have 8 times the level of cadherins as type A cells”.

When setting contact energy parameters, one needs to consider the likely overall order of surface tensions between objects in the model, and how these might be affected by changes in adhesion. Let’s return to the first model described in the Results chapter. This contained 3 types of object: the cells, the collagen gel and the air above the cells. As such, you may recall that we took the highest contact energy to be between cells and the air. The second highest was between the collagen gel and the air. These assumptions are made on the basis that the air does not stick to anything, and the cells are less deformable objects than the collagen gel.



Figure_Apx A.2 The simulation of an endothelial monolayer contains 3 types: Cells (green), Collagen Gel (Blue) and Air (or Medium, invisible)

Next, we assume that if cells were to lose all adhesion to each other, they would in fact have high contact energies with each other (again, due to their low deformability), but not as high as between the collagen gel and the air. We proceed from these assumptions to experiment with switching the order of contact energies, and observe the qualitative behaviour of the simulated system. Of course we can't know (without detailed measurements) that the objects in real system exhibit the same relative differences in surface tension that we use in the model. However, what we can show is that this model is sufficient to capture the different cell behaviours of migration and invasion. The model is sufficient for this purpose, by taking account of the assumptions about surface tensions, and without including any active migration in the model. The Temperature parameter, and any cell constraints (e.g. the volume constraint), are set at reasonable levels for the scale of the contact parameters used (see above table) – so that the cells maintain their volumes, and the system doesn't freeze (Temperature too low or constraints too high) or exhibit fragmentation (Temperature too high or constraints too low).

Later simulations explored a range of parameters for contact energies, Temperature and cell constraints. Here the objectives were to understand how the model behaves under different conditions, and to adjust parameters one at a time in order to calibrate the model with the available data. This was achieved by comparing cell shape and migration metrics recorded in simulation results with those measured in the *in vitro* system.