University of Dundee

DOCTOR OF PHILOSOPHY

A multiscale systems biology study of in vitro cell migration and cancer cell invasion

Schlueter, Daniela K.

Award date:
2013

Awarding institution:
University of Dundee

Link to publication

DOCTOR OF PHILOSOPHY

A multiscale systems biology study of in vitro cell migration and cancer cell invasion

Daniela K. Schlueter

2013

University of Dundee

A Multiscale Systems Biology Study of *in vitro* Cell Migration and Cancer Cell Invasion

Daniela K. Schlüter

Division of Mathematics

University of Dundee

Dundee

A thesis submitted for the degree of Doctor of Philosophy

April 2013

Contents

Acknowledgements

I would like to thank my supervisor Prof. Mark A.J. Chaplain for giving me the opportunity of working on this project and for having enough faith in me to give me a free hand to approach it as I saw fit. I would also like to thank him for his support and understanding especially during the tougher times of my PhD. Furthermore I would like to thank everybody in the Division of Mathematics at the University of Dundee for making me feel welcome from my first day onwards. Although he was only there for a short period of time, I would especially like to thank Dr. Paul Macklin for many helpful discussions. I also learned a lot from and enjoyed working with Dr. Ignacio Ramis-Conde and I would like to thank him for a fruitful collaboration and a great week in Cuenca. Furthermore I would like to thank Nick for helping me with anything computer-related, patiently answering all my stupid questions and being a great person to talk to about anything and everything. It has been very important for me to have some medical and biological links that kept me motivated throughout the project and for that I would like to thank Prof. Alastair Thompson, Dr. Susanna Fagerholm and Hwee San Lek.

For completing my Phd thesis, my personal relationships were just as important as the scientific ones. Therefore I would like to thank Mariya, Raluca, Gibin and Niall for making every day at the office and the weekends in Dundee enjoyable. I would also like to thank Chloe, Shabnam and James for distracting me from my work when necessary. A large thank you also goes to everybody at Appletree

Declaration

I, Daniela K. Schlüter, hereby declare that this thesis has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

Signature of candidate .... **Date** ....

Certification

This is to certify that Daniela K. Schlüter has complied with all the requirements for the submission of this Doctor of Philosophy thesis to the University of Dundee.

Signature of supervisor .... **Date** ....

Publications

[1] D. K. Schlüter, I. Ramis-Conde and M. A. J. Chaplain. Computational modeling of single cell migration: the leading role of extracellular matrix fibers. *Biophys. J.*, 103(6):1141-1151, 2012.

Chapter 1

Introduction

According to the World Heath Organisation, 7.6 million people died of cancer in 2008. This makes cancer one of the leading causes of death in the world. 90% of these deaths are due to metastases (secondary tumours) that develop from cells of the primary tumour which have invaded the surrounding tissue and spread throughout the body. A prerequisite for this spread is the cells' ability to migrate. It is important for cells to be able to migrate during development and wound healing and also in many other biological contexts. However, in the pathological case of cancer a disturbed balance of cell-cell and cell-matrix interactions leads to cell migration and invasion that can be fatal for the patient. In this thesis we take a systems biology approach to study cell migration and cancer cell invasion in an *in vitro* setting in order to shed more light on the underlying mechanisms and causes. *"Systems biology...is about putting together rather than taking apart, integration rather than reduction. It requires that we develop ways of thinking about integration that are as rigorous as our reductionist programmes, but different....It means changing our philosophy, in the full sense of the term"* (Denis Noble) [Noble, 2006]. In accordance with this definition of systems biology we develop a multiscale model of cell migration and cancer cell invasion. A considerable number of models of these processes have been developed

previously. However, generally the focus of these models is somewhat narrow and the influence of only one factor on migration and invasion is considered. Such a factor could for example be the interaction of the migrating cell or the mass of invading cells with a continuous field of chemoattractants or extracellular matrix elements or the loss of cell-cell adhesion. In this thesis, we explicitly couple cell-matrix interactions with cell-cell interactions and intracellular dynamics in order to study the combined influence of these different factors on migration and invasion. We believe that only by considering all of these processes can the observed behaviour of invading cancers be truly understood. We consider cells and matrix fibres as individual elements, which have a mutual interaction. The matrix fibres guide the cell movement but are in turn reorientated by the cells moving across them. Additionally, we include intra- and intercellular dynamics of cell-cell adhesion and repulsion and couple it with cell division. Furthermore, Western Blot data concerning the integrin signalling pathway, which is important in regulating cell-matrix interactions, is studied and used to develop a model of part of this complex pathway for future integration in the cell migration model. The model development is done in stages in the different chapters of this thesis. Chapters 2 and 3 give an overview of the biological and mathematical background, respectively. In Chapter 4 we develop a single cell migration model which is extended to a two-cell migration model in Chapter 5. This work is used in Chapter 6 as the basis of a multiscale model of cancer cell invasion. Finally in Chapter 7 we study the integrin pathway. With each model, data analysis techniques are developed in order to be able to quantify the simulation results and make them comparable to experimental data. We conclude the thesis with a chapter concerning further possible extensions of the models and future work.

Chapter 2

Cell migration, Carcinogenesis and Cancer Cell Invasion

2.1 The eukaryotic cell

The cell is the basic functional and structural unit of all living organisms. Cells are very complex entities and whole books have been written on their evolutionary development and function (e.g. [Alberts et al., 2002, Lodish et al., 2012]). There are two main groups of cells - prokaryotic and eukaryotic cells. Organisms whose cells are eukaryotic are called eukaryotes. Animals as well as plants, fungi and protozoans belong to the eukaryotes. Eukaryotic cells can vary in shape and are generally between 10 -100μm in diameter. What defines a eukaryotic cell is that it has a distinct nucleus which is surrounded by a membrane and is thus set apart from the rest of the cell, the cytolasm. Figure 2.1 shows a schematic diagram of a eukaryotic cell.

Animal cells are distinct from other eukaryotic cells in that they have a flexible membrane surrounding the cytoplasm rather than a rigid cell wall. In addition
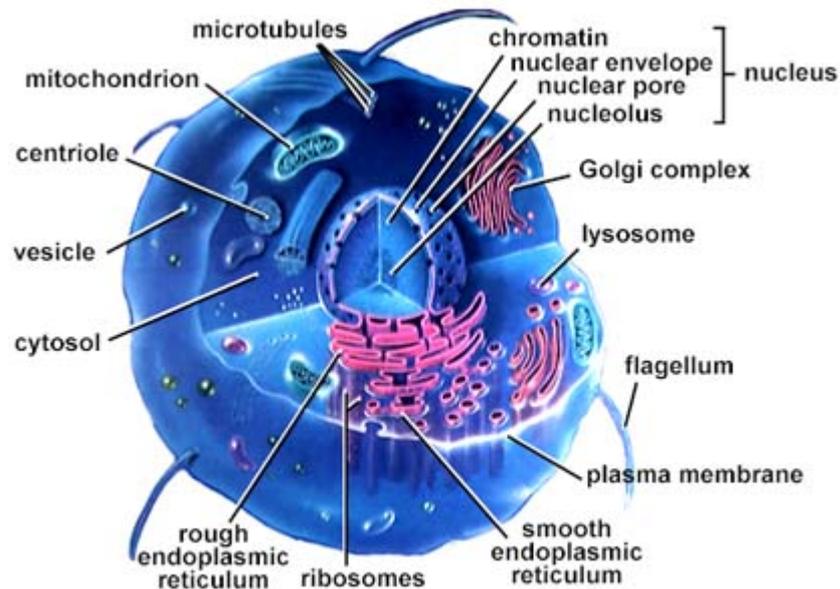
Figure 2.1: *Three dimensional schematic diagram showing the structure of a general eukaryotic cell. Image taken from* On-line Biology Book.

to being surrounded by a plasma membrane, they have several subcellular compartments, called organelles, enclosed by internal membranes. The nucleus is the largest of these organelles and it contains the cell's DNA. Not only does it contain the DNA but it is also the place where sections of the DNA are transcribed into m(essenger)RNA, which are then transported into the cytoplasm to be translated into proteins at the endoplasmatic reticulum. The proteins are in turn transported to specific sites in the cell where their function is required. The transport is helped by an array of fibrous proteins, collectively called the cytoskeleton, which forms a complex network in the cell's cytoplasm. The different types of these proteins have different functions. One specific type, the microtubules, for example, form the mitotic spindle which plays a major role in dividing the DNA equally between the two daughter cells during cell division. Thus without the cytoskeleton, cells could not reproduce. Correct cell division is
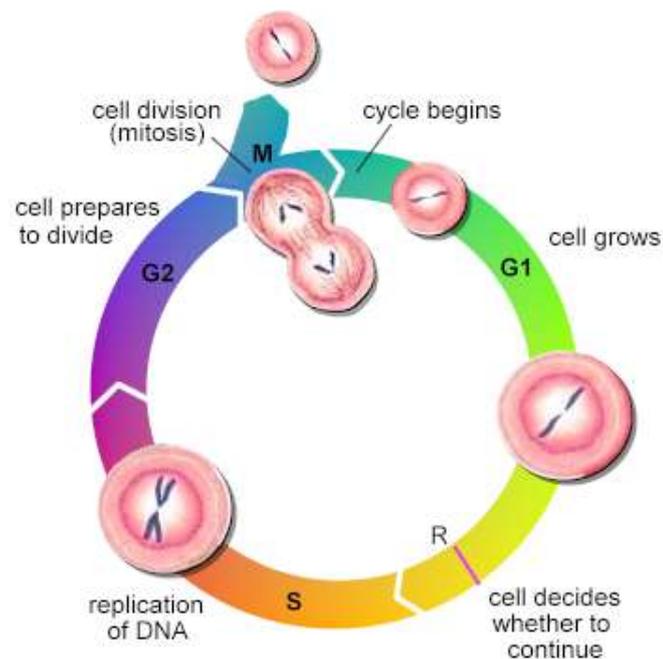
Figure 2.2: *Schematic representation of the four stages of the cell cycle. Image taken from the* Encyclopaedia of Science

however not only guided by this one type of protein, but also by the tightly controlled activation, inhibition and degradation of a number of proteins giving rise to the so called cell cycle which ensures the division into two equal daughter cells. The cell cycle is generally subdivided into four stages (see Figure 2.2). There are two gap phases, G1 and G2, during which the cell grows and mRNA and proteins are produced. The S phase, which follows G1, is the synthesis phase. During this time the structures which carry the DNA, the chromosomes, are duplicated. After the second gap phase the cell divides. This process is called mitosis. Sometimes a fifth stage is considered which is termed the G0 phase or the resting phase. In other considerations G0 is encompassed in G1. Cells enter the resting phase if they have developed and grown to their designated size but external or internal factors do not favour cell division. Such factors are, for example, stresses like low oxygen levels or anti-growth factors secreted into the environment by other cells
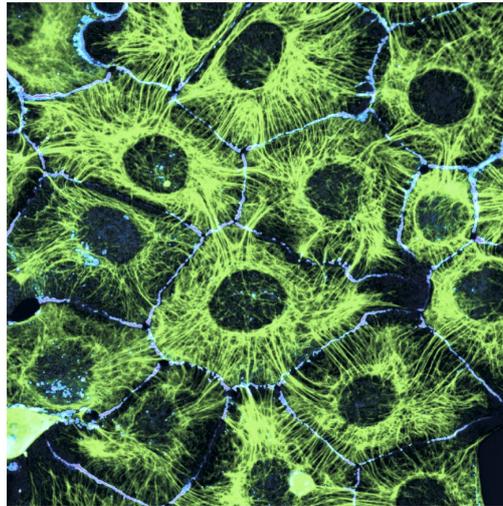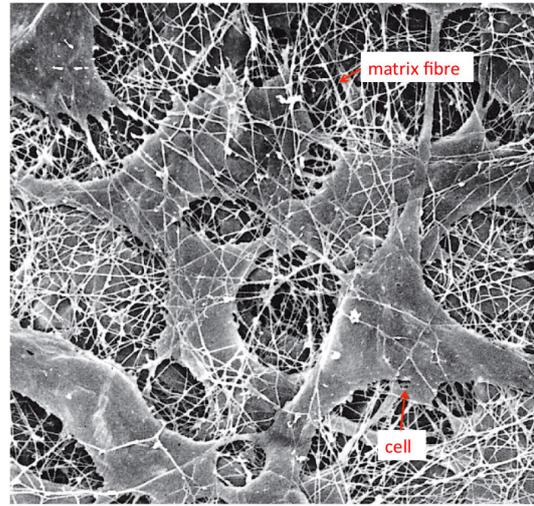
(a)	(b)

Figure 2.3: *Images showing the intracellular and extracellular scaffolding. (a) Image of a sheet of epithelial cells, where keratin, an element of the cytoskeleton, is detected with a keratin-specific antibody (green) and a plasma membrane protein is bound by a second antibody (blue). Image reproduced with copyright permission from Kathleen J. Green. (b) Image of the extracellular matrix meshwork in which fibroblasts (connective tissue cells) are embedded. Image adapted from Weinberg [2007].*

to limit growth and thus preserve tissue integrity.

In addition to guiding the movement of structures within the cell, the cytoskeleton has many more functions, one of which is giving the cell strength and rigidity to help it maintain its shape. It is also vitally important for organising individual cells into multicellular tissues. The sheets or chains of cells that form these tissues are held together by cell adhesion molecules which are, in turn, connected to, and in part, controlled by, the cytoskeleton. Figure 2.3(a) shows a sheet of cells where one specific cytoskeletal protein is tagged in green and the cell membrane is shown in blue. It can be seen that the green fibrils exist at the sites where

the cells are linked together. Not only cell–cell adhesion but also cell–matrix adhesion is controlled by the cytoskeleton. The cells in tissues are surrounded by a complex meshwork of proteins and polysaccharides which are secreted by the cells into spaces between them (see Figure 2.3(b)). This is called the extracellular matrix (ECM).

2.2 The extracellular matrix

The extracellular matrix is very versatile. Depending on the relative amounts of its various components, it can form a variety of shapes and have many different characteristics, each adapted to the functional requirement of that particular tissue. One of the roles of the ECM is to serve as a scaffold for the tissues but it also affects their development and physiology as well as influencing the cell's migratory properties, function and proliferation. The different components of the extracellular matrix fulfil different roles. Figure 2.4 shows a schematic image of the extracellular matrix and its components. The major proteins in the ECM are collagens. This is a family of fibrous proteins which is, for example, the main component of skin and bone. Its role in the ECM is mainly structural and it can form long fibrils which can assemble into highly ordered arrays. Another structural component of the ECM is elastin, which forms the main part of the elastic fibres, giving the tissues such as skin and blood vessels their elasticity. Fibronectin and laminin, two further proteins found in the ECM, have mainly adhesive functions. Fibronectin exists as both a soluble form in the blood and as insoluble filaments on the surface of certain cells which then deposit it in the matrix. Fibronectin is not only important for cell–matrix attachment but also for guiding cell migration.

The meshwork of these ECM protein fibres is embedded in a hydrated gel consisting of negatively charged polysaccharide chains covalently linked to proteins.
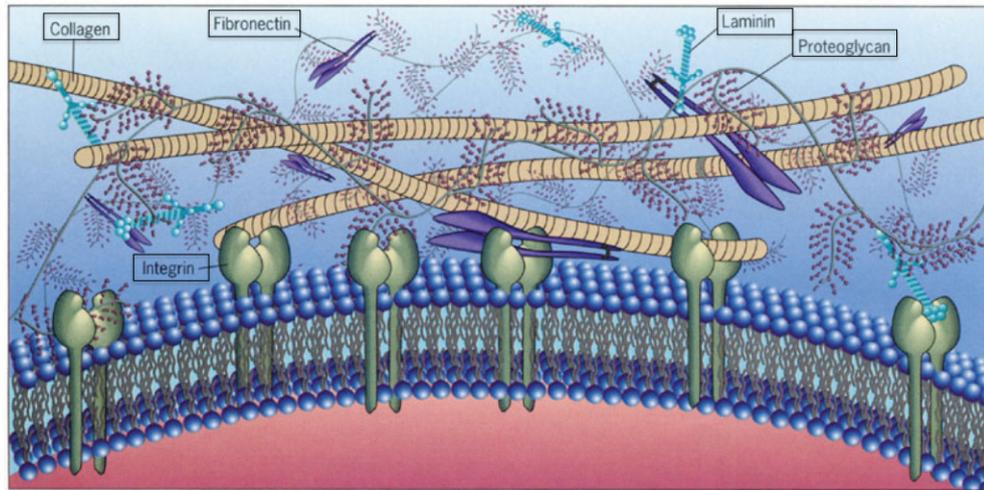
Figure 2.4: *Schematic image of a cell membrane and the extracellular matrix surrounding the cell. Different components of the extracellular matrix are shown - the proteins collagen, fibronectin and laminin as well as the proteoglycans. Image adapted from the on-line course of cell biology by Sichuan University.*

The complex of polysaccharide chains and a core protein forms a proteoglycan molecule. The polysaccharide chains adopt very extended conformations and thus fill a very large volume compared to their mass. Due to their negative charges, they attract osmotically active cations. This leads to water moving into the matrix and thus creates a swelling pressure that enables the matrix to withstand compressive forces. In addition to keeping the space around the cells hydrated, the proteoglycan molecules can also form porous gels of different pore sizes and charge densities which allows them to regulate the traffic of molecules in the extracellular matrix. This can, for example have an effect on cell signalling [Alberts et al., 2002].

Although cells are highly influenced by the surrounding ECM, the interaction is not unidirectional. In addition to secreting extracellular matrix components, cells also rearrange and organise the extracellular matrix just like the ECM guides and influences the cells.

2.3 Cell migration

The migration of individual cells occurs in a wide variety of biological contexts ranging from development and wound healing to malignant diseases such as cancer [Lauffenburger and Horwitz, 1996, Yang and Weinberg, 2008, Baum et al., 2008, Lee et al., 2006]. Figure 2.5 shows a schematic diagram of the four-step movement cycle. In order to migrate, a cell first needs to acquire front-rear polarity, which is in itself a very complex process [Etienne-Mannevielle, 2008, Tanos and Rodriguez-Boulan, 2008]. The direction that a cell polarises in can be determined by extracellular cues such as growth factors, chemical gradients and extracellular matrix components through spatially limited activation of signalling complexes inside the cell [Huttenlocher, 2005]. The polarity is stabilised and sustained during migration by multiple feedback mechanisms that include, among others, integrins - cell–matrix adhesion molecules, which maintain the spatial molecular asymmetry [Huttenlocher, 2005, Lauffenburger and Horwitz, 1996]. Complexes at the front of the cell interact with the actin filaments of the cytoskeleton leading to polymerisation and extended membrane protrusions [Huttenlocher, 2005, Mogilner, 2007]. This is achieved by controlling the pool of free actin monomers as well as the number of free actin filament ends, for example by preventing self-nucelation and by targeting monomers to the free ends [Ridley et al., 2003]. Certain capping proteins terminate this process and also ensure that it only occurs in filaments close to the membrane. A different set of complexes ensures the disassembly of older filaments at the rear of the cell to ensure the supply of monomers [Ridley et al., 2003]. The resulting protrusions can be large, widespread lamellipodia or thin filopodia (see Figure 2.5). It is suggested that these two structures have distinct functions. The broad 'brush-like' structure of the lamellopodium can push on large areas of the plasma membrane thus inducing directional migration whereas the spike-like filopodia might be acting more as
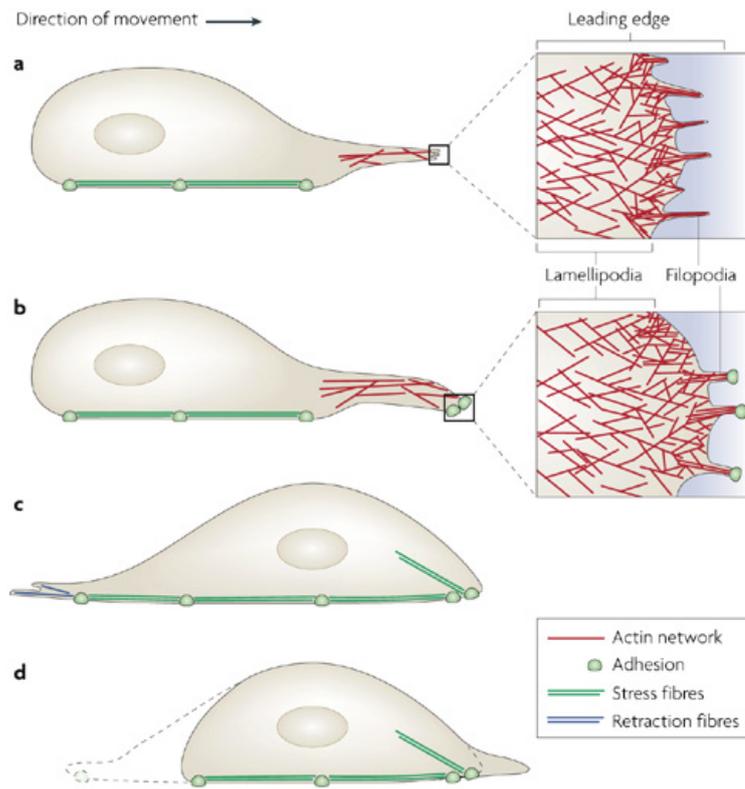
Figure 2.5: *Image showing the four main steps of cell migration. Once a cell has established front-rear polarity, protrusions are formed through actin polymerisation. Then new adhesions are formed with the matrix at the front of the cell. Subsequently contraction leads to translocation of the cell body and the rear is retracted by breaking the cell–matrix bonds and disassembling the actin filaments. Reproduced with copyright permission from Mattila and Lappalainen [2008].*

sensors to explore the environment [Ridley et al., 2003]. Both types of protrusions bind to the extracellular matrix through integrins. Integrins are transmembrane cell–matrix adhesion molecules which are activated by binding to the extracellular matrix. These molecules consist of two subunits : an α and a β unit. In mammals there are multiple types of both, the different combinations of which lead to 24 different kinds of integrins. All of these different types bind to distinct subsets of ECM ligand, although they are partially overlapping. Once integrins are activated they cluster and initiate signalling processes on the inside of the cell leading to the accumulation of adaptor proteins and proteins involved in actin polymerisation which is why they can influence and initiate cell polarisation [Guo and Giancotti, 2004]. Integrins bind to the actin filaments via these adaptor proteins establishing a solid link between the cell's cytoskeleton and the extracellular matrix. The binding of adaptor molecules furthermore increases integrin activity by changing its conformation to the high affinity state [Ridley et al., 2003]. This process takes place over a period of a few minutes during which small focal complexes stabilise to form focal contacts [Friedl and Wolf, 2003, Ridley et al., 2003, Zamir and Geiger, 2001]. Depending on the type of cell these focal contacts are more or less pronounced. Fast migrating leukocytes, for example, only have very few visible integrin clusters. Thus very small adhesions are probably important for their migration [Ridley et al., 2003]. The connection between the cytoskeleton and the ECM can be used for two purposes - it gives the cell traction to migrate and it also enables the cell to gather mechanical information about the environment it is in. Following focal contact formation, cell contraction leads to the generation of traction forces and therefore the forward movement of the cell body whereby any cell–matrix bonds at the rear of the cell are released [Friedl and Wolf, 2003, 2009, DiMilla et al., 1991, Lauffenburger and Horwitz, 1996]. In a three-dimensional matrix cells are surrounded by matrix elements and thus their

11

movement is obstructed by the lack of space to move into. Cells have evolved different mechanisms to deal with this, which depend on the cell's characteristics as well as the density and rigidity of the matrix [Friedl and Wolf, 2010]. Some cells adapt to their environment by becoming highly deformable so that they can squeeze through gaps in the matrix provided the matrix is pliable and the gaps are of a reasonable size. This is called amoeboid migration. Other cells however create a space to migrate into. They break down the matrix fibres that block the preferred path. This process is called focalised proteolysis and the type of migration is termed mesenchymal migration [Friedl and Wolf, 2009, 2010]. Figure 2.6 shows the different types of migration by which cells can migrate in two- and three-dimensional tissues (Figure 2.6(a)) and the conditions under which cells acquire these different migration modes (Figure 2.6(b)).

A key component of all cell migration is the interaction with the individual fibres of the matrix. This becomes clear when closely examining the interactions between a cell and the individual matrix fibres as is done, for example, in the experiments shown in Figure 2.7. These images clearly show individual cells interacting with single fibres and reorienting these fibres which changes the environment for themselves and also for other cells. Recent experimental studies have investigated in more detail the importance of this remodelling of individual fibres, of cell adhesion and of force generation on two-dimensional surfaces [Friedrichs et al., 2007, Ludwig et al., 2008, Kirmse et al., 2011, Jiang et al., 2004, Lo et al., 2000, Poole et al., 2005]. It was shown, for example, that cells align the matrix during migration by reorienting individual collagen fibrils. Surprisingly the deformation of the matrix occurred asymmetrically revealing a matrix anisotropy which supports directional cellular traction and cell polarisation. Furthermore it was found that the directional motility of certain cells on a two-dimensional collagen matrix not only depends on the global structure given to the matrix by the

(a)

Morphology   **Migration mode**

Amoeboid (blebby)

Amoeboid (pseudopodal, filopodal)

Mesenchymal

Multicellular streaming

Collective
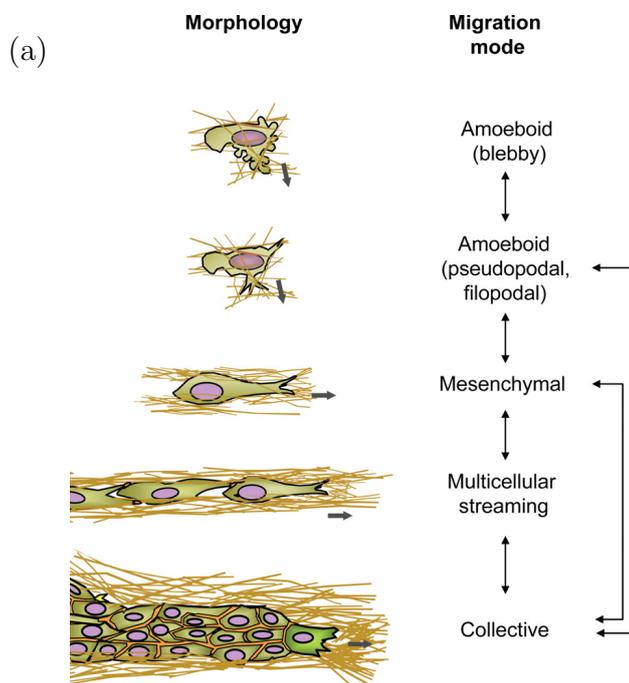
Figure 2.6: *Schematic diagram showing the different types of migration cells can acquire.*

(a) Representation of the different migration modes in which malignant tumour cells can invade the surrounding tissues.

(b) Table of the determining factors for the different migration modes. Images reproduced from Friedl and Wolf [2010] in accordance with RUP copyright policy.

(b)
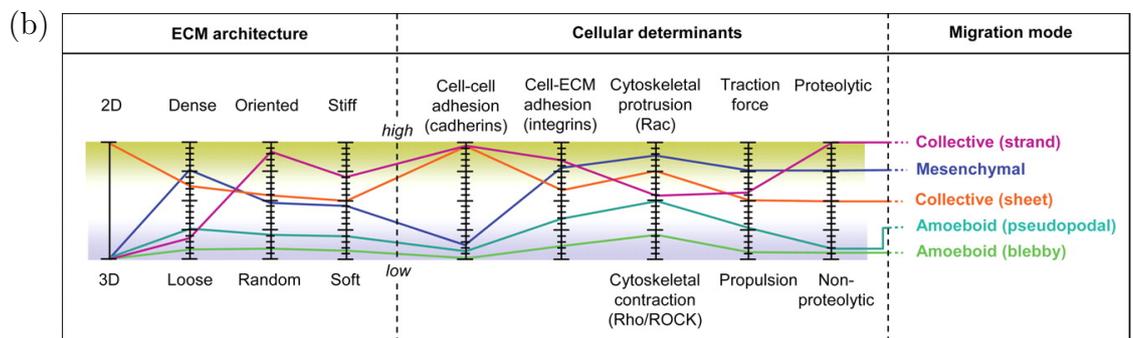
individual collagen fibrils, but also on the molecular-scale structure of the fibrils. These recent studies show that there is still a lot unknown about cell migration even in the simplest setting of a single cell on a two-dimensional matrix and that therefore this field sparks evermore biological and medical interest.

13
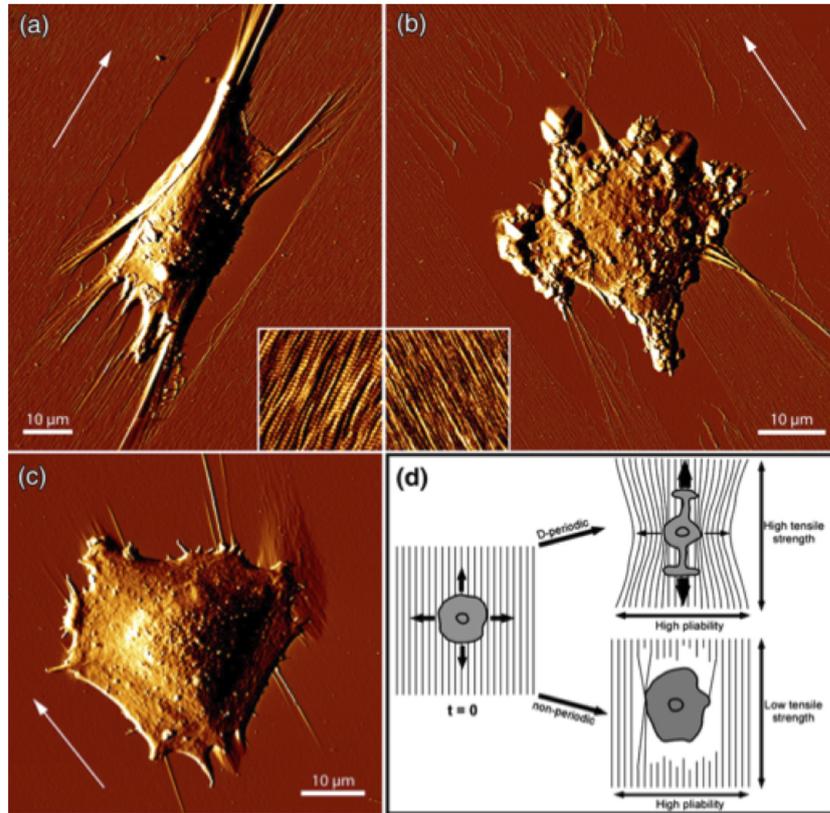
Figure 2.7: *Experimental images of individual cells interacting with collagen matrices of differing fibre alignment. (a) D-periodic fibres (i.e. anisotropic) (b) non periodic fibres (i.e. isotropic) (c) glutaraldehyde-fixed D-periodic fibres. Reproduced with copyright permission from Friedrichs et al. [2007].*

2.4 The epithelial phenotype

The human body is made up of trillions of cells all of which can be organised into over 200 types. However, all of these different cell types can be grouped together as components of five main classes of tissues: epithelial tissue, connective tissue, muscular tissue, nervous tissue and blood [Lodish et al., 2012]. Epithelial cells form sheets that line the hollow organs in the body, for example the lung, kidney and the gallbladder and the external surface of organisms [Weinberg, 2007]. Through this they protect the underlying tissue from the contents of the cavity

14
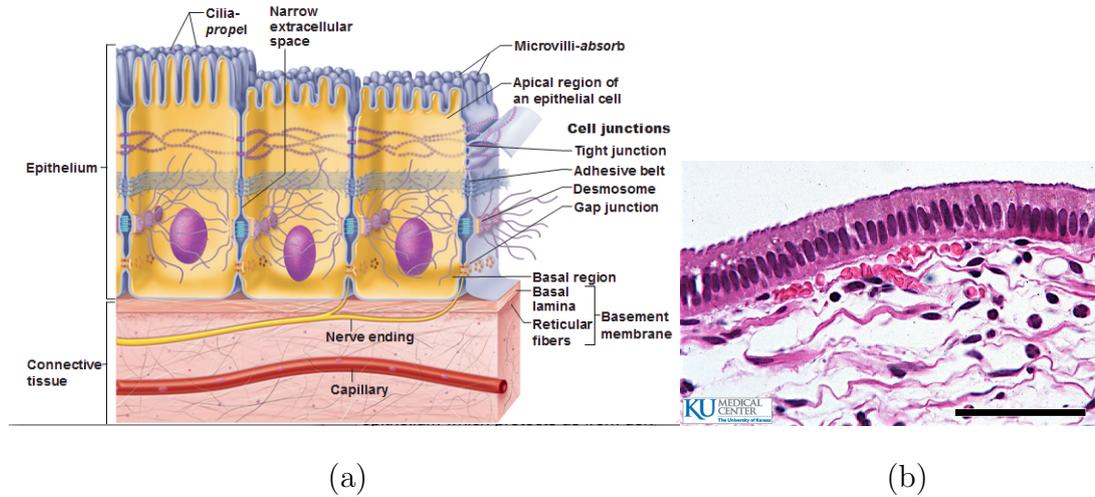
Figure 2.8: *The epithelium. (a) Schematic diagram of the epithelium with the underlying basement membrane and loose connective tissue. Image taken from Antranik.org. (b) Image of the endothelium lining the gall bladder with the underlying basement membrane and loose connective tissue with blood vessels, bar=100 μm. Image taken from the University of Kansas Medical Centre website.*

and transport fluids and essential nutrients [Weinberg, 2007, Freshney, 1992]. The skin is for example an epithelium which protects us from dehydration due to loss of water as well as from any external factors. Beneath the sheet of epithelial cells lies the basement membrane which serves as a scaffold for the cells and separates them from the connective tissue underlying it [Weinberg, 2007]. The basement membrane is a very dense layer of extracellular matrix. Epithelial cells that form a sheet situated on a basement membrane are fully differentiated and generally have a cuboid or columnar shape (see Figure 2.8)[Freshney, 1992]. Their nucleus is located in the lower third of the cytoplasm. On the apical side, the side facing inward into the lumen of a cavity, they have microvilli. The apicolateral surfaces form strong adhesion bonds with their neighbours through multiple junctional complexes. [Freshney, 1992]. There are strong gradients in the molecule concentrations from the apical to the basal side of the cell, and vice versa, which are

initiated and maintained by the cell–cell adhesion complexes at the apicolateral sides of the cell as well as the cell–matrix adhesion at the basal end. This gives the cell a characteristic apical-basal polarity. Different types of junctions exist that mediate the cell–cell interactions. The three main classes are anchoring junctions, tight junctions and gap junctions. While gap junctions allow the fast transport of small molecules between the cytoplasms of two neighbouring cells, anchoring junctions and tight junctions are responsible for holding the tissue together. Again there are multiple different types of anchoring junctions, the most studied of which are probably the adherens junctions. This special type of junction consists of adhesion proteins in the plasma membrane, called cadherins, which bind to proteins on the neighbouring cell's surface, adaptor proteins, which connect the adhesion proteins with the cytoskeleton, and the cytoskeleton filaments themselves [Lodish et al., 2012]. Epithelial cells express E(pithelial)-cadherin which is in this case not only a cell–cell adhesion protein but also a marker for the epithelial phenotype as well as an inducer of the apical-basal polarity.

2.4.1   E-cadherin and β-catenin

E-cadherin is one of the most important adhesion molecules in epithelial tissues [Pecina-Slaus, 2003]. It is a parallel homodimer with five cadherin repeats bound together by calcium ions forming the "stiff, rod-like" [Gumbiner, 2005] extracellular region of the protein. The intracellular domain binds to a complex of different catenins which link it to actin binding proteins and the actin cytoskeleton. Figure 2.9 gives a rough representation of this. In adherens junctions the extracellular domain of one E-cadherin dimer forms a homotypic bond with an E-cadherin dimer on the neighbouring cell (see Figure 2.9). One of the many roles of E-cadherin is the binding of β-catenin. This not only forms a functioning cell–cell adhesion site but it also prevents β-catenin from interacting with other proteins.
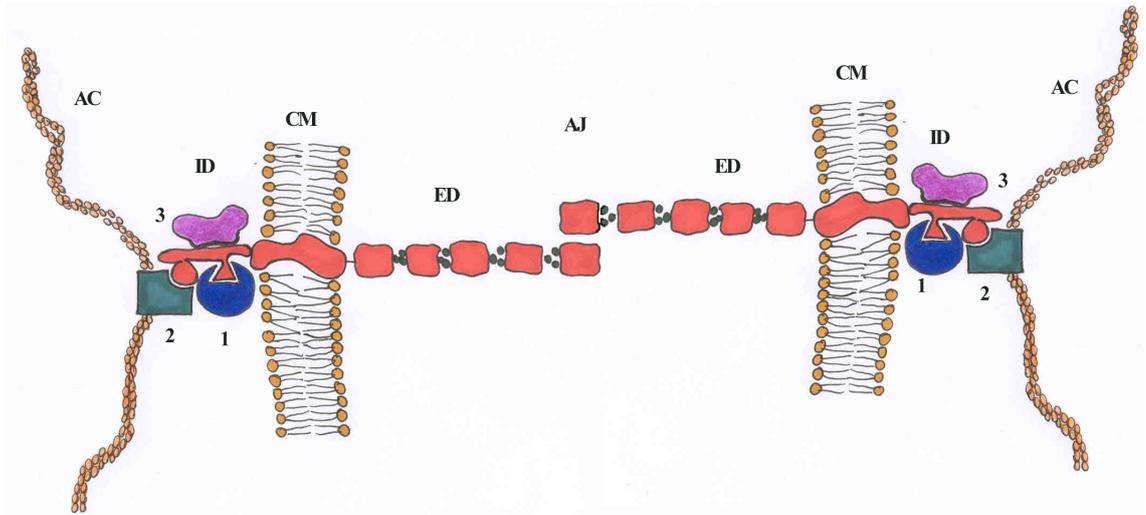
Figure 2.9: *Schematic diagram of an adherens junction and the proteins involved. AJ=adherens junction, ED=extracellular domain, CM= cytoplasmic membrane, ID= intracellular domain, AC=actin cytoskeleton, 1=β-catenin, 2=α-catenin, 3=p120-catenin. Image reproduced from Pecina-Slaus [2003] in accordance with the Open Access license.*

β-catenin exists in three pools in the cell: 1) bound to E-cadherin at the cell membrane, 2) in the nucleus associated with the LEF/TCF transcription factors, 3) in the cytoplasm where it can associate with the adenomatous polyposis coli (APC) tumour suppressor gene product and other proteins which mark it for degradation [Juliano, 2002, Wong and Gumbiner, 2003]. It has been suggested that the binding of β-catenin to E-cadherin and the binding to APC are processes that cooperate to maintain tissue integrity and prevent tumour development [Huelsken et al., 1994]. However the activity of β-catenin as transcription a factor has rather opposite effects. Therefore the E-cadherin–β-catenin pathway has been studied extensively. A brief overview over E-cadherin–β-catenin dynamics starts with E-cadherin and β-catenin binding at the endoplasmatic reticulum immediatly after production [Hinck et al., 1994]. The complex is then trafficked to the cell membrane [Hinck et al., 1994, Chen et al., 1999]. Other molecules such as α-catenin

can then bind and when cell–cell contact takes place the complex can form adherens junctions with E-cadherin-β-catenin complexes on the neighbouring cell [van Roy and Berx, 2008]. Junction disassembly followed by endocytosis lead to the disruption of the E-cadherin-β-catenin complex and the components can either be degraded, recycled for cell–cell adhesion or reused in different signalling contexts. As mentioned above, the fate of the free β-catenin molecules needs to be tightly regulated and is partly dependent on the presence or absence of proteins of the Wnt family. Wnt proteins bind to the cell surface receptor Frizzled which in turn activates the Dishevelled family proteins. Active Dishevelled proteins can inactivate a complex of axin, GSK-3β and APC. This complex, when activated, leads to the phosphorylation of β-catenin which can subsequently be ubiquinated and thus marked for degradation. The presence of Wnt therefore inhibits the degradation of free β-catenin and a cytoplasmic pool can stabilise. If the free β-catenin level is above a certain threshhold, some β-catenin molecules can enter the nucleus and can interact with the transcription factors of the TCF/LEF family to promote the transcription of specific genes. These genes are pro-migratory and pro-invasive (e.g. MMP7)[Hulsken and Behrens, 2000]. High levels of nuclear β-catenin can be found at the invasive front of a tumour whereas in central areas it is located at the membrane in E-cadherin–β-catenin complexes [Brabletz et al., 2001]

2.4.2 Epithelial cell migration and the epithelial to mesenchymal transition

"Epithelial cells do not migrate!" This statement was made by a famous developmental biologist in 1990 [Quaranta, 2002]. The citation has been used to introduce a paper on the motility cues of the tumour environment in which it is explained that this statement is no longer seen as the ultimate truth and that it

is now agreed that epithelial cell migration does occur for example in the context of wound healing as well as in malignant cancer cell invasion [Quaranta, 2002]. In order for the cell to migrate as explained above, changes in the actin cytoskeleton and cell adhesion are required. It is now generally agreed that a prerequisite is for the cell to undergo some measure of epithelial to mesenchymal transition which includes loss of the characteristic apical-basal polarity and cell–cell adhesion. In wound healing this transition generally only occurs in the leading cells and it is only partial with minor loss of cell–cell contact. Malignant cells undergo a more complete transition which leads to the migration of cell sheets and strands as well as individual cells. Depending on their characteristics and their environment, the single cells can migrate either in a mesenchymal or amoeboid fashion (see Figure 2.6). In migrating cell clusters, the leading cells are generally highly mobile whereas the following cells are more passive.

The epithelial to mesenchymal transition (EMT) was first described in the context of embryology. Here it plays a crucial role, for example, in the formation of the multilayered structure from a single epithelial layer during gastrulation [Levayer and Lecuit, 2008]. In order for this structure to form, epithelial cells have to lose their epithelial characteristics and become motile mesenchymal cells. Similarly, cells from epithelial cancers generally have to undergo EMT before they can start invading and migrating through surrounding tissue.

Many molecular processes are involved in the transition from an epithelial to a mesenchymal cell. Certain transcription factors are activated and, among others, specific cell surface proteins as well as cytoskeletal proteins are expressed. One of the main markers of epithelial cells, E-cadherin, is furthermore lost or down-regulated. This can be seen at the periphery of many invasive cancers as well as in embryonic cells undergoing EMT [Weinberg, 2007]. Instead, these cells acquire mesenchymal properties such as front-rear polarity and spindle shape together
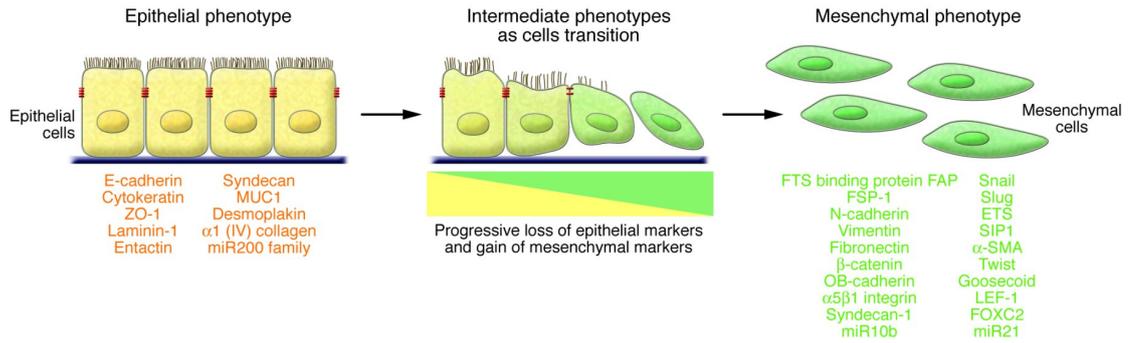
Figure 2.10: *Schematic diagram of the process of the epithelial to mesenchymal transition and the biomarkers linked with each stage. Reproduced with copyright permission from Kalluri and Weinberg [2009].*

with the expression of mesenchymal markers such as N-cadherin and integrins. All of these changes in the protein expression pattern are used as biomarkers for the progression of the EMT (see Figure 2.10) [Kalluri and Weinberg, 2009]. Many different triggers for this switch have been suggested. Whereas a lot of work has been focussed on identifying molecular players in the switch and many E-cadherin antagonists have been found [Kang and Massague, 2004, Schmalhofer et al., 2009], it has also been linked to the cells' microenvironment. The role of stromal cell types like fibroblasts, macrophages and inflammatory cells in inducing up-regulation of EMT drivers has been shown [Weinberg, 2007]. Also the possible importance of the extracellular matrix has been identified [Nakaya and Sheng, 2008]. This work proposed that the disassembly of the basement membrane rather than that of the adherens junctions formed by E-cadherin, is a key driver of EMT. This was the conclusion after noticing that the breakdown of the basement membrane was the first noticeable event during EMT in chick embryos and thus preceded the loss of adherens junctions as well as that of epithelial markers such as E-cadherin. Whether the induction of EMT after basement membrane breakdown is due to the changes in the mechanical forces or because the barrier that holds back certain growth factors and other chemicals has been breached is

not yet clear, as one can never be observed without the other [Ingber, 2002]. The EMT is seen as being completed when the basement membrane underlying an epithelial layer is partly degraded and a mesenchymal cell has formed which can migrate away from the layer it originated from [Kalluri and Weinberg, 2009]. However, this process is reversible. Not much is known about the mesenchymal to epithelial transition (MET) but it is an important process during development where cells within certain tissues move back and forth between epithelial and mesenchymal phenotypes [Kalluri and Weinberg, 2009]. It has also been suggested that it plays a part in the formation of metastases (secondary tumours). Some of the cells in the primary tumour might turn into a more mesenchymal phenotype after receiving EMT-inducing signals. This allows them to invade the surrounding tissues and travel to other sites in the body. Here the EMT-inducing signals might be absent which could lead to the reverse process, the MET. Thus cells at metastatic sites revert to the phenotype shown by other cells inside the primary tumour [Weinberg, 2007].

2.5 Carcinogenesis

Carcinogenesis is defined as "the process by which normal cells are transformed into cancer cells" (NCI Dictionary of Cancer Terms). It is a multi-step process which generally takes place over a series of decades. The progression sees normal cells evolve and form benign tumours which are localised and non-invasive. They then become more and more malignant through gaining the ability to invade the local tissue and form metastases at distant sites in the body. Malignant tumours are called cancers [Weinberg, 2007]. The metastases of these malignant tumours are the main causes of death. About 90% of cancer deaths are due to these secondary tumours. The reason is that whereas the primary tumours generally have

to be of a very large size in order to affect the normal tissue function, the metastases often cause great devastation, most frequently by disrupting vital organs like the brain or the liver. An example is breast cancer which usually would not endanger life, but metastases arising from it are likely to form in the bone tissue which can cause erosion and finally skeletal collapse [Weinberg, 2007].

Tumours are classified by their origin. The most common human tumours arise from epithelial tissues and are called carcinomas. They are responsible for 80% of cancer deaths. Most carcinomas fall into one of two categories that are associated with the function of the epithelial cells. Squamous cell carcinomas arise from epithelia whose main function is the lining of cavities and channels. These cancers include those of the skin, the nasal cavity and the cervix. The second category are the adenocarcinomas which arise from epithelia that contain specialised cells which secrete substances into the cavities that they line, mainly in order to protect the epithelial cells (e.g. from the acidity in the stomach). These include the colon, the breast, the pancreas and the stomach [Weinberg, 2007].

Different theories exist as to how tumours are initiated and promoted. The most widely held view sees cancer development following darwinian evolution. In this context the evolving entities are individual cells. Random mutations occur in the different cells of a tissue leading to genetic variation in this cell population. In this heterogeneous population selection may favour the prominence of an individual cell which has advantageous traits concerning proliferation and survival in the existing microenvironment. This cell and its descendants form a clone which can expand rapidly. Once this clone has reached a certain size another advantageous mutation might occur in one of these cells leading to a new, doubly mutated clone which can outgrow the clone it originated from and so on. Certain foods and other external factors that have been shown to increase cancer development might speed up the mutation rate and influence the development that way. This

is a very simplistic view of cancer development, yet it can easily be extended to include new findings coherently.

Another theory sees cancer development not driven by mutations alone but also by its environment. Thus cancer may be seen as a disease of deregulation of the process of cell integration into tissue or tissue integration into organs [Ingber, 2002]. It should therefore be studied as a disease of the tissue and not of an individual cell [Sonnenschein and Soto, 2011]. Reasons for this are that the review of experimental evidence has shown that at a genetic level, cancer cells cannot in fact be distinguished from normal cells just as little as cells from malignant tumours can be distinguished from those of benign tumours. In addition many cell behaviours seen in cancer development (growth, motility etc.) are normal behaviours during embryonic development and can be induced *in vitro* by altering the physical interactions between the cells and the extracellular matrix [Ingber, 2008]. Similarly it has been shown that the cancer phenotype can be reversed by integrating cancer cells into normal tissue [Sonnenschein and Soto, 2011]. In general, mechanical factors can lead to the diverse array of cell shapes and arrangements occurring during organ development and thus have a strong influence on cell behaviour at a tissue level.

A relatively new but quickly evolving concept is that of cancer stem cells. This theory evolved due to the fact that tumours form tissue clones. Thus the cells in a tumour are not as heterogeneous on a genetic, epigenic, phenotypic or any other level as one would expect if the development was driven by the accumulation of mutations in the cells and the concept of darwinian evolution. Normal tissues are generated and maintained by stem cells. These cells undergo coordinated processes to produce progeny that can either be stem cells or cells that will become highly differentiated depending on the requirements. Cancer stem cells are defined in the same way with the difference that they do not produce normal

tissue cells but can generate a fully malignant cell population. Thus in order to eradicate a cancer, the cancer stem cells have to be targeted by the therapies. Whatever the precise underlying causes, the general progression in carcinogenesis is the same. The development begins with abnormal cell proliferation. However the cells are constrained to stay on the epithelial side of the basement membrane. At this point carcinomas are still benign. Tumours at this stage are called hyperplastic or metaplastic. Hyperplastic tissues appear to be normal apart from an excess number of cells. In metaplastic tissues, cells are reversibly replaced by another differentiated cell type which is normal but usually not encountered at that site. Dysplasia is the next step in carcinogenesis and forms the transition from benign to premalignant tumours. Dysplastic tissues are cytologically abnormal
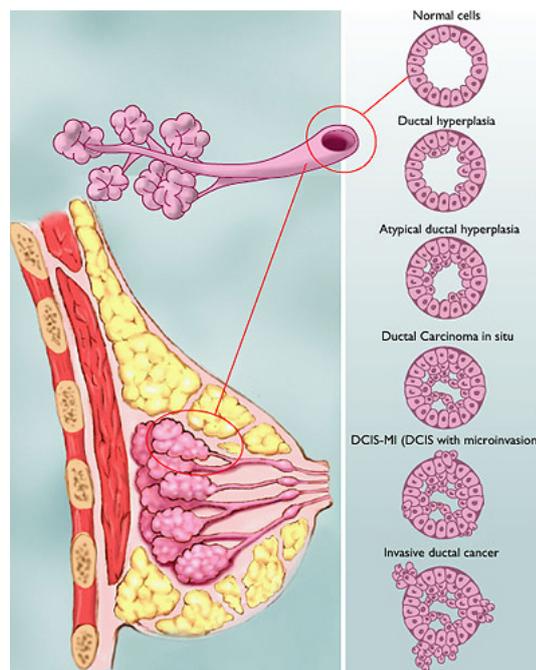
Figure 2.11: *Schematic diagram showing the development of ductal carcinoma. The normal cells first turn hyperplastic and then develop into a ductal carcinoma in situ, a high grade dysplasia. Then they acquire the ability to breach the basement membrane and lead to microinvasion before turning into a fully invasive ductal carcinoma. Image taken from the BC Cancer Agency website.*

but still benign as they respect the boundary created by the basement membrane. A tumour turns malignant once it has gained the ability to breach the basement membrane and invade the surrounding tissue. The final degree of abnormality is reached when metastases are formed [Weinberg, 2007]. Figure 2.11 shows this development in the case of a specific kind of breast cancer, ductal carcinoma. Here the different stages can be seen very clearly. In bigger and flatter tissues the process is the same (imagining cutting the ducts open and flattening them out gives an image of the development in flat sheets).

The breach of the basement membrane in carcinogenesis is the beginning of the "invasion-metastasis cascade" [Weinberg, 2007] (see Figure 2.12). Although most avascular tumours are able to induce the growth of a vascular network on the stromal side of the basement membrane and thus secure the nutrient and oxygen supply, breaking through the membrane and moving through the extracellular matrix enables them to get direct access to the blood vessels and lymphatic vessels. This does not only mean that they have a better nutrient supply and can thus proliferate even more, but it also means that some cells might be able to force their way through the vessel wall, a process called intravasation. Once inside the blood stream (or lymph system), the cancer cells are transported around the body. At distant sites they can extravasate (move through the vessel wall into the tissue) and form secondary tumours. Only a small number of cells are usually capable of surviving this journey and even fewer are able to lead to a cell population of detectable size at a new site where the environment is different to that at its site of origin [Weinberg, 2007]. However, enough cells successfully manage this to lead to 157,275 cancer deaths in the UK alone in 2010 (according to Cancer Research UK). Statistics show that cancer is one of the main causes of death in the western world and understanding how cells migrate and invade local tissues as well as how they get into the blood vessels is of utmost importance in

primary tumor formation • localized invasion • intravasation • transport through circulation • arrest in microvessels of various organs • extravasation

interaction with platelets, lymphocytes, and other blood components

formation of a micrometastasis

colonization – formation of a macrometastasis
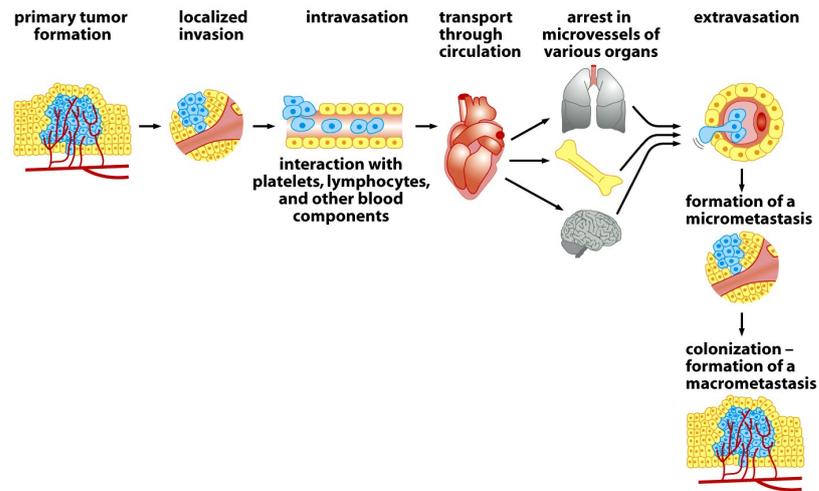
Figure 14-4 The Biology of Cancer (© Garland Science 2007)

Figure 2.12: *Schematic diagram showing the invasion-metastasis cascade. Image reproduced with copyright permission from Fidler [2003] as adapted by Weinberg [2007].*

order to block the spread of cancer through the body.

Chapter 3

Mathematical and Computational Modelling of Cell Migration and Invasion

As explained in the previous chapter, cell migration is a very important process in many biological systems ranging from embryonic development to wound healing to diseases such as cancer. Because of its complexity and importance in such a wide range of processes, it has been studied and modelled mathematically and computationally as part of different biological and medical scenarios as well as independently of the specific context.

Models which study cell migration itself, exist across a range of spatial scales from the subcellular level to the level of tissues and organs. Models at the subcellular level include those that study actin and myosin dynamics during lamellipod protrusion and cell contraction (e.g. [Novak et al., 2008, Mogilner and Edelstein-Keshet, 2002, Peskin et al., 1993]). Actin filaments fill the lamellipod with a rectangular network. The protrusion is driven by constant polymerisation and

capping of the leading filaments. Polymerisation leads to the front end of the filament bending against the cell membrane creating an elastic force which leads to the extension of it. Models of this process were, for example, able to predict that the filaments at the front of the cell have to grow at an angle to the leading edge as parallel growth can occur fast but does not lead to protrusions and normal growing filaments are too rigid and do not bend enough to create the elastic force needed. The models employ a variety of modelling techniques and are reviewed extensively in Mogilner [2007].

Moving up to the cellular level, detailed models of single cell mechanics coupled with adhesion receptor dynamics on the surface of the cell during migration have been developed. One of these techniques models the cell as a chain of elastic springs and viscous dashpods (e.g. [DiMilla et al., 1991]). With this type of model it was possible to predict the speed of migrating cells and the biphasic dependence on multiple factors such as receptor and ligand densities. The shape change of a cell has also been modelled as a moving boundary problem [Mogilner, 2007].

Cell migration models that take into account the cell's microenvironment by modelling the structure and organisation of the extracellular matrix have also been developed. These models generally fall into the two classes (1) hybrid discrete-continuum (e.g. [Dallon et al., 1999, McDougall et al., 2006]) and (2) continuum models (e.g. [Hillen, 2006, Chauviere et al., 2007, Painter, 2009]). The hybrid discrete–continuum models consider cells as discrete entities but the microenvironment is modelled as a continuum. This technique has been used, for example, to study cell migration in the context of wound healing in order to understand why scar tissue shows a higher degree of alignment than unwounded, normal tissue [Dallon et al., 1999]. Finding the answer to this question is important for the development of anti-scarring therapies. In the model the cells move across the

matrix by contact guidance and produce matrix elements as well as rearranging the matrix locally. Studying different scenarios with this model led to the results that the alignment of scar tissue is influenced by cell speed and initial matrix orientation among other things. Although this model specifically concentrates on the process of wound healing, it gave some initial ideas to the work developed in this thesis concerning the modelling of cell–matrix interactions.

Continuum models do not consider individual cells but cell densities as well as a continuous description of the matrix and thus study the problems at a tissue level. With these models it has, for example, been studied how a group of cells behaves under the influence of a chemotactic field in addition to a heterogeneous fibrous matrix [Chauviere et al., 2007] and what patterns arise under the different migration modes, amoeboid and mesenchymal [Painter, 2009].

Cell migration in the context of cancer cell invasion has also mainly been considered in discrete-continuum or continuum models which employ systems of reaction-diffusion-taxis partial differential equations (e.g. [Ramis-Conde et al., 2008a, Anderson et al., 2000, Chaplain and Lolas, 2005, Gatenby and Gawlinski, 2006]). Here cell migration is generally governed by random motility and a directed response to gradients in the extracellular matrix (haptotaxis). These models clearly have the advantage of being computationally very efficient and they thus offer the opportunity to study the influence of a number of different factors, e.g. chemotaxis, haptotaxis, matrix degradation, on the cell's behaviour at an individual cell level (hybrid models) and at the tissue level (continuum models). They however neglect the fact that these processes take place at the level of single cells interacting with individual matrix fibres.

In this thesis models are developed that are based on specifically modelling the individual elements. Other individual-based models of cell migration and invasion can be found in the literature and a brief review of the different individual-based
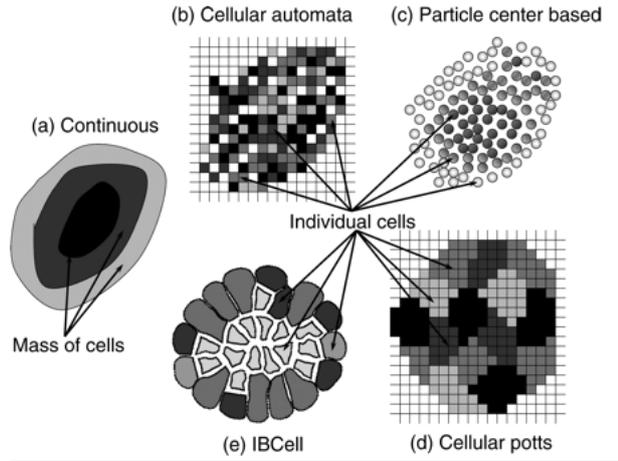
Figure 3.1: *Schematic diagram illustrating different modelling techniques. (a) Continuous models consider a mass of cells whereas individual-based models ((b)-(e)) consider individual cells. The cells are modelled with different levels of detail in the different individual-based modelling techniques. Image reproduced with copyright permission from Rejniak and McCawley [2010].*

modelling techniques and some specific applications are given in the next section.

3.1 Individual-based models

Two main categories of individual-based models exist, on- and off-lattice models. Figure 3.1 illustrates the different techniques. The grid spaces of lattice-based models are chosen either to be the size of a cell or a cellular sub-compartment and can be either square or hexagonal. Cellular automaton models (see Figure 3.1(b)) were first proposed by John von Neumann and are probably the simplest individual-based modelling approach. In these models a grid space generally represents one cell and rules govern how the cells move between grid points. Due to its computational efficiency this modelling technique allows for the simulation of very large systems. In addition intracellular signalling pathways can be encoded in each cell and/or continuous descriptions of nutrient fields and other

environmental factors that influence the cells' behaviour can be added, leading to hybrid and multiscale models. Although relatively simple, this approach can provide models in which interesting collective behaviour can emerge at a tissue level by bringing together interacting components governed by certain rules. The draw-back is, however, that the cells are considered as points and thus the whole set-up is rather artificial and the results can often be directly linked to the rules employed or to the fact that important cellular characteristics (e.g. biomechanical properties) are ignored. Cellular automaton models have however been used to study cancer cell invasion and have, for example, been able to form hypotheses as to how the branching tumour morphology develops and how metabolic changes might lead to tumour invasion. A review of these models can be found in Hatzikirou et al. [2009].

Another lattice-based modelling technique is the cellular Potts model (see Figure 3.1(d)) which can be seen as an extension of the cellular automaton model. Here generally multiple lattice sites make up one cell and the overall idea is that the cell behaves in a way that lowers its effective energy. This energy takes into account the volume and shape of the cell, cell–cell and cell–environment interactions and any further elements encoded in the model. Simulations of a cellular Potts model are a series of attempts to extend the cells' boundaries. If the extension lowers the total effective energy it is accepted. However, if the extension increases the effective energy it is only accepted with a certain probability according to the Boltzman acceptance function. In contrast to the cellular automaton model, this modelling technique takes into account that cells are spatially extended physical objects rather than points and the behaviour of the cells is driven by the effective energy rather than rules alone. It is however very difficult to link the parameters used in the model to experimentally measured parameters and the fact that it is lattice-based means that, similar to the cellular automaton model, it is quite an

abstract modelling technique. Cellular Potts models of cancer cell invasion have been developed. With one of these models the dependency of invasion on cell–cell and cell–matrix adhesion has been investigated and it has been found that cell–cell adhesion is less influential on the invasive behaviour than cell–matrix adhesion and matrix proteolysis [Turner and Sherratt, 2002]. By extending the model to include proliferation the simulations have furthermore allowed the formation of the hypothesis that proliferation does not necessarily increase the depth of invasion but may, in some cases, reduce invasiveness. A detailed description of the cellular Potts model, its applications and also multiscale extensions can be found in Scianna and Preziosi [2013].

Lattice-free models are generally computationally far more expensive than lattice-based ones. They are, however, biologically more realistic and allow for the inclusion of cell mechanics. An example of a lattice-free model is the immersed boundary method. The IBCell model (see Figure 3.1(e)), an immersed boundary model of a cell, models the cell membrane as a network of linear elastic springs whereas the cell's interior, the cystoplasm, is modelled as an incompressible fluid [Rejniak, 2005, 2007]. Discrete receptors on the boundary of the cell allow the interaction between cells and also the cell and its environment [Rejniak, 2007]. This model has the advantage of being very detailed concerning the cell mechanics, especially the way the cells change shape. Due to the computational complexity, however, only a very few cells can be considered in a given simulation. The Subcellular Elements Model is another example of an off-lattice modelling technique [Newman, 2005]. In this technique each cell is represented by a number of subcellular elements. The elements interact via intra - and intercellular potentials. These interactions represent elastic responses to inter- and intracellular biomechanical forces and each element moves according to them. This allows the cells to have an adaptive shape, yet, depending on the number of subcellular

elements, being computationally slightly less expensive than the IBCell model. A third type of lattice free model has first been described in 2001 and is now widely used, for example in the simulation package CHASTE [Pitt-Francis et al., 2009]. This modelling technique uses the cell centres as points of reference and the individual cell centres are connected by linear, over-damped springs [Meineke et al., 2001]. A Voronoi tessellation of the space the cells are situated in gives each cell an individual shape. It is assumed that the cells aim to have an equal distance from one another which is enforced by the viscoelastic forces of the springs. Cell division unbalances this pattern and thus gives rise to cell migration. This modelling technique has been used to study the colonic crypt and the development of colorectal cancer [Mirams et al., 2012]. It is, however, unsuited to study individual cell migration and invasion as the shape and the movement of the cells depend solely on their interactions with neighbouring cells and it is thus not applicable to problems that study individual cells and cell-matrix interactions. Another off-lattice technique is the particle centre-based or force-based model (see Figure 3.1(c)) which is more flexible than the one described last, but computationally less expensive than the Subcellular Element Model and the IBCell model. It will be explained in more detail as this is the technique used in the modelling work in this thesis. For a more extensive review of individual-based modelling techniques see Anderson and Rejniak [2007].

3.2 Force-based models

The particle centre-based or force-based modelling technique for simulations of cell populations was originally developed by D. Drasdo. In these models cells are represented by a fixed spherical shape and are numerically referred to by their centre and radius. The movement of a cell takes place as a result of the forces acting upon it which is why this technique is also called force-based. Newton's

second law of motion is applied to derive the equation governing the cell's motion, i.e.

$$\mathbf{F} = m\mathbf{a} \tag{3.1}$$

where $\mathbf{F}$ is the net force acting on the body, m is its mass and $\mathbf{a}$ its acceleration. The law states that the acceleration of a body is proportional to the net force acting on it. Since the periods of time during which the cells start and stop their movement are very short in comparison to the time they spend moving with uniform velocity, these short phases of acceleration and deceleration shall not be considered here and thus:

$$\mathbf{F} = \mathbf{0}. \tag{3.2}$$

This equation implies that all the forces acting on a cell j must balance. In general, the forces acting on cells are drag forces $\mathbf{F}_{drag}$, cell–cell interaction forces $\mathbf{F}_{ij}$ for all neighbouring cells i, cell–matrix interaction forces $\mathbf{F}_{fj}$ for all neighbouring matrix fibres f and any chemotactic or haptotactic forces $\mathbf{F}_{taxis}$. In addition a noise term $\mathbf{f}_j$ can be added to include further smaller influences on cell motion that are not included in the forces mentioned above. Forces that can direct the cell's movement are positive if, when going back to the original form of Newton's law, they increase the acceleration whereas forces that decrease acceleration are negative. This leads to the equation:

$$\mathbf{f}_j + \mathbf{F}_{taxis} + \sum_{fnnj} \mathbf{F}_{fj} + \sum_{innj} \mathbf{F}_{ij} - \mathbf{F}_{drag} = 0, \tag{3.3}$$

or

$$\mathbf{F}_{drag} = \sum_{innj} \mathbf{F}_{ij} + \sum_{fnnj} \mathbf{F}_{fj} + \mathbf{F}_{taxis} + \mathbf{f}_j, \tag{3.4}$$

where the summation term $fnnj$ denotes all fibres f that are 'nearest neighbours' to cell j and similarly the term $innj$ denotes all cells i that are 'nearest neighbour' to cell j. While any chemotactic and haptotactic forces are forces which may be described in terms of a gradient of a prescribed external concentration field, the
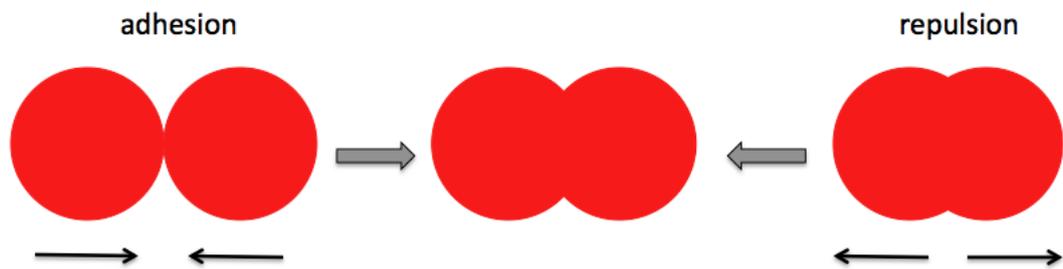
Figure 3.2: *Illustration of the cell–cell interactions, adhesion and repulsion. Initially when cells come into contact they can start adhering to each other via cell–cell adhesion molecules. This pulls them closer together and increases the cell–cell contact area which in turn increases the number of adhesive bonds and thus the adhesive force. However the cytoskeleton is not limitlessly deformable and thus repulsion sets in when cells come too close. A balance of these two forces leads to a stable position of the two cells relative to each other.*

other forces are slightly more complex. Cell–cell interactions exist in the form of adhesion and repulsion. Due to the adhesion receptors that cells express on their surfaces, they can start adhering to each other as soon as they come into contact. The adhesion force between the two cells increases, pulling the cells closer together as the contact area gets larger and the number of cell–cell bonds increases. But given that the cells are spheroidal in isolation, a large contact area stresses their cytoskeleton and their membranes. Furthermore it has been shown that cells only have a small compressibility [Byrne and Drasdo, 2009]. These two elements cause a repulsive force to arise between the two cells, should their contact area get too large. Thus the cell–cell interaction dynamics are those of adhesion and repulsion (see Figure 3.2). To capture this, the cell–cell interactions are modelled using potential functions which have been used in a number of different models [Galle et al., 2005, Drasdo and Hoehme, 2005, Ramis-Conde et al., 2008a,b, 2009, Macklin et al., 2012]. In this thesis we will use the

extended Hertz model to formulate the potential function as has been done in some previously published work [Galle et al., 2005, Ramis-Conde et al., 2008b, 2009]. The extended Hertz model calculates the cell's positive potential due to compression using the classic Hertz model [Hertz, 1881, Landau and Lifschitz, 1959]. Originally this model was developed for two elastic infinite semi-spheres and thus the diameter of the area of contact is assumed to be much smaller than the radius of the semi-spheres. Furthermore, it is assumed that the surfaces are continuous and non-conforming. Thus in order to be able to use this model as an approximation, cells have to be modelled as elastic spheres and the deformations have to be small as this ensures that most of the assumptions the Hertz model is based on hold. Adhesive interactions are governed by expression levels of cell–cell adhesion proteins such as E-cadherin (see Ramis-Conde et al. [2008b]). The potential V_{ij} between two cells i and j with radius R_i and R_j respectively is calculated as follows:

$$V_{ij} = \underbrace{(R_i + R_j - d_{ij})^{5/2} \frac{2}{5\tilde{E}_{ij}} \sqrt{\frac{R_i R_j}{R_i + R_j}}}_{\text{repulsive interaction}} + \underbrace{\epsilon_{ij}}_{\text{adhesive interaction}} . \qquad (3.5)$$

The first term on the right hand side is the repulsive interaction given by the Hertz model with

$$\tilde{E}_{ij} = \frac{3}{4} \left(\frac{1 - \sigma_i^2}{E_i} + \frac{1 - \sigma_j^2}{E_j} \right).$$

d_{ij} is the distance between the centres of the two cells, σ_i and σ_j are the Poisson ratios of the spheres and E_i and E_j are the elastic moduli. ϵ is the adhesive contribution which is, for example, determined by the number of E-cadherin-β-catenin bonds at the site of cell–cell contact and the energy that is released during bond formation. This term for the adhesive interaction is negative. Figure 3.3(a) shows a one-dimensional example of the shape of such a potential function. The biophysical properties of the cells determine the exact values of the potential depending on the distance between the cells. The force $\mathbf{F}_{ij}$ acting on the cell

36
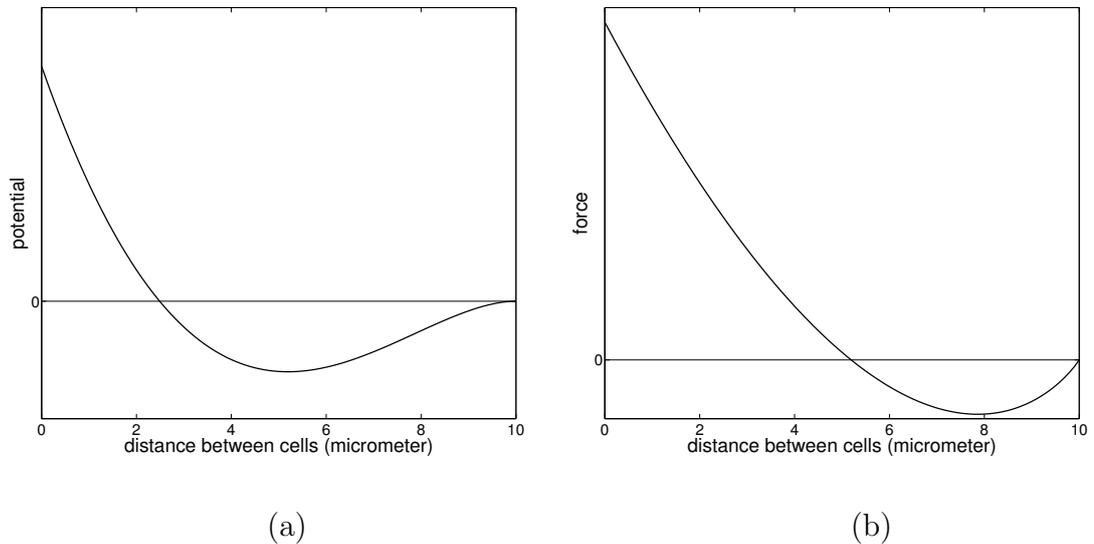
Figure 3.3: *Figures showing a one-dimensional example of the potential and force between two cells of radius 5 μm. (a) Plot of the potential calculated using the extended Hertz model. (b) Plot of the force calculated as the negative derivative of the potential shown in (a).*

due to interactions with its neighbours is given by the negative derivative of the potential.

$$\mathbf{F}_{ij} = -(d(V_{ij})/d(d_{ij}))(\partial d_{ij}/\partial x, \partial d_{ij}/\partial y, \partial d_{ij}/\partial z) \qquad (3.6)$$

Figure 3.3(b) shows the dependence of this force on the distance between the cells. As can be seen in this figure, the cell seeks to lower its potential and thus the stress it is under and therefore is forced to move away from the other cell (positive force) when the repulsion is higher than the adhesion and vice versa.

3.3 The model by Ramis-Conde et al.

Ramis-Conde et al. developed a multiscale force-based model of cancer cell invasion by using the technique explained above and adding intracellular E-cadherin and β-catenin dynamics to the model [Ramis-Conde et al., 2008b]. For simplicity,

it is assumed in the model that E-cadherin and β-catenin bind at the cell membrane. Therefore three different states exist for the E-cadherin molecules: free E-cadherin in the cytoplasm ($[E_c]$), free E-cadherin at the membrane ($[E_m]$) and E-cadherin in complex with beta-catenin at the membrane ($[E/\beta]$). Upon cell–cell contact, E-cadherin moves to the cell surface. The amount that is translocated is proportional to the contact area. Once at the cell membrane, E-cadherin can bind β-catenin and can then form bonds with the neighbouring cells. If the cells detach again, the complexes are dissociated and its components are free in the cytpoplasm. It is assumed that E-cadherin can be recycled and thus the total E-cadherin concentration does not change over time ($E_T = [E_c] + [E_m] + [E/\beta]$). β-catenin however, can be degraded after forming a complex with the proteasome. The Wnt pathway is taken into account here by assuming that if Wnt is activated, no degradation takes place. Furthermore β-catenin is also produced at a constant rate.

These dynamics are modelled using ordinary differential equations governing the concentrations of $[E_c]$, $[E/\beta]$, β-catenin and the proteasome–β-catenin complex. For the equations of this model and more detail see Ramis-Conde et al. [2008b]. In the model cells in isolation are assumed to be spherical with a radius R. When a cell divides, the two daughter cells have a radius $R/2^{\frac{1}{3}}$ for volume conservation. They then grow up to radius R before they can divide again. Cell–cell interactions are governed by equations (3.5) and (3.6). The adhesive contribution to the extended Hertz model is controlled by the intracellular dynamics and is calculated by $\epsilon_s = \varrho_m A_{ij} W_s$ where W_s is the energy of a single bond, A_{ij} is the contact area between two cells and ϱ_m is the density of $[E/\beta]$ in the contact area.

The movement of the cells is then given by an equation of the same type as (3.4).

$$\underbrace{\underline{\underline{\Gamma}}_{is}^f \underline{\nu}_i}_{c-sfriction} + \sum_{jnni} \underbrace{\underline{\underline{\Gamma}}_{ij}^f (\underline{\nu}_i - \underline{\nu}_j)}_{cell--cellfriction} = \sum_{jnni} \underbrace{\underline{F}_{ij}}_{forces} + \underbrace{\underline{f}_i(t)}_{noise} + \underbrace{\chi \nabla Q(t)}_{chemotaxis} \qquad (3.7)$$

Simulations using this model showed the increase in nuclear β-catenin concentration at the outer rim of a tumour and how this leads to invasion of the surrounding tissue. For a detailed explanation of the equation terms, parameter values and results, see Ramis-Conde et al. [2008b].

3.4   The modelling framework used in this thesis

In this thesis we will extend the work by Ramis-Conde et al. [2008b] that has been described above. In addition to the cell-cell interactions in the model we will consider explicit interactions between individual cells and individual matrix fibres. As the review of cell migration and invasion models in the above section shows, this is a process that has not been considered before. Therefore we will first model the interactions of a single cell with individual matrix fibres before moving on to studying two cells on a sheet of extracellular matrix and then finally including these cell-matrix interactions in the cell-cell interactions model by Ramis-Conde et al. [2008b]. We will also take a closer look at the way the cell-cell interactions are modelled and coupled to the intracellular dynamics and will improve this part of the model by Ramis-Conde et al. [2008b]. This will allow us to investigate the role of cell-cell adhesion proteins and pathway dynamics together with the role of cell-matrix interactions in invasive cell populations.

Chapter 4

An Individual-Based Model of *in vitro* Single Cell Migration on a 2D Matrix

4.1   Introduction

In this chapter we develop an individual-based model of single cell migration on two-dimensional substrates which is driven by forces acting upon the cell. In addition individual matrix fibres are reoriented due to cell traction forces. Though focussing only on the most fundamental processes that drive cell migration, the model allows for running multiple computational 'experiments' by varying cell and matrix characteristics. After describing the model and the computational simulation algorithm, we will explain how the simulation data, which is a time series of cell positions, can be analysed to produce quantitative measures that can be compared to experimental data. These measures, namely persistence time and cell speed, will also be compared for different matrix stiffnesses and architectures, fibre lengths and fibre densities as well as different maximum cell speeds.

Furthermore we will show that this model can reproduce the experimental results that cells prefer stiff to soft matrices. Though biological research has focused on the coupling of biomechanics with cell signalling to explain these observations, our findings lead to the hypothesis that the simple physical interactions between cells and matrix fibres and the reorientation of fibres, or lack thereof, captured in this model might be sufficient to produce this behaviour.

4.2   Model description

The basis of the model is an individual cell which moves according to the forces acting upon it. The modelling approach used is a force-based approach similar to the one used to study epithelial cell populations [Galle et al., 2005], cancer cell invasion [Ramis-Conde et al., 2008b], the process of intravasation [Ramis-Conde et al., 2009] and also cell migration in three-dimensional matrices without the explicit inclusion of matrix elements [Zaman et al., 2005]. In contrast to these previous models, the forces considered here are those that a cell generates through interactions with individual matrix fibres to pull itself forward on a layer of extracellular matrix.

4.2.1   Modelling the cell and the extracellular matrix

The shape of an individual cell is relatively flat and hemispherical and we assume that its base has a radius of 15μm and it has a height of 2.6μm as has been measured for Madin-Darby Canine Kidney (MDCK) cells by Schneider et al. [2000]. We explicitly model individual matrix fibres, which could represent fibronectin, collagen, laminin or other fibrous matrix components. These fibres are represented by thin cylinders, the lengths of which are normally distributed

41

with mean 75μm and standard deviation 5μm and the widths are 200nm (as measured for collagen in Friedl et al. [1997]). We consider a spatial domain of size $1000\times1000\mu$m in which we place 15,000 fibres, where one of the end-points of each fibre is randomly positioned following a uniform distribution in space. For isotropic matrices, the direction of each fibre is given by a normalised vector with uniformly distributed x and y components (see Figure 4.1(a)). Ordered matrices are also generated where the matrix fibres are either biased in the sense that they form an angle between 90° and 180° with the x-axis or are fully aligned so the direction of the fibres forms a 135° angle with the x-axis (see Figure 4.1(b), (c)).

(a)   (b)   (c)
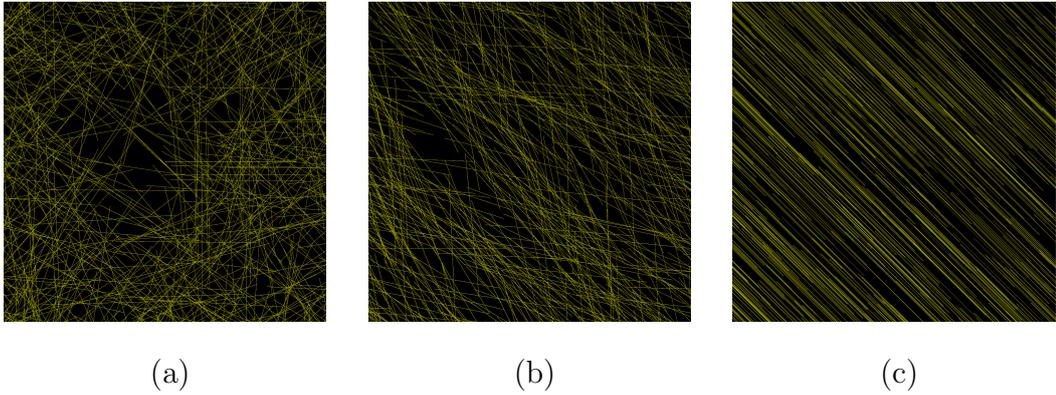
Figure 4.1: *Figure showing the computer generated initial extracellular matrices with (a) randomly distributed fibres (b) biased fibres (c) aligned fibres. The figures show a representative $150\times150\mu$m square sub-domain taken from the entire domain of size $1000\times1000\mu$m.*

4.2.2   Modelling the cell movement

Cell movement is governed by the total force acting on an individual cell. As explained above, by calculating all the forces acting on a cell and then applying Newton's Second Law of Motion, an equation for the cell velocity is obtained (see equation (3.4)). Hence, by integrating this equation the displacement of an

individual cell over time can then be calculated.

The system we are modelling consists of an individual cell interacting with individual matrix fibres and so the forces on the cell consist of a drag force which is balanced by the overall force generated by an individual cell through contact with the matrix fibres and a term accounting for underlying "noise". The system considered is an *in vitro* set up where chemotactic and haptotactic gradients can be controlled and are non-existent at this instance. Thus taxis terms are not included. Therefore the governing equation of motion has the general form:

$$\mathbf{F}_{drag} = \sum_{fnnj} \mathbf{F}_{fj} + \mathbf{f}_j(t), \tag{4.1}$$

where $\mathbf{F}_{fj}$ is the force generated by an individual cell through contact with an individual matrix fibre, with the sum taken over the fibres which are in contact with the cell, and $\mathbf{f}_j(t)$ is the term accounting for "noise". These terms are described in detail below.

We assume that as in *in vitro* set-ups, the layer of matrix fibres and the cells migrating on the matrix are embedded in a gel-like suspension. Thus the drag force that acts on the cell can be modelled using Stokes' Law (cf. Zaman et al. [2005]):

$$\mathbf{F}_{drag} = c\eta\mathbf{v}_j, \tag{4.2}$$

where c is the shape factor which is $6\pi r$, with r being the radius of a spherical object, η the fluid viscosity and $\mathbf{v_j}$ the velocity of cell j. However, here it is important to remember that cells moving over a 2D substrate are not spherical, whereas Stokes' law is defined for spherical objects. We therefore use a variation of Stokes' law for nonspherical objects as developed in Leith [1987] and Payne and Pell [1959] by assuming that the cell has a symmetric, almost hemispherical shape with flat extension around the periphery. With this we can simplify the

variation down to changing only the shape factor to $c = 16.7 \times d$, where d is the radius of the frontal area circle and 16.7 is chosen to be between the estimated factors for a hemispherical cup and a flat disk given in Payne and Pell [1959]. In order to calculate the radius of the circle with the same area as the area of a slice of the cell perpendicular to its velocity, its height and width have to be known. These have been measured for migrating MDCK cells by Schneider et al. [2000]. From these measurements we know that the fontal area of a migrating MDCK cell is approximately the area of half an ellipse with minor radius 2.6μm and major radius 15μm and thus $A_{\frac{1}{2}Ellipse} = \frac{1}{2}\pi \times 2.6\mu\text{m} \times 15\mu\text{m}$. Therefore the radius of the circle with the same area as the area of the biggest slice of the cell perpendicular to its velocity is $\left(\frac{1}{2} \times 2.6 \times 15\right)^{1/2} \mu$m for these measures. Or more general $\left(\frac{1}{2}ab\right)^{1/2}$ where a and b are the minor and major axes of the ellipse, respectively. This leads to the following shape factor

$$c = 16.7\left(\frac{1}{2}ab\right)^{1/2}, \tag{4.3}$$

where a and b are the minor and major axes of the ellipse that is given by a thin slice of the cell at its highest point. For the fluid viscosity η we assume a value of 10^2 poise which is one order of magnitude lower than the viscosity of the three-dimensional matrix used in Zaman et al. [2005]. This reflects the fact that the cells considered here are embedded within a gel rather than a 3D matrix which has many components that the cell interacts with which effectively increases the viscosity.

Other factors can influence the migration of cells. Even in an *in vitro* setting where no growth factors are added or chemical gradients set up, small fluctuations in concentrations or small impurities can occur that will have a slight influence on the cell's behaviour. Therefore we add a noise term $\mathbf{f}_j(t)$, which is uncorrelated and has zero mean, to the equation of motion.

We now describe the calculation of the force $\sum\limits_{fnnj} \mathbf{F}_{fj}$ generated by an individual

cell through contact with the nearby matrix fibres. In order to migrate a cell needs to establish front-rear polarity and form focal contacts with the matrix. The process of focal complex formation and its stabilisation to a focal contact, which enables the cell to obtain the traction needed to move along the fibre, takes place in the order of minutes [Friedl and Wolf, 2003]. In the model we assume a time of 10 minutes for this process. In these single cell migration simulations this process only occurs at the very start when an unpolarised cell is placed on the matrix. In simulations with more cells in later chapters, a cell can lose its polarity when it contacts another cell and cannot extend any more protrusions in that direction (contact inhibition of locomotion). In this case the cell will have to reestablish its front-rear polarity. Once a cell is polarised, the extracellular matrix influences the cell's movement through contact guidance. The distance between the cell's centre and the fibre determines whether a cell is in contact with a fibre or not. If the distance is equal to or smaller than the cell's radius then the fibre and the cell are considered to be in contact.

A cell can move along a fibre in two directions. We assume, as Dallon et al. [1999], that a cell contacting a fibre will chose the direction of the fibre that requires the cell to make the smallest change in direction. The same is done when a cell is in contact with more than one fibre. We again choose the orientation of each fibre that would require the smallest directional change of the cell. Then a projection of the cell's previous polarity axis onto each of these orientated fibres is made (see Figure 4.4(c)). Adding up the projections and normalising the resulting vector, gives the cell's new polarity axis, $\mathbf{p}_j^{new}(t)$ for cell j as:

$$\mathbf{p}_j^{new}(t) = \frac{1}{\|\mathbf{p}_j^{new}(t)\|} \sum_i \frac{\mathbf{p}_j^{old}(t).\mathbf{w}_i}{|\mathbf{w}_i|^2} \mathbf{w}_i, \qquad (4.4)$$

where $\mathbf{w}_i$ is the direction of fibre i.

The projections onto the fibre directions weight the 'input stimuli' that the cell encounters so that more weight is given to fibres closely aligned with the cell's polarity axis. The reason for this procedure is that we assume a cell's movement to be biased towards continuing moving in the same direction, because all the integrins and other pro-migratory proteins are already clustered here. On the other hand, a change in direction would, in most cases, be due to a strong stimulus that causes the intracellular machinery to rearrange itself.

Note that equation (4.4) gives a deterministic description of the cell's polarity axis. The stimulus that the cell encounters can however be stronger in one direction, not just due to the number of fibres, but also due to local fluctuations in the fibronectin and integrin distributions. We incorporate this by adding noise to the procedure explained above through the multiplication of the projection onto each fibre by $(1+\chi)$, where χ is a Gaussian noise term with zero mean and a standard deviation of one. This gives a higher stochasticity to the influence of the fibres more closely aligned with the polarity axis than to those that are at a greater angle, i.e.

$$\mathbf{p}_j^{new}(t) = \frac{1}{\|\mathbf{p}_j^{new}(t)\|} \sum_i \frac{\mathbf{p}_j^{old}(t).\mathbf{w}_i}{|\mathbf{w}_i|^2} \mathbf{w}_i (1 + \chi_{ji}(t)). \tag{4.5}$$

For cell j the net force generated through contact with the nearby matrix fibres f is therefore given by:

$$\sum_{fnnj} \mathbf{F}_{fj} = M\mathbf{p}_j^{new}(t), \tag{4.6}$$

where M is the magnitude of the force, the calculation of which we will explain below.

We assume that for small time intervals $t \in (t_i, t_i + \delta t)$, the cell maintains on average a uniform movement following the direction determined by its interaction with nearby fibres. This is modelled by the use of a constant force term $\mathbf{K}$ derived from the contribution of each pseudopodium attached to a fibre i.e.

$$\sum_{fnnj} \mathbf{F}_{fj} = \mathbf{K}_{t\in(t_i,t_i+\delta t)}. \tag{4.7}$$

Hence, the equation of motion is now:

$$\mathbf{F}_{drag} = \mathbf{K}_{t\in(t_i,t_i+\delta t)} + \mathbf{f}_j(t). \tag{4.8}$$

Since $\mathbf{f}_j(t)$ has zero mean, by assuming the expected value, the force generated by a cell through interaction with the matrix fibres is proportional to the cell's velocity:

$$\mathbf{F}_{drag} = c\eta\mathbf{v}_j = \mathbf{K}_{t\in(t_i,t_i+\delta t)}. \tag{4.9}$$

The direction of the cell's movement is given by the polarity axis and thus:

$$\mathbf{v}_j = |\mathbf{v}_j| \times \mathbf{p}_j^{new}(t). \tag{4.10}$$

The direction of the net force is also along the polarity axis. Thus for $t \in (t_i, t_i + \delta t)$,

$$\mathbf{K}_t = M\mathbf{p}_j^{new}(t). \tag{4.11}$$

From equations (4.9), (4.10) and (4.11) it can be seen that the magnitude of the force, M, is proportional to the cell's speed, i.e.

$$c\eta \times |\mathbf{v}_j| = M. \tag{4.12}$$

Therefore experimental measurements of the speed can be used to calculate the magnitude of the force.

The speed at which a cell moves through extracellular matrix follows a biphasic behaviour and depends on integrin and ligand concentrations and the precise distribution of integrins over the cell surface [DiMilla et al., 1991, Palecek et al., 1997, Zaman et al., 2005]. Apart from possible small local differences, we assume that the ligand density is the same on all fibres and that 95% of expressed integrins are at the front of the cell and therefore the speed is solely dependent on

the integrin expression levels and the number of fibres a cell is in contact with. We further assume in all simulations an integrin expression level of 50% which leads to the maximum cell speed in respect to this parameter [Palecek et al., 1997]. In two dimensions the maximum speed a cell can reach on fibronectin is ≈ 20 μm/h which happens at approximately half the maximum adhesion force [Palecek et al., 1997]. In the simulations the effect of varying the value of this maximum cell speed is however also investigated. We assume that the maximum cell–matrix adhesion is reached when a cell is in contact with 120 fibres as this would cover 80% of the cell–matrix contact area if the fibres were aligned.

Using this and the experimental measurements of cell speed dependent on adhesion force, it is known how fast a cell should be moving given the number of fibres it is in contact with. Thus the magnitude M of the net force that a cell has to generate to pull itself forward at this speed is calculated from equation (4.12). The plots in Figure 4.2 show the relationship between the number of fibres that a cell is in contact with and its speed and the relationship between the speed and the force.
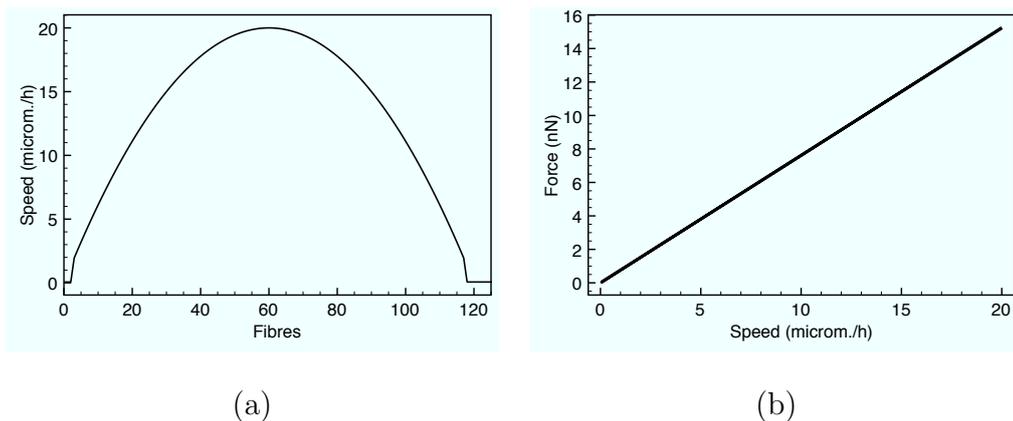

(a) (b)

Figure 4.2: *Plots showing the relationship between (a) the number of fibres a cells is in contact with and it's speed and (b) the speed and the force.*

4.2.3 Modelling the matrix rearrangement due to cell traction forces

When a cell contracts during migration, traction forces are transmitted to the substrate through the adhesion complexes and the pulling on the fibres realigns the matrix [Friedl and Wolf, 2009, Lauffenburger and Horwitz, 1996] (see Figure 2.7). These traction forces of a cell point towards its centre [Lemmon et al., 2009, du Roure et al., 2005] and thus the fibre is pulled inward. We model the fibre as a lever that is rotated about the axis of its moment of force. With this assumption, the end of the fibre that is furthest away from the cell acts as the fulcrum. We assume that the realignment of the fibre is proportional to the integrin expression of the cell and it is also proportional to (1-S) where S is the matrix stiffness which is a non-dimensional value between zero and one. The realignment decreases the closer the fibre is to the cell's midpoint as the traction forces decrease from the cell's periphery inwards [Lemmon et al., 2009, du Roure et al., 2005]. Hence, the angle of rotation ϕ of a fibre is given by:

$$\phi = \Theta - \arcsin\left(\frac{(1 - 0.1 \times I \cdot (1 - S)) \cdot D}{d}\right), \qquad (4.13)$$

where Θ is the current angle between the straight line connecting the fulcrum and the cell's midpoint and the fibre, I is the percentage of integrins expressed by the cell, S is the matrix stiffness, D is the shortest distance between the fibre and the cell and d is the distance of the fulcrum from the cell's midpoint. The other parameter used, the factor 0.1, was estimated to give an appropriate reduction of the reorientation per time step. The influence of a 10% and 20% change in this parameter is however investigated in Section 4.6.2.

The change in ϕ for different matrix stiffnesses can be seen in Figure 4.3(a). In this figure the initial distance between the cell and the fibre (D) is taken to be 15μm which is the radius of the cell, the distance between the cell and the fulcrum

(*d*) to be 50μm (which gives a Θ value of 17.1887 degrees or 0.3047 radians) and the integrin expression (I) to be 0.5. The stiffness S varies between zero and one. We follow the development over 5 consecutive simulation time steps. The change of the angle between the fibre and the line connecting the fulcrum and the cell's centre (Θ) over these time steps can be seen in Figure 4.3(b).

(a)   (b)
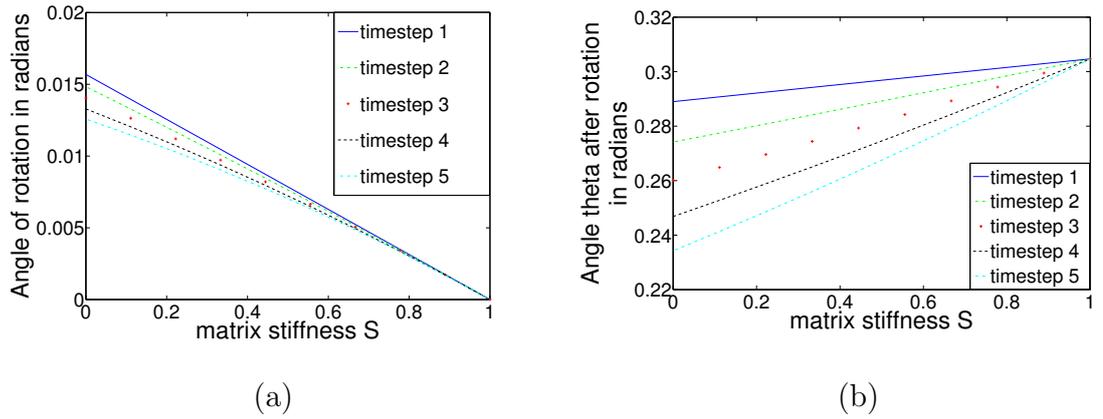
Figure 4.3: *Graphs showing the effects of matrix stiffness on the realignment of the fibre by equation (4.13) over five consecutive time steps. (a) Graphs showing the angle of rotation ϕ depending on matrix stiffness. (b) Graphs showing the angle between the straight line connecting the fulcrum and the cell's midpoint and the fibre (Θ) after the rotation.*

The matrix stiffness S can either be a constant value throughout the domain or, more realistically, we can calculate it for each fibre depending on the number of fibres it has cross-links with. For less than 15 cross-links we assume a matrix stiffness of the number of cross-links $\times$ 0.06. For more than 15 cross-links, the fibre is assigned a stiffness of 0.95. This maximum of 15 cross-links was chosen under the consideration of the number of cross-links the fibres generally have - only a fraction of fibres has a higher number of intersections with other fibres. However, the effect of a 10% and 20% change in this parameter is investigated in Section 4.6.2.

4.3 Computational simulation algorithm

The code used for the computational simulations of the model is implemented in C++. The Mersenne Twister pseudorandom number generator is used for the generation of random numbers in the simulations [Matsumoto and Nishimura, 1998]. The length of a time step has been chosen to ensure that the distance a cell can travel during one time step is one magnitude smaller than 200nm which is the width of a fibre and the smallest unit considered. Thus the length of a time step is fixed at 3 seconds (which allows a maximum movement of $\approx 0.017\mu$m per time step for a cell that reaches a maximum speed of 20μm/h).

The procedure during each time step can be summarised as follows:

Step 1:

Each fibre is examined to see whether the cell was in contact with it during the last time step and whether the cell is polarised. If both of these conditions are met then it means that the cell has exerted force on it during that time step. All the fibres for which this is the case are reorientated in the way explained in equation (4.13).

Step 2:

All the fibres that the cell is in contact with are found. If it is in contact with at least one fibre and has established front rear polarity either through previous polarisation or through the new formation of focal contacts over 10 minutes, the new polarity axis is calculated using equation (4.5).

Step 3:

If the cell is polarised, the net force that it generates for its movement on the matrix is calculated, using the knowledge from the previous step of how many fibres the cell is in contact with. In order to do so, the speed that the cell should be moving at is first calculated using the given parameter determining the maximum speed the cell can reach and the number of fibres the cell is in contact with

(see image 4.2(a)). From this the magnitude of the force the cell must generate is derived (see Figure 4.2(b)). Equation (4.6) is then used to calculate the net force.

Step 4:

The cell is moved according to the forces calculated in step 3. This is done by first solving equation (4.1) for the cell velocity and then applying the Forward Euler method. This gives the new position of the cell at the end of this time step.

This procedure is illustrated schematically in Figure 4.4

4.4 Data analysis

Experimentalists who study single cell track data and the influence of intra- and extracellular components on cell migration generally use two measures to analyse and compare their results, namely persistence time and root-mean-squared cell speed [Harms et al., 2005, Gail and Boone, 1970, Bergman and Zygourakis, 1999, Stokes and Lauffenburger, 1991]. Together these give a complete description of the cell movement over a certain length of time. Thus we will use the same measures here to quantify the results from our computational simulations.

Cell speed determines the total length of the cell path during the time interval of interest whereas persistence time can be considered as the amount of time the cell spends moving without changing its direction by more than $60°$ [Ware et al., 1998]. Both, speed and persistence time, are related to the displacement of the cell during a known time interval which can be calculated from the positions of the cell at either end of that interval. In experiments these positions can be determined from images taken at the beginning and the end of that time interval by using image processing tools to find the x and y coordinates of the cell centroid
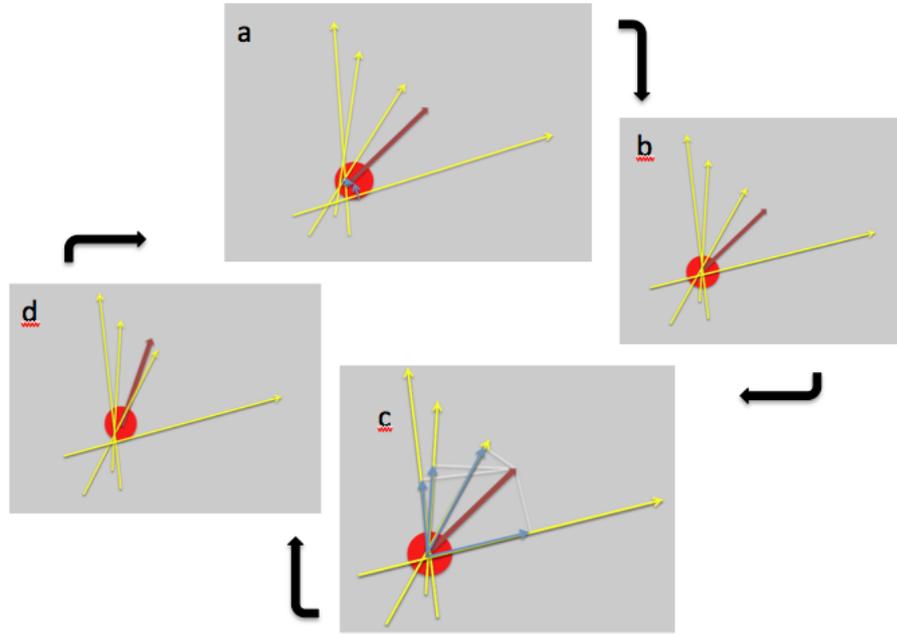
Figure 4.4: *Schematic diagram showing the re-orientation of fibres by an individual cell and the subsequent calculation of the polarity axis and cell migration in this new direction. (a) Re-orientation of the fibres by forces exerted by the cell (forces shown by blue arrows). (b) Choice of fibre orientation with the polarity axis of the cell being indicated by the dark red arrow. (c) Calculation of the new polarity axis by adding up the projections of the old polarity axis over all the fibres (blue arrows). (d) Movement of the cell along the new polarity axis according to the forces calculated.*

[Harms et al., 2005]. Producing the same kind of data from the simulations is achieved by writing the coordinates of the cell position to an output file at the time points of interest. Just as in experiments [Harms et al., 2005, Ware et al., 1998], the mean-squared displacements are calculated from this cell track data of the simulations using the method of non-overlapping intervals as described by Dickinson and Tranquillo [1993]. Given N consecutive positions of the cell with constant time interval Δt, following this method, the mean-squared displacement
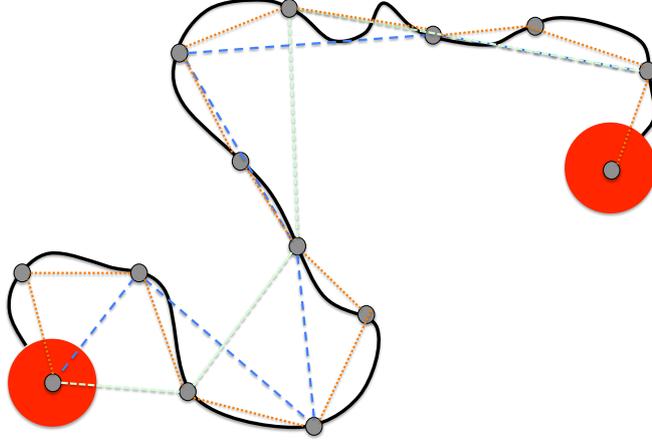
Figure 4.5: *Schematic diagram illustrating the method of non-overlapping intervals for obtaining mean-squared displacements for i=1,2,3 from cell tracking data. The two red circles show the cell at the first and the Nth (here 14th) position. The small grey circles show the other given positions of the cell with constant time interval Δt and the black line is the actual cell path. The dotted orange lines show the displacements $\delta_{(1+i(j-1))\Delta t \to (1+ij)\Delta t}$ for $i = 1$ and $j = [1, ..., 13]$, the blue dashed lines show the displacements for $i = 2$ and $j = [1, .., 6]$ and the green dotted lines show the displacements for $i = 3$ and $j = [1, .., 4]$.*

during a time interval of length $t_i=i\Delta t$ is calculated by:

$$< \delta(t_i)^2 >= \frac{1}{n_i} \sum_{j=1}^{n_i} [\delta_{(1+i(j-1))\Delta t \to (1+ij)\Delta t}]^2 \qquad (4.14)$$

where $i = [1, 2, ..., N-1]$ and $n_i = \lfloor \frac{N-1}{i} \rfloor$.

Figure 4.5 illustrates the term $\delta_{(1+i(j-1))\Delta t \to (1+ij)\Delta t}$ for i=1,2 and 3. It is determined in the same way for $i = [4, ..., N-1]$ to give an $N-1$ dimensional vector of displacement values. The vector of mean-squared displacements $< \delta(t_i)^2 >$ is derived from this as stated in equation (4.14).

In order to calculate the root-mean squared cell speed s from these measurements, the root mean-squared displacement for the case $i = 1$ is divided by the tracking interval Δt. The cell's persistence time is then fitted to the persistent random walk model [Othmer et al., 1988, Alt, 1990, Dickinson and Tranquillo, 1993] by substituting its speed into the relationship:

$$< \delta(t_i)^2 > = 2s^2 P[t - P(1 - e^{-t/P})], \tag{4.15}$$

where s is the cell speed, P is the persistence time and $< \delta(t_i)^2 >$ the mean-squared displacement. The fitting is done using an unconstrained nonlinear optimisation routine. For this the inbuilt MATLAB (MATLAB R2010a, The Math-Works, Natick, MA) function '*fminsearch*' is used which itself uses the Nelder-Mead Simplex algorithm to find a parameter or a set of parameters that minimises the objective function value. For a function of an n-dimensional variable, this algorithm creates a simplex of $n + 1$ vertices around the initial guess x_0. It does this by adding 5% to each component $x_0(i)$ and using these n vectors together with the original x_0. Then it continually either discards the point of the simplex with the worst function value and replaces it by a new one which is generated by reflection and possibly also expansion or contraction of this point or, under certain conditions, it discards all points and calculates new vertices by shrinking the old simplex [Lagarias et al., 1998]. As an initial guess x_0 we use a randomly generated value for P. The function to minimise in this case is the $N - 1$ dimensional vector of the root-mean squared difference between the 'measured' root-mean-squared displacements and the ones calculated from formula (4.15) using the actual cell speed and the estimated persistence time. This is done component-wise, by fitting the persistence time.

In the present model cell migration is studied in an *in vitro* situation where no external chemical gradients and no growth factors etc. exist. In most published

migration experiments, the cells are however kept in serum which most likely influences the migratory behaviour of the cells. Thus a specific set of experiments by Harms et al. [2005] has been chosen to compare the simulation results to where the cells were serum-starved and re-suspended in serum-free medium before the migration assays were done. During the migration assays the cells were tracked for six hours and imaged every 15 min. In order to be able to compare the simulation data to these experiments, we also use a time interval Δt of 15 minutes and a total cell tracking time of six hours to calculate the persistence time. For most of the computational 'experiments' 15 simulations are run with 15 different seeds for the random number generator used in the noise terms of the cell movement. Thus the simulation results are 15 different sets of cell tracking data for each 'experiment'. The actual cell speed and persistence time are determined for each of these 15 cell tracks. As outliers occur in the resulting persistence times for certain sets of 'experiments' and the distribution of the results is partly very asymmetric, box plots are used to visualise the data. They are done with R (version 2.13.1, [R Core Team, 2012]). The default value of 1.5 is used for the range, meaning that the whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box. Thus any data points beyond that value are classified as outliers. For the cell speeds the mean and the standard deviation give a good description of the data and are calculated in all cases. This is visualised in plots using MATLAB. Whenever it can be of additional interest to look at an image of the cell track or the underlying matrix fibres, the images are produced with OpenGL.

4.5 Computational simulation results

In these single cell migration simulations the added noise term was initially ignored. At the beginning of all simulations a non-polarised cell was placed in the

56

middle of a 1000μm$\times1000\mu$m area of extracellular matrix. The cell was then left to polarise and start migrating over a time of 3 hours and was then tracked over 6 hours of real time to make it comparable to experiments which were run for the same length of time [Harms et al., 2005]. All parameter values used in the simulations are given in Table 4.1 along with the reference from where they were obtained.

parameter	value	reference
Radius of a cell base (R)	15μm	[Schneider et al., 2000]
Height of a cell	2.6μm	[Schneider et al., 2000]
Matrix fibre diameter	200nm	[Friedl et al., 1997]
Suspension viscosity (η)	10^2 Poise	derived from [Zaman et al., 2005]
Maximum cell speed in 2D	20μm/h	[Palecek et al., 1997]
Focal complex formation time	10 min	[Friedl and Wolf, 2003]

Table 4.1: *Table detailing the parameter values used in the computational simulations.*

4.5.1 The influence of matrix stiffness on persistence and migration speed

First we investigated the effect of matrix stiffness on cell migration. 15 simulations were run for each of the following cases where the matrix composition was varied: (i) a very loose matrix (S=0), (ii) a medium stiff matrix (S=0.5) and (iii) a very stiff matrix that cannot be reorientated (S=1). In addition 15 simulations were run in which the stiffness was calculated individually for each fibre as explained above. This was done for four different matrix architectures by seeding the random number generator which is used to place the matrix fibres with four different numbers. The random number generator used for the noise terms in the cell movement was given the same 15 seeds in the four studies. The results are
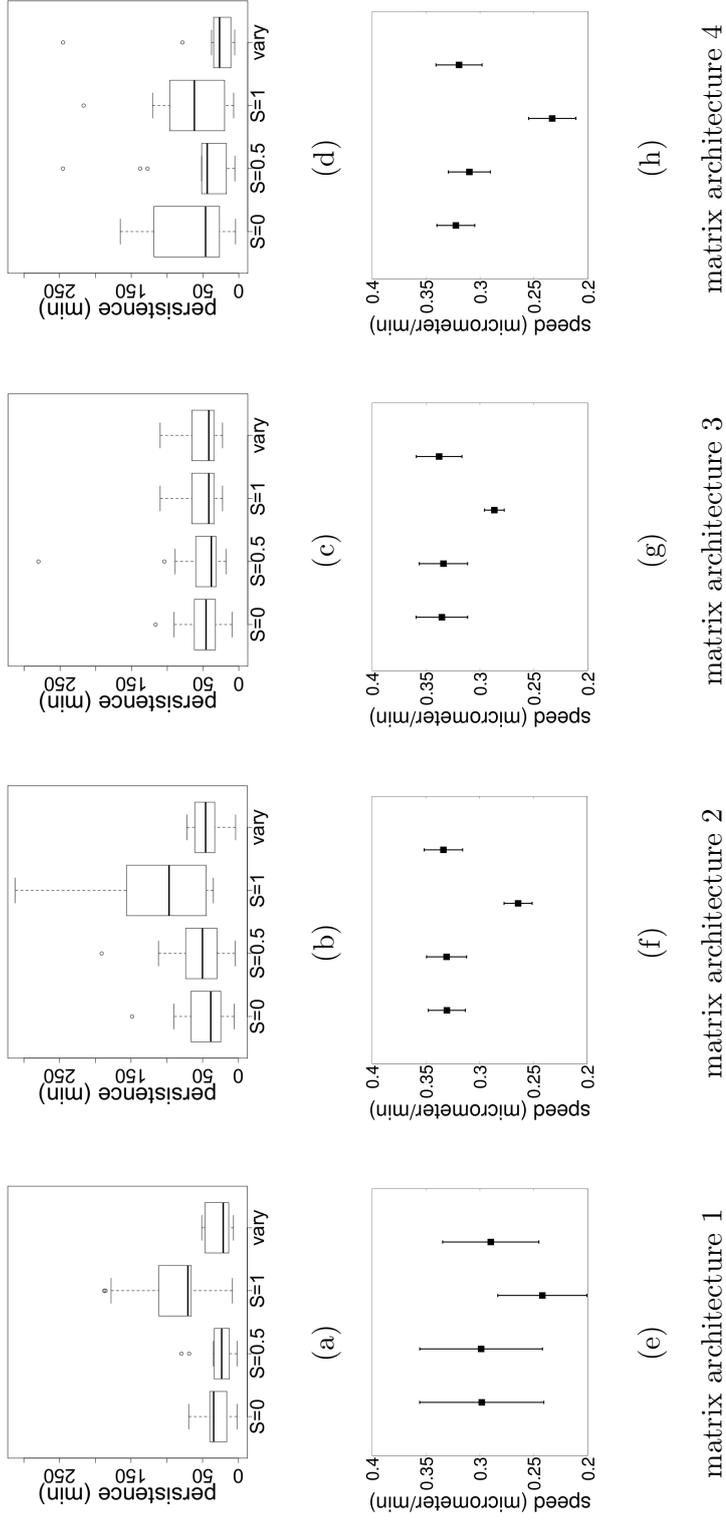
Figure 4.6: *Plots showing how the persistence time and mean actual speed of cell migration vary with different matrix stiffnesses in the 4 different random matrix architectures. Each matrix architecture is produced by using a different seed for the random number generator that gives rise to the position and orientation of the matrix fibres. (a)-(d) Plots showing the persistence times for cells on loose (S=0), medium stiff (S=0.5) and stiff (S=1) matrices as well as for matrices where the stiffness depends on matrix interconnectedness at a given point. In plot (d) two outliers exist outside of the y-axis range for S=0. They are at 367.24 min and 430.87 min. Each box plots represent the results of 15 simulations. (e)-(h) Plots showing the actual mean speed of the cells during the simulations which led to the box plots above.*

given in Figure 4.6 with outliers denoted by small circles.

It can be seen from Figure 4.6 that the persistence times on a stiff matrix have the biggest variation over the four different sets of simulations which means that the tracks of cells on a stiff matrix are the most dependent on the matrix architecture. This is not unexpected. The least variation can be seen in the simulations in which the stiffness of the matrix is calculated for each fibre independently. Here persistence times between 3 and 109 minutes (apart from one outlier which has a persistence time of 245 minutes) arise with the majority, the lower and upper quartiles, being between 20 and 60 minutes. Experiments by Harms et al. [2005] measured a persistence time of roughly 8 to 20 minutes in Chinese hamster ovary (CHO) cells on fibronectin unstimulated by epidermal growth factor (EGF) and a persistence time of about 19 to 50 minutes in cells stimulated by EGF [cf. Figure 2D, pg. 1483, Harms et al. [2005]]. Thus although the simulations predict a slightly higher persistence time than is observed in unstimulated cells, the values are of the correct order of magnitude. It is clear from the results, however, that the reorientation of the matrix fibres is crucial for this. Similar values as mentioned above are also found for a matrix stiffness of 0.5 and a very loose matrix, but a stiff matrix that allows no reorientation gives much more variation and much higher persistence times than observed in the experiments.

A more detailed study of the persistence time and mean actual speed of migration on matrices of varying stiffness S between 0.5 and 1 is presented in Figure 4.7. For these simulations the first matrix architecture from previous simulations was considered. The plots in Figure 4.7(a) and Figure 4.7(b) show that both persistence time and mean speed are very similar between 0.5 and 0.9, but there appears to be a transition to higher persistence times and lower cell speeds between 0.9 and 1. Interestingly, in all simulations it was found that cells on a very stiff matrix move significantly more slowly than cells on a matrix that allows reorientation
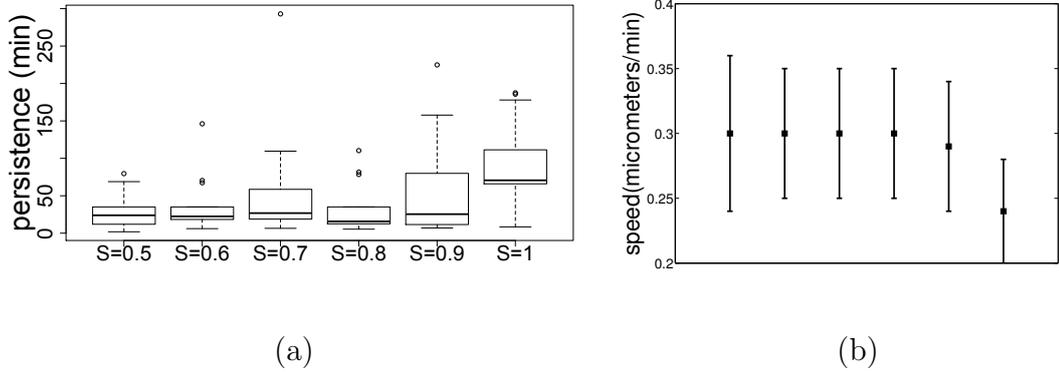
Figure 4.7: *Plots showing (a) the persistence time and (b) mean actual speed of migration of an individual cell on matrices of varying stiffness S between 0.5 and 1.*

(Figures 4.6(e)-4.6(h), Figure 4.7(b)). Presumably this is due to the fact that the cells are in contact with fewer matrix fibres in this case. This follows from the fact that in soft matrices the re-orientation allows for preserving the contacts between the fibre and the cell for longer time. Whether or not this is also true in experiments remains to be seen, although the results of Lo et al. [2000] [Table 1, pg. 148] appear to substantiate this.

In the simulations cells can also be tracked over longer periods of time and it can be seen that they exhibit a random walk behaviour. The period of time chosen was 3 days. In these simulations a specific seed was chosen for the random number generator which produces the stochasticity in the cell movement. It was chosen to be the seed which led to the median persistence time in previous simulations of cells on a very stiff matrix and also in the simulations of cells on a matrix where the stiffness for each fibre was calculated independently. Also here the difference becomes clear between the path of a cell that reorients the matrix (Figure 4.8(b)) and that of a cell that does not (Figure 4.8(a)). When matrix reorientation occurs, it leads to very sharp, non-smooth turns in the cell path

(denoted with *green asterisks* in the image). The movies of these simulations can be found in the supporting material of Schlüter et al. [2012] as Movies S3 and S2, respectively.
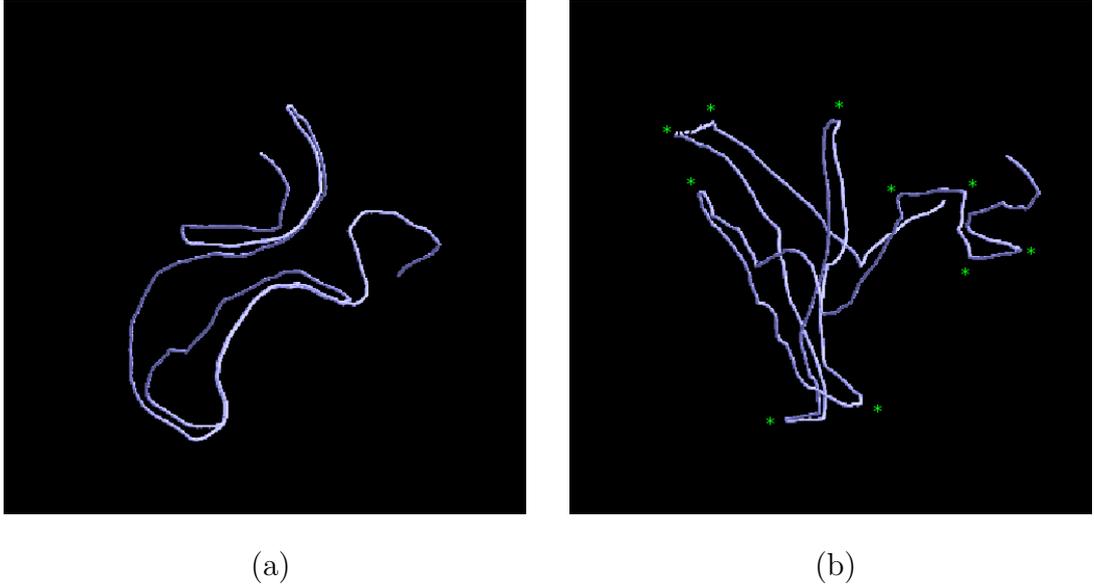

(a)   (b)

Figure 4.8: *Plots showing the cell paths which have developed over 3 days (a) without and (b) with matrix reorientation (using matrix architecture 1 from previous Figure 4.6). In case (b), the matrix stiffness S is dependent on fibre connectedness. The plots show that without matrix reorientation, the cell path is much smoother and does not contain any sharp turns by the cell (a). In contrast, if the cell reorientates the matrix (b), it undertakes many more sharp turns and changes of direction, denoted by asterisks. The black squares are each an area of 500×500µm.*

4.5.2   On whether cell movement is guided by substrate rigidity

Experiments carried out by Lo et al. [2000] have shown that individual cell movement can be guided solely by physical interactions between the cells and the
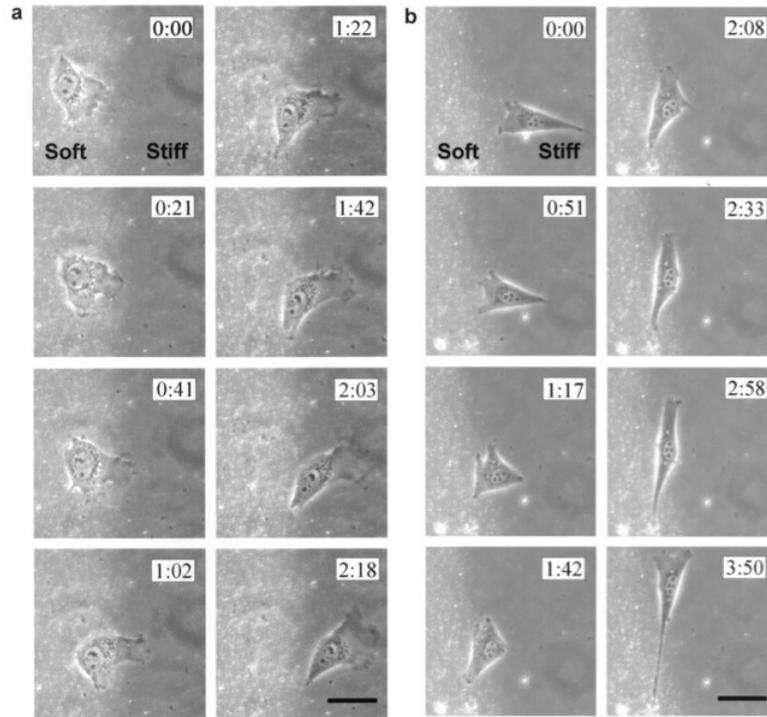
Figure 4.9: *Figure showing experiments by Lo et al. [2000] where cells have been placed close to the gradient on a matrix with two different stiffnesses. Panel (a) shows a cell being placed on the softer side of the matrix which then over time migrates onto the stiffer side of the matrix. Panel (b) shows a cell being placed on the stiffer side of the matrix initially and how, over a period of time, it moved towards the gradient but then stayed on the stiffer side of the matrix. Reproduced with copyright permission from Lo et al. [2000].*

underlying substrate. 3T3 fibroblasts were placed in the middle of a collagen-coated polyacrylamide substrate sheet where one half of the sheet was "soft" and the other half of the sheet was "stiff". The results showed that the cells either migrated onto the stiffer side when starting on the soft side or stayed on the stiff side when starting there (see Figure 4.9) i.e. cells tend to prefer stiff matrices to softer ones which has been termed "durotaxis".

These experimental results provided a scenario that could be tested with this

modelling approach. We created a two-dimensional domain similar to that used in Lo et al. [2000] with a different stiffness of matrix in each half, creating a "transition of rigidity" across the middle. This configuration is shown in Figure 4.10(a). A single cell was placed close to the transition zone, indicated by a red asterisk in Figures 4.10(b) - (e). Initially the left side of the matrix was assigned a stiffness of 0.25 and the right side of the matrix was assigned a stiffness of 0.75 which meant that the cell started out on the stiffer side. 15 simulations were run over 3 days of real time using this configuration and the final locations of the cell in the domain were noted. Then the stiffness properties of the matrix were switched around leading to the cell starting on the softer side as the right side of the matrix was then assigned a stiffness value of 0.25 and the left side a value of 0.75. The results are shown in Figure 4.10(b) and (c). It can be seen that there is a slight preference for the stiffer side of the matrix as the cell stayed on the stiffer side in eight of the 15 simulations where it started on the stiffer side and it moved to the stiffer side in nine of the 15 simulations where it started on the softer side. However, when the discontinuity in rigidity was increased, the results became much clearer as can be seen in Figures 4.10(d) and (e). Here the soft side was given a stiffness of 0.05 and the stiff side was given a value of 0.95. In the simulations where the cell started on this very stiff side, the cell stayed on that side in 12 out of the 15 simulations. In the set of simulations where it started on the soft side, it still ended up on the stiffer side in 13 out of the 15 simulations. Qualitatively, these results mirror those found in Lo et al. [2000] i.e. there is an apparent preference of cells for a stiff substrate. The simulation results indicate that the reorientation of the matrix or the lack thereof on stiffer matrices may play an important role in durotaxis. The fact that we are able to reproduce these results by using this computational model suggests that the physical structure of the ECM is a sufficient condition for a cell to choose a particular location within

(b)

(c)

(d)

(e)
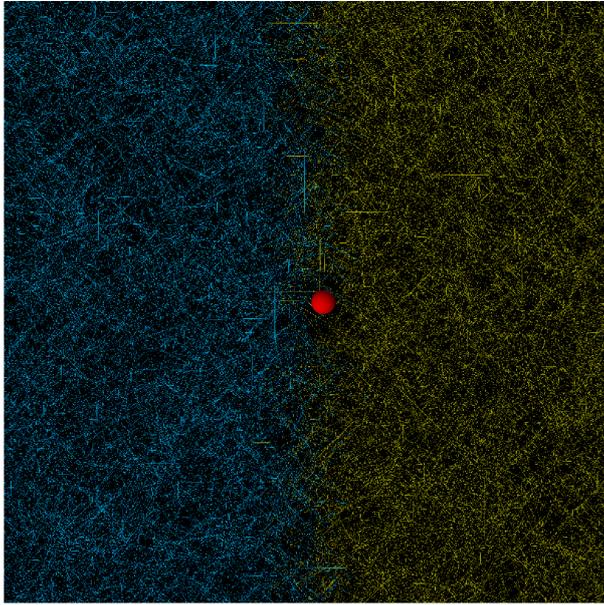
(a)

Figure 4.10: *(a) Figure showing how the matrix was divided into two sides of different stiffness on the left hand side and right hand side of the domain with the cell being placed initially just on the right hand side of the domain (different stiffnesses denoted by different colours.). The cell was always placed initially in the same position but the matrix properties were altered. Simulations were run with the left side of the matrix being soft and very soft and the right side being stiff and very stiff and vice versa. The results are shown in plots (b)-(e). The squares are end-points of the cells after 3 days and the asterisks show the cell's starting position.*

64
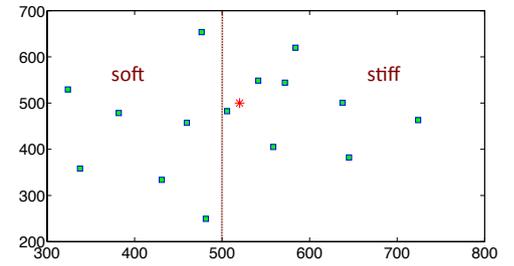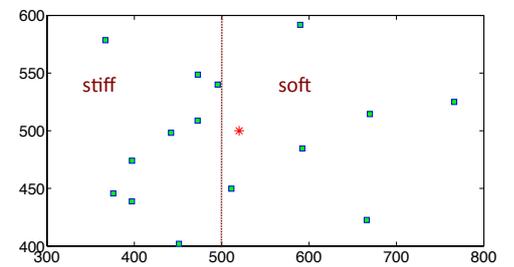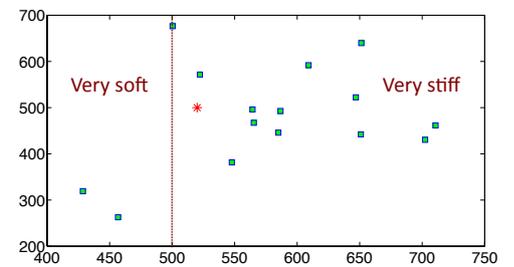
the surrounding environment. It is difficult to conclude this from biological observations alone, since in experiments the internal cell dynamics, internal cell biomechanics, cell phenotypic properties, etc., are all factors that could possibly play a role and these cannot be ignored experim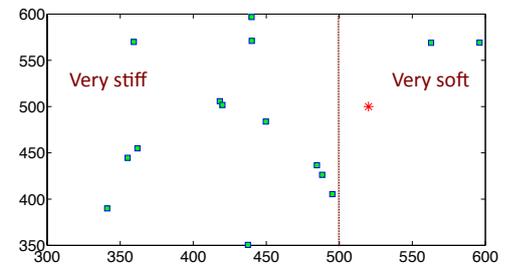entally. In this model the location of the cell depends uniquely on the physical structure of the matrix and on the cell movement. The findings do suggest that these are sufficient to explain the cells' preference for stiffer matrices.

4.5.3 Nonlinear dependencies of persistence time and cell speed on matrix composition and architecture

In order to investigate the influence of cell speed and matrix characteristics on persistence time, we ran a number of simulations of cells migrating on matrices, where the cells were given different maximum cell speeds (denoted by s_{max}) and the composition of the matrices was varied in terms of different fibre lengths and densities. In all of these simulations the matrix stiffness was calculated independently for each fibre as explained above and the first matrix architecture from the previous simulations was used. In the first set s_{max} was increased from 10μm/h to 20μm/h in steps of 2.5μm/h. At the same time the matrix fibre length was varied between 25μm and 100μm in steps of 12.5μm, whereby the number of matrix fibres was always increased or decreased accordingly so that the overall density of matrix fibres is not altered. For each combination, 10 simulations were run and the persistence times and mean actual speed were calculated. The results can be seen in Figure 4.11. The plots in Figure 4.11(a) show that for "slow" cells (maximum speeds of 10, 12.5, 15μm/h), those cells that migrate on shorter fibres are more persistent than the cells migrating on longer fibres. For maximum cell speeds of 17.5 and 20μm/h, fibre length becomes less important and persistence times are more or less independent of fibre length. However, from the plot in

Figure 4.11: *Plots showing how the mean persistence time and actual mean speed of cell migration vary with matrix fibre length and maximum cell speed. (a) Plot of the mean persistence time in minutes varying with maximum cell speed and mean fibre length; (b) Plot of the mean actual cell speed during the same simulations varying with maximum cell speed and mean fibre length.*

Figure 4.11(b), the mean actual speed of the cells seems to depend in a bimodal manner on the fibre length.

In the second set of simulations the cell speed was again increased from 10μm/h to 20μm/h in steps of 2.5μm/h. However this time the matrix density was changed by varying the total number of fibres in the domain. Specifically we placed a number of fibres of length 75μm, ranging from a total of 7500 to 22500 using increments of 3750, in the domain of size $1000{\times}1000\mu$m.
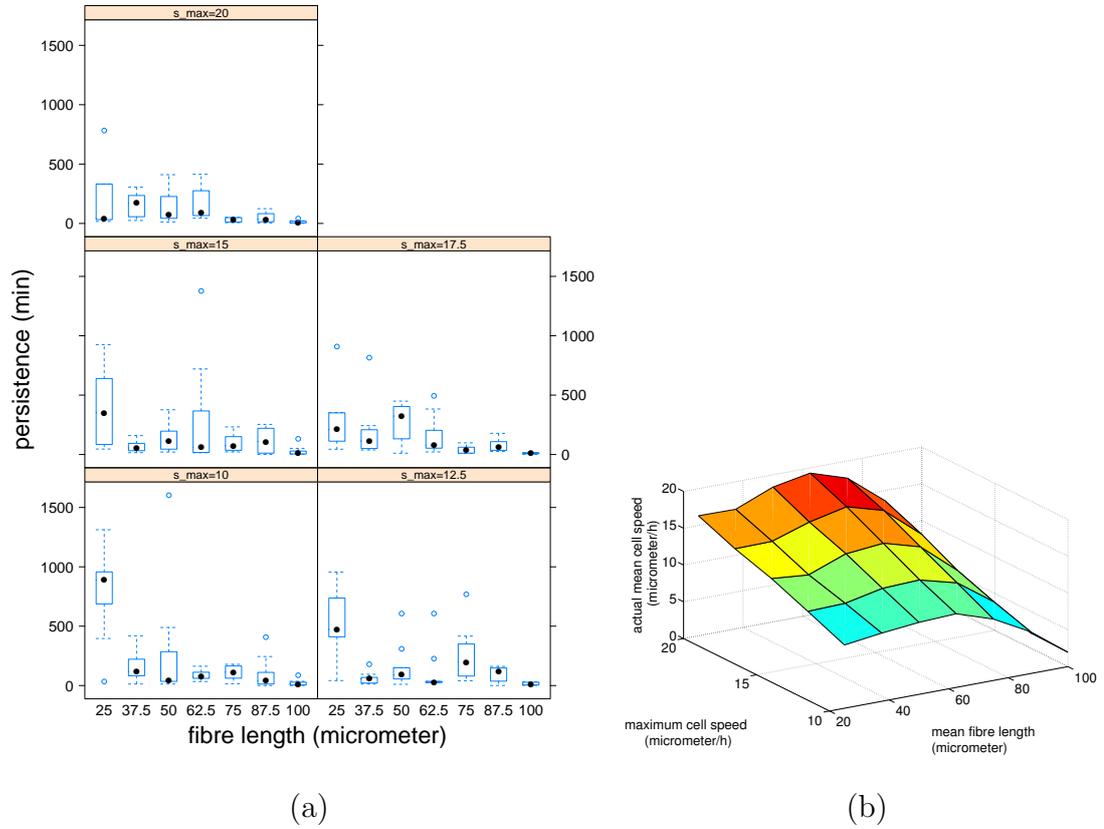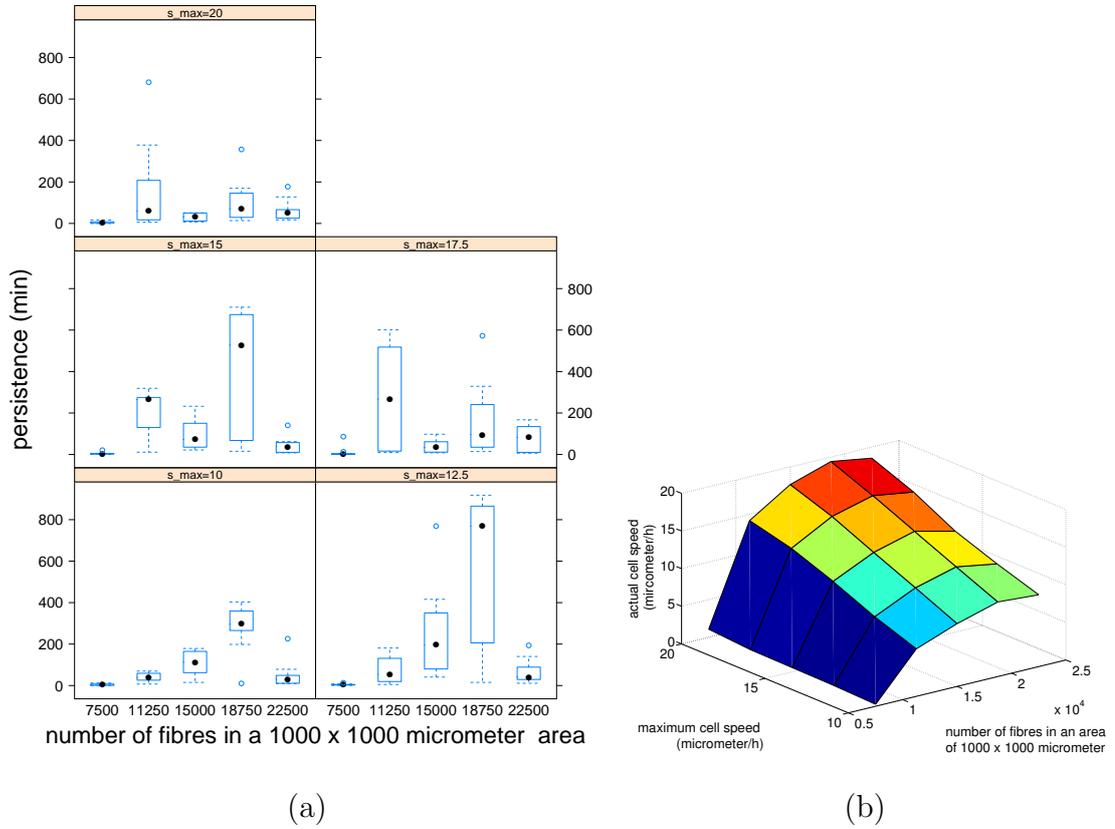
Figure 4.12: *Plots showing how the mean persistence time and mean actual speed of cell migration vary with matrix density and maximum cell speed.*

(a) Plot of the mean persistence time in minutes varying with maximum cell speed and matrix density, i.e. number of matrix fibres; (b) Plot of the mean actual cell speed during the same simulations varying with maximum cell speed and matrix density.

Again 10 simulations were run for each combination and the persistence times and the mean actual speed were calculated. The results can be seen in Figure 4.12. The plots in Figure 4.12(a) show that for cells moving at low maximum speeds (10, 12.5, 15μm/h) there is an increase in persistence with the density before there is a drop at the highest density (22500 fibres in the domain). For high maximum cell speeds (17.5 and 20μm/h) this effect is lost and persistence times

67

are more or less independent of matrix density. From the plot in Fig.4.12(b) the actual cell speed clearly shows a biphasic dependence on the matrix density, a fact observed in experiments [Palecek et al., 1997].

Finally the behaviour of cells on matrices of varying degrees of anisotropy was investigated. Specifically, we compared the persistence of cells on a random matrix to that on a biased matrix and to that on a fully aligned matrix (see Figure 4.1 for the initial conditions of each type of matrix). 15 simulations were run for each matrix type and the cell tracks were examined in each case. Representative plots are given in Figure 4.13. Comparing the cell tracks in each plot, it can be seen that on average the cells persist in a given direction for a longer period of time on the aligned and biased matrices than on the random (isotropic) matrices. These results indicate that persistence in a given direction (unsurprisingly) decreases with matrix randomness. A formal analysis using persistence time calculated from equation (4.15) is not possible since this equation is derived only for motion in isotropic environments [Alt, 1990].
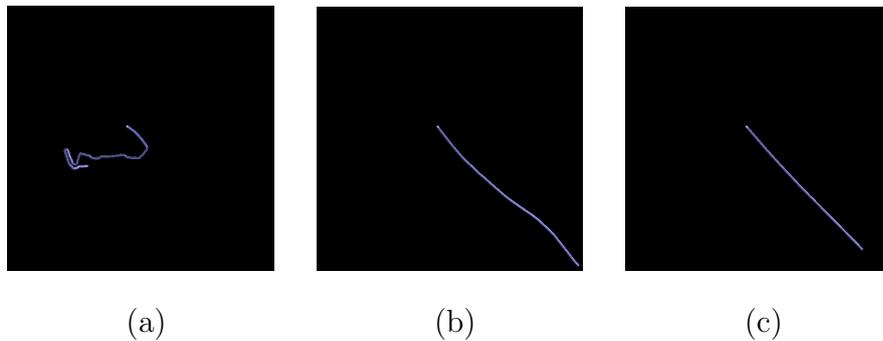

(a) (b) (c)

Figure 4.13: *Figures showing the plots of cell tracks over a period of 6 hours on extracellular matrices with (a) randomly distributed fibres, (b) biased fibres, (c) aligned fibres.*

4.6 Sensitivity of the results towards so far unconsidered parameters

4.6.1 Sensitivity towards the noise terms

The noise term in the calculation of the cell's polarity axis as well as the noise term $\mathbf{f}_j(t)$ could potentially have a strong influence on the persistence of the cell movement. Therefore we investigated these terms in more detail by running multiple sets of simulations varying the standard deviations. First the sensitivity of the results towards the noise term in the calculation of the polarity axis was considered. This is the noise term χ that is added to the weight of each fibre in the calculation of the cell's polarity axis (see equation (4.5)), thus influencing the cell's direction of movement. In order to ensure that this does not have too big an impact on the results, simulations were run with 25%, 50% and 100% increase and decrease of the standard deviation used to generate normally distributed numbers with zero mean. Again 15 simulations were run for each value of the standard deviation and the persistence times were calculated. The results are given in Figure 4.14.
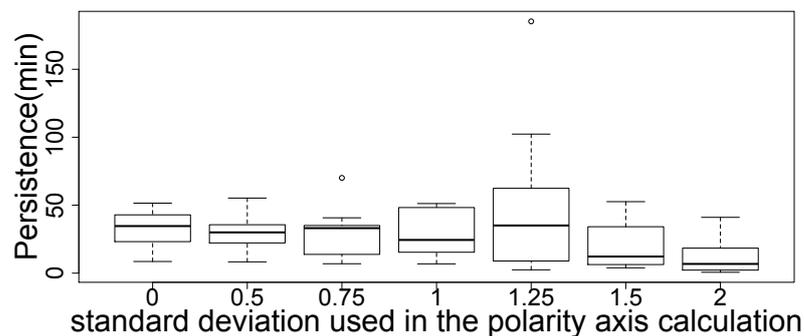

Figure 4.14: *Plots showing the mean persistence time (in minutes) for different values of the standard deviation used for generating χ.*

In the simulations with standard deviations close to one, the persistent times are quite stable and no real impact of this parameter can be seen. If a bigger range is taken into account, it becomes obvious that an increase in standard deviation leads to a decrease in persistence although even here the difference is not very pronounced as the range of persistence times measured is very similar in all cases. Therefore it is clear that although this parameter has a slight influence on the results, it is by no means the determining factor.

We also investigated the influence of the noise term $\mathbf{f}_j(t)$ on the cell's behaviour. For this the standard deviation was increased from 0, which had been used so far, to 0.0001, 0.001, 0.01, (which was then of the same order of magnitude as the cell velocity per time step) and 0.1. Again 15 simulations were run per standard deviation and the results for the persistence time can be seen in Figure 4.15.
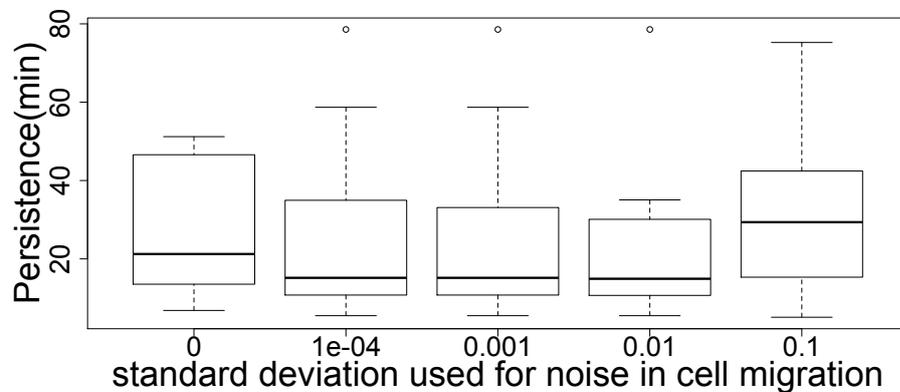
Figure 4.15: *Plots of the persistence time (minutes) of cell migration where the standard deviation in the noise term $\mathbf{f}_j(t)$ is varied from 0 to 0.1.*

For small standard deviations the influence of this noise term is minimal. Surprisingly the range in which the persistence times lie decreases for a standard deviation of 0.01. However as soon as the standard deviation becomes higher than the maximum cell speed per time step, the range of values for the persistence

time increases which is what one would expect. Nonetheless, it is encouraging to see that for smaller standard deviations up to the order of magnitude of the cell velocity, the noise does not have a great influence on the results and the model is thus insensitive to small perturbations in this parameter.

4.6.2 Influence of parameters in modelling the matrix re-arrangement due to cell traction forces

Finally the influence of two further parameters on the results of the simulations was investigated. The parameters in question are used in calculating the matrix rearrangement due to cell traction forces. The first parameter to be considered was the factor 0.1 which reduces the reorientation (equation (4.13)). The results of varying this parameter by 10% and 20% are shown in Figure 4.16. It can be seen that the results are not very sensitive to this parameter for, although the spread increases by varying it, the median changes very little.
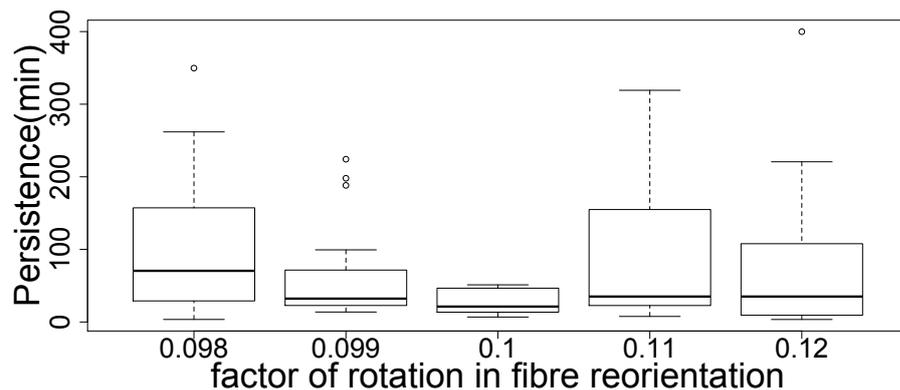
Figure 4.16: *Plots of the persistence time (minutes) of cell migration where the factor that reduces the reorientation (given a value of 0.1 in equation 4.13) is varied by 10% and 20%*

The second parameter to be investigated closer is the number of fibre cross-links at which the maximum stiffness of 0.95 is reached. Again this number was varied from its initial value of 15 by 10% and 20%. The results are shown in Figure 4.17. In this case the spread of the results increases again but, similarly, the median changes very little.

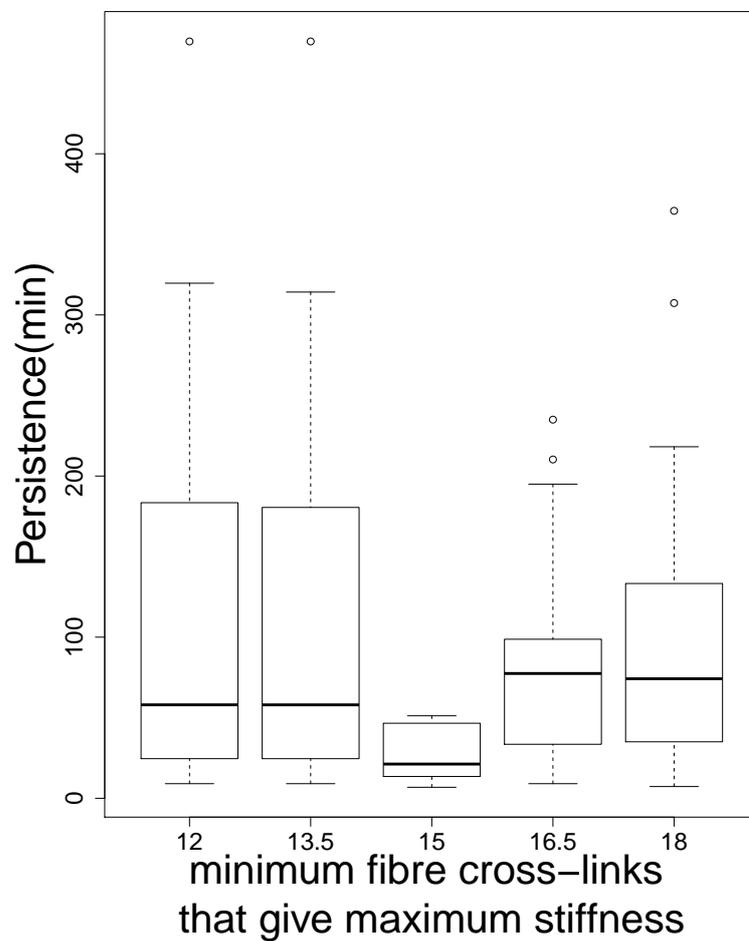**Figure 4.17:** *Plots of the persistence time (minutes) of cell migration where the number of fibre cross-links, at which the maximum stiffness is reached, is varied by 10% and 20% above and below the original value 15.*

4.7 Discussion

In this chapter we have formulated a modelling framework for single cell migration on two-dimensional matrices in which the cell as well as the matrix fibres are individual elements or agents. Using this approach, we investigated the influence of matrix stiffness on cell migration and found that the reorientation of the matrix fibres due to cell traction forces might be an important part of this process as very stiff, non-reorientable, matrices led to very variable and occasionally very high persistence times which do not agree with experiments. We also ran simulations to test whether the model could reproduce experiments that showed the preference of cells for stiffer matrices. The results agreed with these experiments and suggest that matrix reorientation or the lack thereof on stiff matrices, may be an important factor in durotaxis. This is particularly interesting as the physical interactions between cells and matrix cannot be isolated in experimental settings. Therefore in those set ups it is difficult to distinguish between the roles of the physical interactions and the intracellular signalling pathways coupled to these. The model makes it possible to ignore the intracellular signalling. It shows that purely the interactions between cells and the matrix together with the physical structure of the matrix, can explain the preference that cells have for stiffer rather than softer matrices.

Furthermore we examined the relationship between persistence time as well as actual cell speed, and maximum cell speed and matrix fibre length or matrix density. In both cases we found that there is a nonlinear dependency of the persistence time on the two factors, especially at low cell speeds, as well as a biphasic dependency of the actual cell speed on the fibre length and matrix density. Additionally we looked at the influence of the matrix structure on the persistence and found that, unsurprisingly, a more ordered matrix leads to a higher persistence. Finally we also investigated the sensitivity of the results towards a change of

certain, so far unconsidered, parameters. We examined the stability of the results depending on both noise terms which influence the movement of the cell and found that they do not have a significant influence on the persistence times measured. Similarly small perturbations in the parameters used to calculate the rearrangement of the matrix fibres due to cell traction forces have little impact on the results.

In this model the focus is on the most fundamental processes underlying cell migration at the level of cell–matrix interactions. Naturally certain simplifications have therefore been made. One of these is that the extracellular matrix fibres were modelled as rigid cylinders that were not connected. This means the cell's application of force onto one fibre does not affect other fibres in the closer environment. Similarly there is no counter force pulling the fibres back into their original place after the cell has moved across them. These are clearly two aspects that might have an impact on the results and are probably also the reason why the results for very loose, medium stiff and variably stiff matrices are very similar. Including matrix elasticity would most likely lead to less realignment in stiffer matrices and might thus alter the results. However, some things can already be learnt from this model or can at least give interesting hypotheses e.g. that matrix remodelling might play an important part even in two dimensional migration. Also that durotaxis might be a process that does not necessarily require intracellular signalling pathways. Furthermore other results indicate a non-linear dependency of persistence time on cell speed and fibre length or cell speed and matrix density. As the results are quantitative and measurable in the laboratory, it would be very interesting to see whether these hypotheses can be confirmed by experiments. This is especially the case as persistence time is often used to characterise and compare cellular behaviour and it is therefore important to understand all the factors that can influence it in an experimental setting. However,

a lot of parameters used in the model had to be estimated as no measurements could be found in experimental literature, especially concerning the forces involved in cell–matrix interactions. Therefore, although the results are a first step towards gaining more insight into this process, the model should be seen mainly as a framework that can lead to quantitative and predictive results given 'real' experimental input data.

Chapter 5

Modelling Two Cells Migrating: "Follow my Leader"

5.1 Introduction

In the previous chapter a computational model was developed for single cell migration on two dimensional surfaces. However, in experiments a single drop of a suspension of hundreds or thousands of cells is generally placed on a petri dish or well plate. Therefore in this chapter we extend the single cell migration model to a two cell model, the results of which are exemplary for the most fundamental building blocks that make up the behaviour of any number of cells migrating on a two dimensional matrix. The simulations show that cells have a tendency to follow each other. This behaviour can often be seen in experiments. In three dimensions, this 'multicellular streaming' [Friedl and Wolf, 2010] is well-described and it is clear that it is a very efficient way for cells to migrate along the tracks and tubes of individual 'leader cells'. However, apart from statements from some biologists (i.e. private communications), we could not find any evidence or explanation of this in the literature available. Therefore, in this chapter, we will

investigate which parameters have an influence on this behaviour by varying the matrix stiffness, the initial distance between the cells, the fibre lengths and fibre densities and by calculating the lengths of time the cells follow each other under these conditions.

5.2   Model description

The basic model used here is the same as in the previous chapter. Again, individual cells are modelled as relatively flat and hemispherical objects with a radius of the base of 15μm. A height of 2.6μm is assumed. Also the individual matrix fibres are again represented by thin cylinders, the lengths of which are normally distributed with mean 75μm and standard deviation 5μm and the widths are 200nm.

For each cell, its movement is again governed by equation (4.1) and the matrix fibres are also reoriented as before using equation (4.13). Here the stiffness of the matrix fibres is always calculated dependent on fibre interconnectedness. All the parameters for which it was shown in the previous chapter that their variation has little impact on the results, are given their original values, i.e. the number of fibre cross-links at which the maximum stiffness is reached is given the value of 15, the factor that reduces the reorientation is given the value of 0.1, the standard deviation used to generate $\mathbf{f}_j(t)$ is 0.01, the standard deviation used to generate χ is 1 and s_{max} is $20\mu m/h$. One of the only two differences between the model here and the one in the previous chapter is that initially two cells are placed on the matrix instead of one. This means that the contact inhibition of locomotion (CIL) which is encoded in the model, now becomes important. As mentioned in the previous chapter, cells have been found to change their direction after contact with another cell [Mayor and Carmona-Fontaine, 2010]. Generally this is due to the fact that the cell cannot extend protrusions in the direction in which

it contacts another cell. Thus this can have an influence on the cell's polarity [Mayor and Carmona-Fontaine, 2010]. Just as in experiments, in the model-based simulations, a cell can lose its polarity when it makes contact with another cell. In this case the cell then has to re-establish its front-rear polarity. Unlike before, when only one cell was considered and therefore CIL could not occur, here it is an important part of the simulations. Furthermore we assume that the matrix reorientation by one cell can be constrained by the presence of the second cell. If both cells are in contact with the same fibre, we assume that one cell should not be able to reorient that fibre in such a away that it loses contact with the other cell. Thus, after calculating the reorientation angle, it is checked whether or not the other cell is still in contact. If this is the case, the reorientation takes place as before. Otherwise, the reorientation is reduced so that the other cell stays in contact with the fibre.

5.3   Computational simulation algorithm

Similar to the previous simulations, the time step is chosen to be three seconds and the procedure during each time step can be summarised as follows:

Step 1:

Each fibre is examined to see whether the cells were in contact with it during the last time step and whether the cells are polarised. If both of these conditions are true for at least one cell then it means that the cell(s) has exerted force on the fibre during that time step. All the fibres for which this is the case are reorientated by one cell at a time in the way explained in equation (4.13). After each reorientation the fibre is checked to see that it has not lost contact with the other cell and, if necessary, the reorientation is decreased.

Step 2:

For each cell all the fibres that the cell is in contact with are found and it is

determined whether the two cells are in contact.

Option 1: If the cells are not in contact and the cell in question is in contact with at least one fibre and has established front rear polarity either through previous polarisation or through the new formation of focal contacts over 10 minutes, the new polarity axis is calculated using equation (4.5).

Option 2: If the cells are in contact and the cell in question has front-rear polarity, defining the rear to be where the contact with the other cell is, and the cell is in contact with at least one fibre, the new polarity axis is calculated using equation (4.5).

Option 3: If the cells are in contact and the cell in question has front-rear polarity, defining the front to be where the contact with the other cell is, the cell loses its polarity. If the cells are in contact and the cell in question has already lost its polarity and is re-polarising, then the direction of the fibres w_i in equation (4.5) is chosen to be pointing away from the contacting cell.

Step 3:

For each cell, if the cell is polarised, the net force that it generates for its movement on the matrix is calculated. This is done using the knowledge from the previous step of how many fibres the cell is in contact with. Together with the given parameter determining the maximum speed the cell can reach, the calculation of how fast the cell should be moving (see Fig. 4.2(a)) and the magnitude of the force the cell must therefore be generating (see Fig. 4.2(b)) can be made. Equation (4.6) is then used to calculate the net force.

Step 4:

Both cells are moved according to the forces calculated in step 3. This is done by first solving equation (4.1) for the cell velocity and then applying the Forward Euler method. This gives the new position of the cells at the end of this time step.

5.4 Data analysis

Initially it was observed that the cells follow each other in some simulations for long periods of time and in some simulations only very briefly and very seldom. Therefore, in order to try and understand the impact that certain parameters have on this behaviour, it was necessary to find a way of quantifying the simulation results. In order to do so a MATLAB code was written that calculated the distribution of how long the two cells follow each other in each 'experiment' defined by varying one parameter. For each case, 15 simulations were run in each of which a different seed for the random number generator was used for the cell movement, as was done in previous simulations. In order to establish whether or not cells were following each other, a distance matrix was determined for each of the 15 simulations by calculating the distance between each point on the first cell's track to each point on the other cell's track. Thus the value at position i, j in the distance matrix is the distance between the first cell at time step i and the second cell at time step j. Then the time points were searched for where the cells were closer than two-cell radii. Once such a time point was found, it was tracked to find out how long these episodes lasted. This was done by counting the number of subsequent time points for i and j, for which the entry (i, j) stayed smaller than two-cell radii. Those elements of the matrix that were counted once, were then marked to ensure they were not counted again. This procedure was followed for all 15 simulations in each set. Then the distribution was determined by creating an array of 'bins' from 0 (cells only come closer than two-cell radii for one time step) to the maximum number of time steps that cells were found to stay closer than two-cell radii. The discovered timespans were then binned for all 15 simulations. This gave the distribution of how long the two cells followed each other in this particular 'experiment'.

After first plotting the normalised distribution of the experiments (see Fig. 5.2), it

became very clear that the distributions are difficult to compare in this way. Given the skewness of the distributions and the long tail, varied width notched box plots were chosen to be the best presentation of the results and were calculated in R. The widths of the box plots give an indication of how big the data set is; in this case how many instances were counted where cells followed each other. A wider box represents a bigger data set than a thinner one as the width is proportional to the square-root of the number of observations in that data set [R Core Team, 2012]. As in normal box plots, the median was drawn and the box encompasses the 25%- and 75%-ile. Here the range of the whiskers was again 1.5 and outliers were drawn as circles. In addition the boxes are 'notched' giving an indication of the significant difference when comparing the medians between two 'experiments'. The notches extend to $\pm 1.58 \times \frac{\text{inter-quartiles-range}}{\sqrt{\text{sample number}}}$ [R Core Team, 2012]. If the notches of the two box plots do not overlap, then this is an indication that the difference between the two medians is statistically significant.

5.5 Computational simulation results

In the first simulation that was run with two cells, one cell was kept stationary for 10 minutes while the neighbouring cell was allowed to polarise and start migrating. In Figure 5.1(a) the stationary cell is shown in red and the other cell in green. As soon as the second cell could polarise it started following its former neighbour along the tracks the neighbouring cell had laid down (Figure 5.1(b)-5.1(i)). It continued to do so until the two cells came into contact with each other which made the following cell lose its polarised state and re-polarise in another direction due to the contact inhibition of locomotion encoded in the model. This behaviour could be observed a multiple of times during the simulation over 3 days of real time. A movie of this simulation can be found in the supporting material of Schlüter et al. [2012] as Movie S1. However, when a more systematic
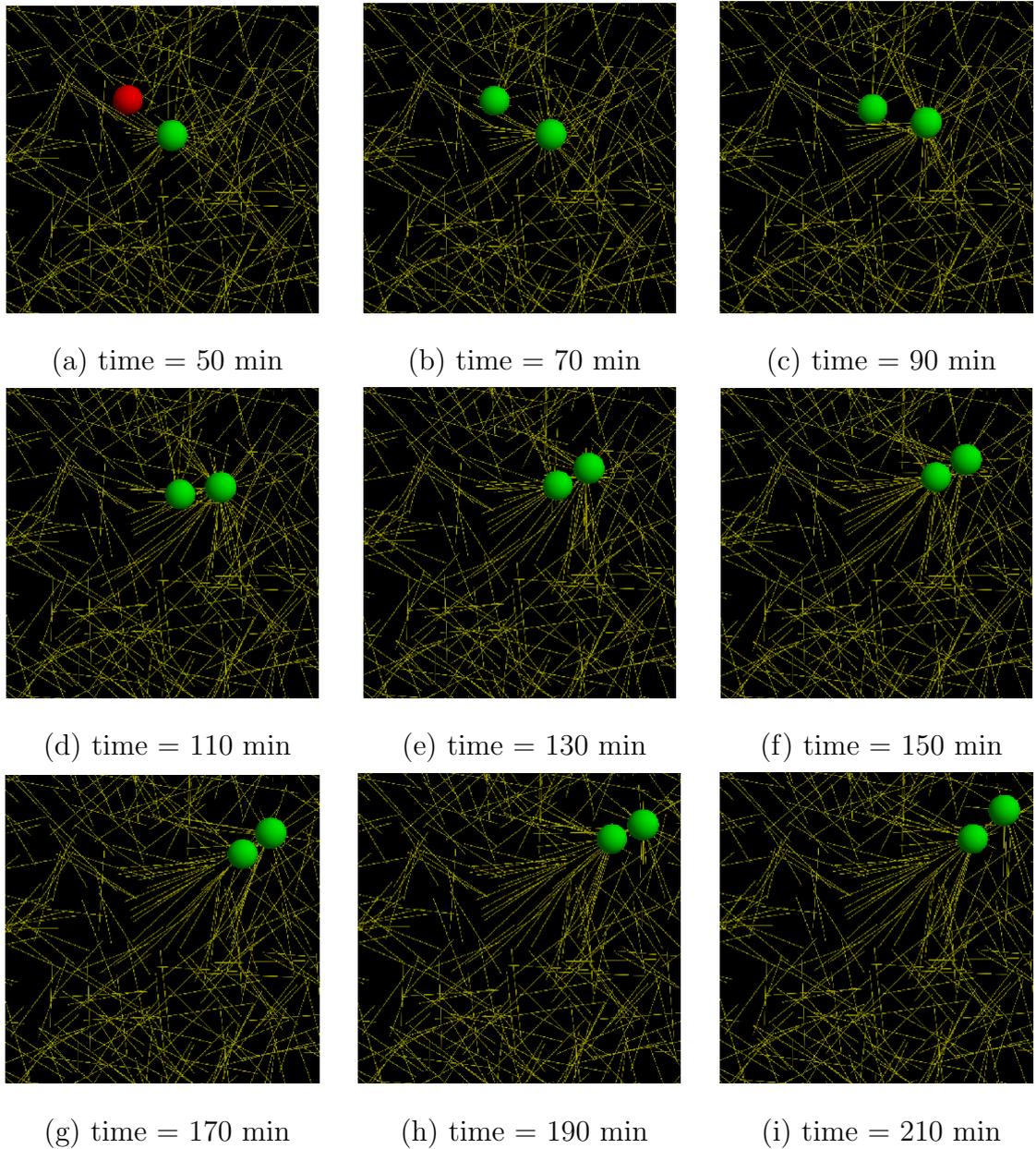
Figure 5.1: *Plots showing snapshots at times t = 50-210 minutes of two cells following each other through the matrix. Green denotes a front-rear polarised cell, while red denotes a non-polarised cell. The plots show that an initially non-polarised cell (t = 50 minutes) becomes polarised (t = 70 minutes) and then follows the path of the initial polarised cell.*

approach was taken and the parameters were all chosen as mentioned in the model description, the simulations did not necessarily show the two cells moving along the same track for a period of time long enough for it to eliminate the notion that this had happened coincidentally. Thus the different parameters which might have an influence on the results were varied and the results are presented below.

5.5.1 The influence of matrix stiffness and initial cell–cell distance on the migration patterns

In order to investigate the influence of certain parameters on the behaviour of the two cells more thoroughly, four different sets of simulations with 15 simulations per set were run. First two cells were placed immediately next to each other and one of them was kept stationary for 10 minutes while the other could polarise and start migrating as was done in the simulation seen in Figure 5.1. In the second set of simulations the two cells were again placed next to each other but the matrix used was a stiff, non-reorientable matrix. The results can be seen in Figure 5.2. If the cells did not come close enough it was not counted at all, and therefore zeros in these plots denote that the cells came together at just one point in time. The plots in Figure 5.2 clearly show that cells on stiffer matrices follow each other for longer periods of time - Figure 5.2(b) shows that cells can spend a time of up to 1300 seconds close to each other (which is over 20 minutes), compared to 700 seconds in Figures 5.2(a). In order to compare the results more easily, varied width notched box plots were generated from these distributions. In addition to the two sets of simulations above, two further sets were added. Instead of placing the cells very closely next to each other, they were placed with a little more distance between them (i.e. the cells were not touching at the beginning of the simulations). This means that there was a larger range of different fibres either cell could have been in contact with. The box plots showing the results of
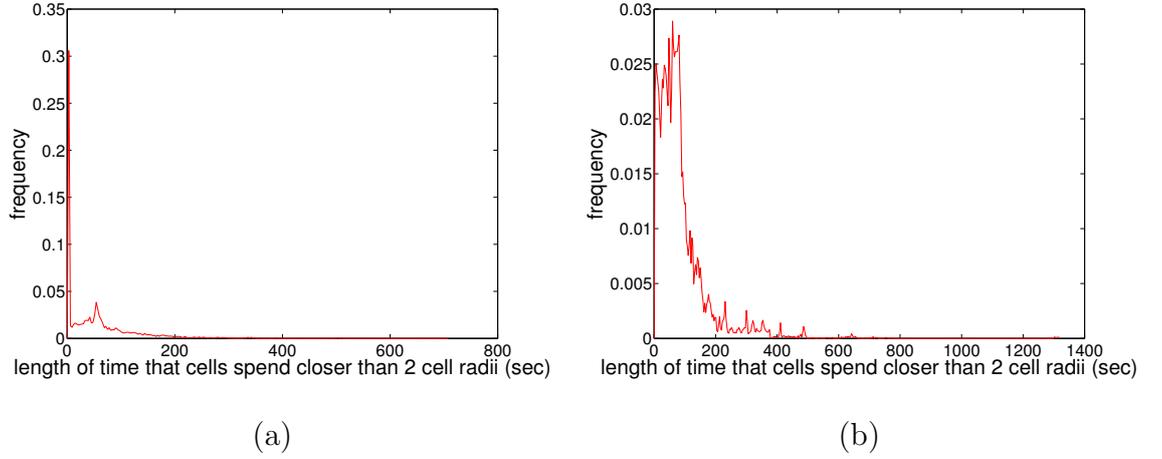
Figure 5.2: *Plots showing the distribution of the amount of time that cells are closer than two cell radii. (a) Distribution of the time lengths for two cells reorienting the matrix during movement. (b) Distribution of the time lengths for two cells on a stiff, non-reorientable matrix.*

these four 'experiments' are given in Figure 5.3. Considering the width of the box plots, it can be seen that the size of the data sets (i.e. the number of occurrences of one cell following the other), is largest for the cells initially being placed right next to each other, both on a re-orientable and a stiff matrix. The smallest, by quite some margin, is the one resulting from the simulations where the cells were placed further apart on a stiff matrix.

The shape of the distributions and the long tails are captured in the large number of outliers that exist for all four sets of simulations. The highest value is reached on a stiff matrix at over 1200 sec as was also shown in Figure 5.2. The highest values of the other three sets lie between ≈ 800 sec (original setup) and ≈ 500 sec (distant). A summary of the other important values in these plots, the median, the 25%ile, the 75%ile and the upper and lower end of the whiskers, are given in Table 5.1. Apart from the lower end of the whisker which is at 3 seconds (1 timestep) for each set of simulations, all the values are highest for the cells
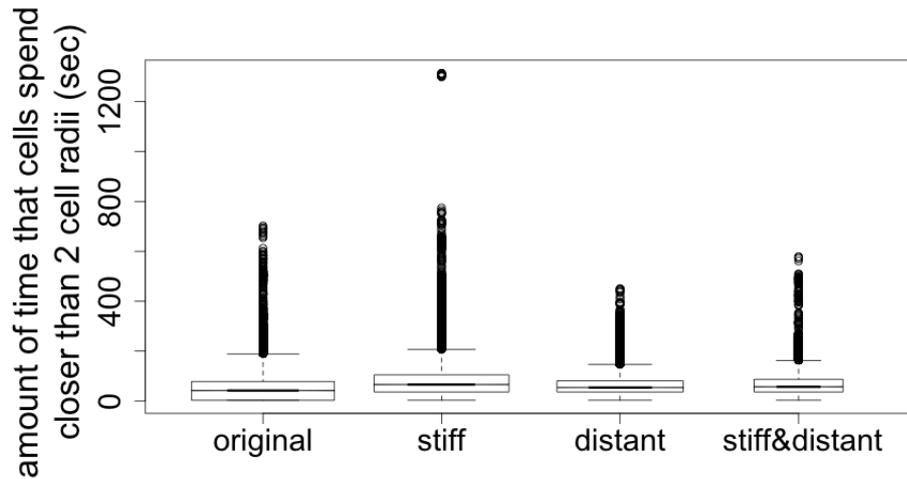
Figure 5.3: *Varied width notched box plots generated from the distribution of the amounts of time cells spend closer than 2 cell radii under the original conditions (*original*), on a stiff matrix (*stiff*), when the cells are placed on a matrix further apart (*distant*) and when the cells are placed further apart on a stiff matrix (*stiff&distant*).*

migrating on stiff matrices. This is followed by cells that were initially placed further apart on stiff matrices which is interesting as this is the set of simulations in which cells followed each other the least often compared to the other ones in this group (see box plot width in Fig 5.3). On softer matrices the spread of the length of time the cells follow each other is longer when cells start off close together than when they are initially placed far apart. This is obvious from the first quartile being much lower in the first case but then with all the other values catching up with the results for the second case. The upper end of the whisker of the box plot for cells starting off close together is considerably higher than that of cells starting at a distance from each other. Whether or not these differences are statistically significant is difficult to say from the present data. However, taking into account the notches in the box plots, which are not much more than lines

Table 5.1: *A summary of the data of the box plots in Figure 5.3*

	original	stiff	distant	stiff&distant
lower end of whisker	3	3	3	3
first quartile	3	36	36	36
median	42	66	54	57
third quartile	78	105	81	87
upper end of whisker	189	207	147	162

in the box plots given here, it can at least be said that there is a statistically significant difference between the median of the original set of simulations and that of the simulations where a stiff matrix was used.

5.5.2 The impact of fibre length on the cells' behaviour

As in the previous chapter for single cell migration here the fibre length could also have an important role to play. Thus, using the softer matrix from the previous simulations, the fibre lengths were varied from 25μm to 100μm in steps of 25μm. The number of fibres placed in the 1000μm$\times1000\mu$m area was increased or decreased accordingly to prevent a change in the matrix density. In all the simulations the cells here were placed close together initially.

The results are shown in Figure 5.4. Here the sizes of the data sets look to be very similar, with all box plots having similar widths. The longest times that cells followed each other were reached by cells migrating on the longest fibres of 100μm. In these simulations the cells followed each other for maximum times of ≈ 800 seconds. A similar value was reached for cells migrating on fibres of length 75μm. In the other two cases the maximum length of time is between 500 and 600 seconds. Apart from these extreme outliers, the results can be divided into
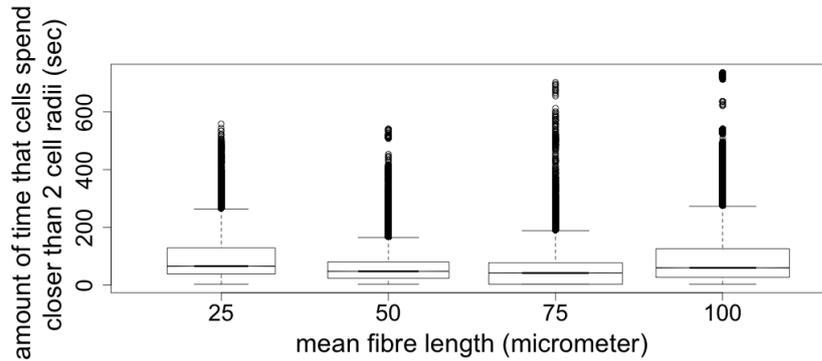
Figure 5.4: *Varied width notched box plots generated from the distribution of the amounts of time cells spend closer than 2 cell radii during migration on a matrix made-up of fibres of mean length 25μm, 50μm, 75μm and 100μm.*

two groups the first one being the results for fibres of length 25μm and 100μm and the second one those for fibres of length 50μm and 75μm. This can also be seen in Table 5.2.

Table 5.2: *A summary of the data of the box plots in Figure 5.4*

	25	50	75	100
lower end of whisker	3	3	3	3
first quartile	39	24	3	27
median	66	48	42	60
third quartile	129	81	78	126
upper end of whisker	264	165	189	273

Especially the values for the third quartile and the upper end of the whiskers show the divide clearly with the difference being almost 100 seconds. Also the differences between the medians of the simulations using 100μm and 25μm fibres and the results for simulations using 50μm and 75μm fibres are statistically significant as can be seen in Figure 5.4. However, in both of these groups a similar

pattern exists. In both cases the resulting values for the shorter fibre in the group are higher up to the third quartile where they become very similar and then for the upper end of the whisker the values for the longer fibres are higher. Thus, although there is a clear distinction between extreme (25 and 100 μm) and moderate fibre lengths (50 and 75μm), in both of these groups, the spread of the lengths of time that cells follow each other is larger for the cells migrating on the longer fibres.

5.5.3 The impact of fibre density on the cells' behaviour

Fibre density is another factor the variation of which might have an impact on the modelling results. Thus again 3 sets of 15 simulations were run in which the matrix density was altered. In one set the original matrix density was used by placing 15000 fibres of mean length 75μm in a domain of 1000$\times$1000μm. For the other two sets the number of fibres was decreased to 7500 and increased to 22500. Figure 5.5 shows the results of those simulations.
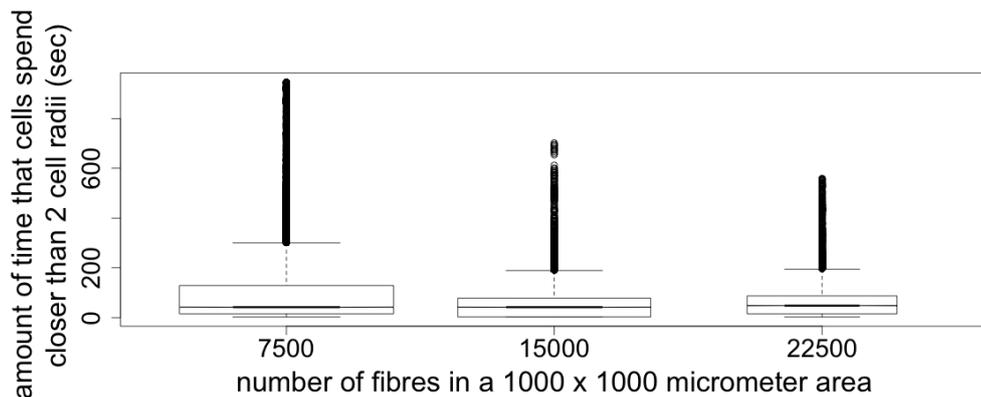

Figure 5.5: *Varied width notched box plots generated from the distribution of the amounts of time cells spend closer than 2 cell radii during migration on matrices of different densities.*

The width of the box decreases with increasing fibre number. Thus, cells are

more likely to follow each other on less dense matrices than they are on very dense matrices. This is not necessarily surprising as fewer fibres mean that the cell has fewer contacts and therefore when encountering a track that is already laid down in a certain way there are fewer stimuli steering it away in a different direction. The maximum amount of time that cells spend following each other also decreases with increasing fibre numbers from ≈ 1000 seconds to ≈ 600 seconds. Table 5.3 shows however that the difference between the three sets of simulations decreases for the upper end of the whiskers and the third quartile.

Table 5.3: *A summary of the data of the box plots in Figure 5.4*

	7500	15000	22500
lower end of whisker	3	3	3
first quartile	15	3	15
median	42	42	48
third quartile	129	78	87
upper end of whisker	300	189	195

The median is almost the same in all three sets and also the notches in Figure 5.5, which are again not much more than lines in the box plots presented here, show that there is no significant difference between them. Thus although the range of the lengths of time that the two cells follow each other is greater when the cells migrate on sparse matrices, the median is not different from when the cells migrate on dense matrices.

5.5.4 The impact of the combined variation of fibre length and density on the cells' behaviour

When changing the fibre length previously, the number of fibres was increased or decreased accordingly so that the density would be unchanged. Changing the fibre length clearly still had an effect on the results. Thus it would be interesting to see what influence a change in both of the parameters has on the results. To this end, the fibre lengths were decreased to a mean of 25μm and the number of fibres in the domain was varied. The highest number of fibres was 45000 which gives the same density as placing 15000 fibres of length 75μm in the domain and was thus used earlier. The next value chosen was 22500 which gives the same density as having 7500 fibres of length 75μm. Also lower numbers of 15000 and 7500 fibres were placed in the domain for a set of simulations. The results are shown in Figure 5.6.
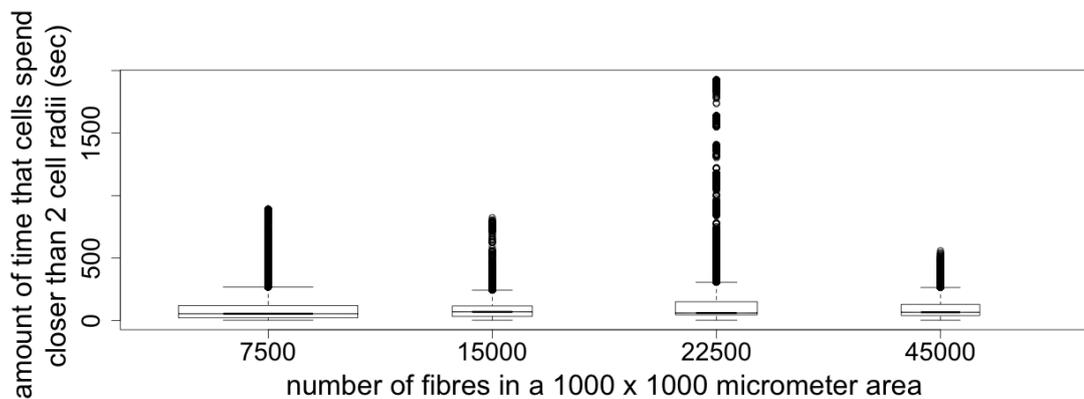

Figure 5.6: *Varied width notched box plots generated from the distribution of the amounts of time cells spend closer than 2 cell radii during migration on matrices of different densities constituted of fibres of mean length 25μm.*

The data set is clearly the largest for the simulations on matrices of 7500 fibres. The reason is probably the same as above, that the fewer contacts the cell has, the fewer stimuli there are to move the cell off a path that has already been laid down.

In the other three sets the width of the boxes is very similar. What is striking is the spread of the outliers in the simulations with 22500 fibres. Cells follow each other in these simulation for up to ≈ 2000 seconds. This is considerably longer than in the simulations with 7500 fibres of mean length 75μm. Therefore it must be the combination of fibre length and density and not the density alone which leads to the results. Also all the percentile values are higher in these sets of simulations than they are when varying the density alone and can be seen by comparing Table 5.4 with Table 5.3.

Table 5.4: *A summary of the data of the box plots in Figure 5.6*

	7500	15000	22500	45000
lower end of whisker	3	3	3	3
first quartile	21	33	45	39
median	54	69	60	66
third quartile	120	117	150	129
upper end of whisker	267	243	306	264

Apart from the value of the outliers, there is generally little difference between all the different sets of simulations run with fibres of mean length 25μm. Thus it seems as though it is the fibre length that generally determines the behaviour of two cells placed next to each other on the matrix. The fibre density on the other hand seems to have an influence on the overall range of the results. This can be concluded from the variation of the maximum lengths of time for which the cells follow each other between the different sets of simulations.

5.6   Discussion

In this chapter we have presented a natural extension to the model in Chapter 4 by placing two cells in the domain. The cells were then tracked over three days of real time and their behaviour classified by calculating the distribution of the lengths of time that the cells followed each other. This was done under different conditions. First the matrix stiffness and the initial distance between the cells was altered. Then different fibre lengths and densities were considered before finally both, the fibre lengths and densities, were varied. The results showed that matrix stiffness has an impact on the cells behaviour. The cells followed each other for longer periods of time on stiff matrices than they did on softer, re-orientable ones. This was the case for both, cells which were initially close together and cells which were initially further apart. Whether the matrix was soft or stiff, the median of the length of time for which the cells followed each other was slightly higher and the range was wider for cells that were initially close together rather than further apart. In an experimental setting it is rather unlikely that cells would start off very close together but it is possible that they 'bump' into each other while migrating which then causes at least one of the cells to stop due to CIL. Then the same situation could occur as at the beginning of the simulations, that one cells keeps migrating (or starts migrating again) while the other one first has to polarise. If both cells stop and have to re-polarise then they would move away from each other afterwards, however if one cell keeps migrating then it is possible for the other cell to re-establish the polarity axis in the same direction as the other cell and thus follow this cell. Possibly by following each other initially, a very well-established path is laid down which makes it more likely that the cells follow each other along this track again at other times during the simulation. This is the only reason we can find why there is an obvious difference between cells starting off close together or further apart.

The second set of investigations which focused on the impact of the fibre lengths on the cells' behaviour, showed that the maximum lengths of time that the cells follow each other (the largest outliers) generally increased with fibre length. All the other characteristic values behaved less linearly though. Cells migrating on short fibres (25μm) and long fibres (100μm) spent longer following each other than they did on matrices consisting of fibres of mean lengths 50μm or 75μm. The spread was larger in these cases and also the median was higher with the difference between the medians from these sets and the ones with mean fibre lengths of 50μm and 75μm being statistically significant. It is difficult to find an explanation for this. One possibility is that very short fibres have fewer intersections with other fibres and are therefore easier to reorient than longer ones. Similarly in the case where long fibres are used, there are fewer of them in order to keep the density unchanged which might also result in fewer cross-links. However previous simulations showed that stiffer matrices led to longer timespans of cells following each other which would contradict this explanation.

When varying fibre density, the maximum length of time that cells followed each other for, decreased with increasing density. As explained above, this could possibly be due to the fact that the cells have fewer contacts to different matrix fibres on sparse matrices compared to dense matrices. Thus when a cell encounters a track that has been laid down by the other cell, it is more likely to follow it on a sparse matrix as there are fewer stimuli pulling the cell in a different direction. However the observed effect is not mirrored in the value of the median, meaning that it is just the range and the spread of the lengths of time that cells follow each other that is larger on sparse matrices than on dense ones.

When combining a change in fibre length and fibre density it was difficult to find a clear pattern. The most distinctive results from this were the high values for the lengths of time that the cells followed each other in the case where 22500

fibres of mean length 25μm were placed in the domain. Values of ≈ 2000 seconds were reached which is higher than in any other set of simulations with fibres of length 25μm and also far higher than in the simulations with 7500 fibres of mean length 75μm which gives the same density as was considered here. Thus it must have been this specific combination of fibre length and density which led to these results. As before, here again it was just the range which showed this interesting behaviour and little difference could be found between the medians of the different sets of simulations. Therefore this leads to the conclusion that the fibre length determines the general cell behaviour and the fibre density influences the range of the behaviour that can be observed.

In summary, one can say that fibre length and matrix stiffness seem to have an impact on the behaviour of two cells placed on the extracellular matrix whereas fibre density mainly affects the range of the lengths of time that cells follow each other for rather than the median of these values. Unfortunately we could not find any experimental or theoretical publications to compare this with and thus a validation of these results is currently not possible. The way cells behave in this context is rather important, though, since, for example, cells that invade their environment during cancer development have been shown to follow each other's tracks. Although this is obviously generally an observation in three dimensions, its two dimensional counterpart as considered here, could possibly complete the picture and shed more light on the underlying factors. Thus hopefully there will be experimental findings in the future to compare these results with.

Chapter 6

A Multiscale Model of *in vitro* Cancer Cell Invasion in a 2D Domain

6.1 Introduction

Using the work from the previous chapters, we now formulate a model of cancer cell invasion in a two dimensional *in vitro* setting. This is done by extending the previous model to include cell–cell adhesion and repulsion and cell division. While the repulsive forces between two cells are governed by their bio-mechanical properties, the strength of adhesion between two cells is determined by the number of E-cadherin–β-catenin bonds they form. The intracellular dynamics underlying the E-cadherin–β-catenin interactions happen on a much faster time scale than the cellular components and are modelled using ordinary differential equations giving the complete model a multiscale nature. In addition, compartmentalisation of the cell's intracellular domain in terms of cell–cell contact areas, leads to a spatial intracellular model. The parameters for this intracellular model are

obtained from fitting the model to data from the available literature. We then use this model to study the development of cell colonies dependent on internal and external cell characteristics. This leads us to find cell behaviour and cell-matrix interactions that trigger invasiveness as is seen in cancer cells.

6.2   The intracellular model

In Chapter 3.3 a brief description was given of the multiscale model published by Ramis-Conde et al. [2008b]. The intracellular component of that model focusses on E-cadherin–β-catenin dynamics as will be the case here. However, in order to develop as simple a model as possible, Ramis-Conde et al. [2008b] made a number of assumptions which do not reflect the recent findings of these dynamics and, most importantly, highly influenced the outcome of the simulations. The main simplification in this model is that the intracellular spatial aspect of the dynamics is not taken into account. All the species - free E-cadherin, membrane bound E-cadherin and E-cadherin–βcatenin complexes at the membrane which form bonds with E-cadherin–β-catenin complexes on the surface of the neighbouring cells - are modelled with one ordinary differential equation each. If a cell has multiple neighbours, the model assumes that this one ordinary differential equation describes the dynamics of the number of E-cadherin–β-catenin complexes, and therefore the number of cell–cell bonds, at each of the cell–cell contact areas. Thus, if one of the neighbours starts detaching, the overall number of E-cadherin–β-catenin complexes decreases in this model by Ramis-Conde et al. [2008b]. Because of the nature of the model, the number of bonds and thus the adhesion decreases between this cell and *all* its neighbours, although in reality it should only be decreasing between this cell and the neighbouring cell that has started detaching.

In order to address these concerns, in this chapter we develop a new, intracellular

spatial model of the dynamics which will be parameterised by first translating the number of E-cadherin–β-catenin bonds between cells into forces and then fitting these to actual data of cell–cell adhesion forces found in the available literature.

6.2.1 The E-cadherin–β-catenin pathway: model description

As shown by Hinck et al. [1994], E-cadherin and β-catenin bind at the endoplasmatic reticulum immediately after production. The complex is then trafficked to the cell membrane [Hinck et al., 1994, Chen et al., 1999]. Other molecules such as α-catenin can then bind to it and upon cell–cell contact, the complex can form adherens junctions with E-cadherin–β-catenin complexes on the neighbouring cell [van Roy and Berx, 2008]. The E-cadherin–β-catenin complex can also be endocytosed and there are multiple possible scenarios for this [Bryant and Stow, 2004]. Here, we assume that endocytosis has no effect on adhesion, i.e. only those complexes that are not involved in adherens junctions are endocytosed (as proposed by Le et al. [1999]) or endocytosed junctional complexes are replaced almost immediately by exocytosis of a newly formed complex. Junction disassembly followed by endocytosis leads to the disruption of the E-cadherin–β-catenin complex and the components can either be degraded, recycled for cell–cell adhesion or reused in different signalling contexts. Since production of neither molecule is explicitly taken into account, the possibility of degradation is also not considered but it is assumed that the overall amount of E-cadherin and β-catenin is at a steady-state. The effect of down-regulation of E-cadherin will be studied by decreasing the total amount of E-cadherin. The dynamics described are modelled with a compartment model and ordinary differential equations that describe the transition between compartments and interactions between the species as shown

in Figure 6.1. Each cell–cell contact site, as well as the cell's cytoplasm, is considered as a separate compartment. Free E-cadherin and free β-catenin only exist in the cytoplasmic compartment. The cytoplasmic compartment can also hold E-cadherin–β-catenin complexes which includes those complexes that are at the cell membrane but not at a site of cell–cell contact. E-cadherin–β-catenin complexes at a specific contact site are in the corresponding cell–cell contact compartment.
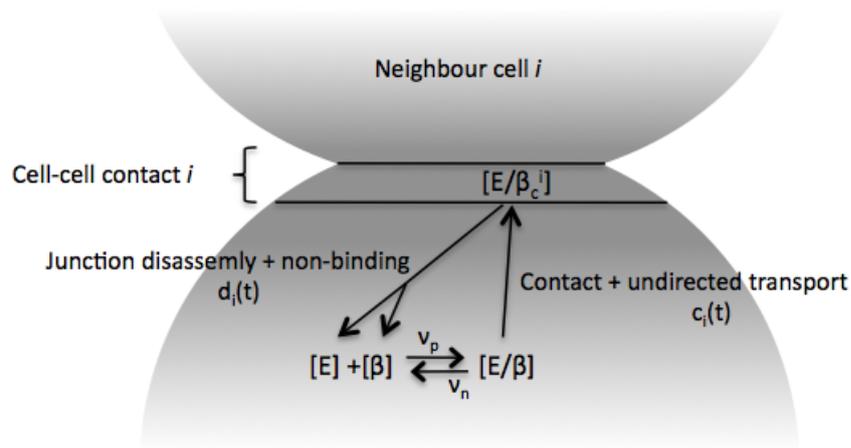
Figure 6.1: *Schematic diagram showing the E-cadherin–β-catenin dynamics as considered in the model. Free E-cadherin (E) and β-catenin (β) in the cytoplasm bind to form a complex (E/β). In adherent cells, in addition to the general transport to the cell surface, E-cadherin–β-catenin complexes are trafficked to the contact area. If the complex is transported to a site of cell–cell contact (it is denoted E/β_C^i at cell contact site i) it can bind complexes on the neighbouring cell's surface. If there is no binding partner, the complex can be internalised again and recycled. The same process takes place when bonds are broken due to junction disassembly.*

The equations governing the described E-cadherin–β-catenin dynamics are as follows:

$$\frac{dE}{dt} = -\nu_{\mathrm{p}} * E * \beta + \nu_{\mathrm{n}} * E/\beta + \sum_{\mathrm{i=1}}^{\#\mathrm{contacts}} d_{\mathrm{i}}(t) * E/\beta_{\mathrm{c}}^{\mathrm{i}}, \qquad (6.1)$$

$$\frac{d\beta}{dt} = -\nu_{\mathrm{p}} * E * \beta + \nu_{\mathrm{n}} * E/\beta + \sum_{\mathrm{i=1}}^{\#\mathrm{contacts}} d_{\mathrm{i}}(t) * E/\beta_{\mathrm{c}}^{\mathrm{i}}, \qquad (6.2)$$

$$\frac{dE/\beta}{dt} = \nu_{\mathrm{p}} * E * \beta - \nu_{\mathrm{n}} * E/\beta - \sum_{\mathrm{i=1}}^{\#\mathrm{contacts}} c_{\mathrm{i}}(t) * E/\beta, \qquad (6.3)$$

$$\frac{dE/\beta_{\mathrm{c}}^{\mathrm{i}}}{dt} = c_{\mathrm{i}}(t) * E/\beta - d_{\mathrm{i}}(t) * E/\beta_{\mathrm{c}}^{\mathrm{i}} \ , \ \forall \mathrm{i}, \qquad (6.4)$$

where E is free E-cadherin, β is free β-catenin, E/β are non-adhesion effective E-cadherin–β-catenin complexes in the cytosol or at the cell membrane at non-contact sites and $E/\beta_{\mathrm{C}}^{\mathrm{i}}$ are E-cadherin–β-catenin complexes at the contact site with cell i. ν_p is the E-cadherin–β-catenin binding rate and ν_n the complex dissociation rate. The complex formation and dissociation terms have also been considered by Ramis-Conde et al. [2008b]. The compartmentalisation of the cell into a cytoplasmic compartment and individual compartments for each cell-cell contact site is however a new concept and the terms associated with these are therefore newly derived here.

Furthermore $d_{\mathrm{i}}(\mathrm{t})$ describes the endocytosis of E-cadherin–β-catenin complexes due to either junctional disassembly or because there is no binding partner at the surface of the neighbouring cell at this contact site. Thus

$$d_{\mathrm{i}}(t) = \underbrace{a_{\mathrm{d,i}} * \rho_{\mathrm{d}}}_{\substack{\text{internalisation due to}\\\text{junctional disassembly}}} + \eta * \underbrace{\frac{1}{1 + exp(-2 * (E/\beta_{\mathrm{C}}^{\mathrm{i}} - E/\beta_{\mathrm{c,i}}^{\mathrm{i}}))}}_{\text{internalisation due to non-binding}} \qquad (6.5)$$

Here the first term on the right side of the equation represents the internalisation due to junctional disassembly and the second term describes internalisation due to non-binding. $E/\beta_{\mathrm{c,i}}^{\mathrm{i}}$ is the density of E-cadherin–β-catenin complexes at the site

of contact with cell i, expressed by cell i. The second term is a smooth approximation to a Heaviside function which is zero when $E/\beta_{\mathrm{c}}^{\mathrm{i}}$ is much smaller than $E/\beta_{\mathrm{c,i}}^{\mathrm{i}}$ and one when $E/\beta_{\mathrm{c}}^{\mathrm{i}}$ is much greater than $E/\beta_{\mathrm{c,i}}^{\mathrm{i}}$, thus giving an internalisation of E-cadherin–β-catenin complexes only when there are fewer complexes at the cell-cell contact site in cell i than there are at that site in the cell currently of interest. η is the internalisation rate in this case. ρ_{d} is the internalisation rate of complexes freed by cell–cell detachment. The two rates are considered to be different due to the fact that complexes involved in adherens junctions have more binding partners and are also linked to the cytoskeleton and therefore their dissociation is a much slower process [Miyashita and Ozawa, 2007]. $a_{\mathrm{d,i}}$ gives the loss of contact area with cell i at time t:

$$
a_{\mathrm{d,i}} = \begin{cases} \|\frac{\partial}{\partial t}\hat{a}(t)_{\mathrm{i}}\|, & \text{if } \frac{\partial}{\partial t}\hat{a}(t)_{\mathrm{i}} < 0, \\ 0, & \text{otherwise,} \end{cases}
$$

where $\hat{a}(t)_i$ is the contact area between the two cells at time t. This term for the internalisation due to junctional disassembly is taken from Ramis-Conde et al. [2008b].

Similarly $c_{\mathrm{i}}(t)$ gives the exocytosis of E-cadherin–β-catenin complexes at the site of contact with cell i. Different mechanisms are possible for this. Experiments that study cell-cell adhesion only ever consider two cells and therefore, to our knowledge, it is not known what happens when more cells adhere to each other. Two different scenarios can be thought of. In the first case the adhesion between two cells remains unchanged when a third cell attaches to them. This would imply that the number of E-cadherin–β-catenin complexes that is trafficked to a cell-cell contact site and forms bonds is limited. Thus if one assumes that a cell should, for example, be able to have six neighbours, then the number of E-cadherin–β-catenin complexes at a given cell-cell contact site could only be a maximum of one sixth of the total possible number of E-cadherin–β-catenin complexes. In the second case, the adhesion between two cells decreases when

a third cell attaches to them and decreases again during the attachment of a fourth cell and so on. This means that almost the total number of possible E-cadherin–β-catenin complexes can be at a single cell-cell contact site, but as soon as contact is made with another cell, the complexes are redistributed. We develop an exocytosis term for both of these scenarios.

Model 1:

$$c_i(t) = \underbrace{a_{c,i} * \rho_c * (1 - \alpha_i(t))}_{\text{directed exocytosis}} + \underbrace{\iota * \frac{\hat{a}(t)_i}{4 * \pi * r(t)^2}}_{\text{undirected exocytosis}}$$

Model 2 :

$$c_i(t) = \underbrace{a_{c,i} * \rho_c}_{\text{directed exocytosis}} + \underbrace{\iota * \frac{\hat{a}(t)_i}{4 * \pi * r(t)^2}}_{\text{undirected exocytosis}}$$

We assume that two processes are involved in the exocytosis, (i) directed trafficking to the contact site which is given by the first term on the right and, (ii) undirected transport given by the second term. The assumption that undirected transport occurs is based on the fact that E-cadherin–β-catenin complexes can be found at the entire cell surface (not just at sites of contact) and also in non-adherent cells [Le et al., 1999]. We model this undirected transport by assuming it to be proportional to the ratio of the contact area to the total cell surface area, with the proportionality constant being ι and $r(t)$ being the cell's radius at time t. The directed trafficking is dependent on the increase in contact area given by:

$$a_{c,i} = \begin{cases} \frac{\partial}{\partial t}\hat{a}(t)_i, & \text{if } \frac{\partial}{\partial t}\hat{a}(t)_i > 0, \\ 0, & \text{otherwise.} \end{cases}$$

ρ_c is the complex translocation rate. This term is again taken from Ramis-Conde et al. [2008b]. As explained above, in Model 1 we assume that trafficking is furthermore dependent on the number of E-cadherin–β-catenin complexes already at the site of contact as the overall number is limited. If the number of complexes at the cell-cell contact site is smaller than the maximum number allowed, then

$\alpha_i(t)$ is the ratio of complexes at the cell-cell contact site to the maximum number allowed. Otherwise it is one.

In addition to the above dynamics, Model 2 has additional equations that govern the redistribution of E-cadherin–β-catenin complexes between different cell-cell contact sites. Two new variables have to be introduced for this: C_+^i is the number of E-cadherin–β-catenin complexes that are translocated from other cell-cell contact sites to the site of contact with cell i and C_-^i is the number of complexes that is moved from cell-cell contact site i to other contact sites.

$$C_+^i = \gamma * \sum_{j=1, j \neq i}^{\#\text{contacts}} E/\beta_C^j * \left(1 - \frac{\frac{E/\beta_c^i}{\hat{a}(t)_i}}{\frac{E/\beta_c^j}{\hat{a}(t)_j}}\right), \qquad (6.6)$$

$$C_-^i = \gamma * E/\beta_C^i * \sum_{j=1, j \neq i}^{\#\text{contacts}} \left(1 - \frac{\frac{E/\beta_c^j}{\hat{a}(t)_j}}{\frac{E/\beta_c^i}{\hat{a}(t)_i}}\right). \qquad (6.7)$$

C_+^i is added and C_-^i is subtracted from the right hand side of equation (5.4) in Model 2 giving the equation

$$\frac{dE/\beta_C^i}{dt} = c_i(t) * E/\beta - d_i(t) * E/\beta_C^i + C_+^i - C_-^i \, , \ \forall i. \qquad (6.8)$$

In the following we will try to fit both models to published cell-cell adhesion data in order to find good estimates for the parameters $\gamma, \eta, \iota, \nu_n, \nu_p, \rho_c$ and ρ_d.

6.2.2 The E-cadherin–β-catenin pathway: parameter estimation and model refinement

The data available from studies of cell-cell adhesion dynamics, are adhesion force data. Therefore we translated the percentage of E-cadherin–β-catenin bonds between two cells into an adhesion force. This is done by first assuming that the number of bonds between two cells is equal to the minimum of the E-cadherin–β-catenin complexes in either cell at the contact site. Published data that show the

force it takes to separate two cells [Chu et al., 2004] are then used and it is assumed that this is the adhesion force between the two cells. The data are interpreted differently for the two intracellular models. In Model 1 it is assumed that only a certain percentage of the total possible E-cadherin–β-catenin bonds can form at any contact site. Assuming that each cell in a monolayer should be able to have six neighbours and some free E-cadherin and β-catenin in the cytosol, only about 15% of the total possible number of complexes can be found at a cell-cell contact site. Thus for this model we assume that in the given data from Chu et al. [2004] only 15% of all possible E-cadherin–β-catenin complexes are at the contact site and generate the measured separation force of 210nN (see Figure 6.2(a)). For Model 2 we assume that 80% of the possible E-cadherin–β-catenin complexes in a cell produce the force of 210nN. Chu et al. [2004] measure the separation force of two cells using a dual pipette assay. As well as studying the influence of different E-cadherin expression levels (100%, 58%, 41%, 38%, 14% and 2%) on the force, they also look at the time course of the force during early cell-cell contact until its maturation after about 60 minutes. The results of the experiments are shown in Figure 6.2. Figure 6.2(a) shows the increase of force between two cells from 0nN at the time of their initial contact, to just over 200nN after 60 minutes. We used the force measurements after 5 minutes, 10 minutes, 30 minutes and 60 minutes to fit the models to. The black curve in Figure 6.2(b) shows the separation force of two cells dependent on the E-cadherin expression level. These measurements were taken 30 minutes after the initial cell-cell contact. When comparing Figures 6.2(a) and (b) it is noticeable that the separation force of two cells that express 100% of E-cadherin after 30 minutes is not the same in both experiments. This is most likely the case because different cells were used for these experiments and for each one specific cells were selected as a control in comparison to which the other cells in that experiment were analysed. In order to be able to use both sets
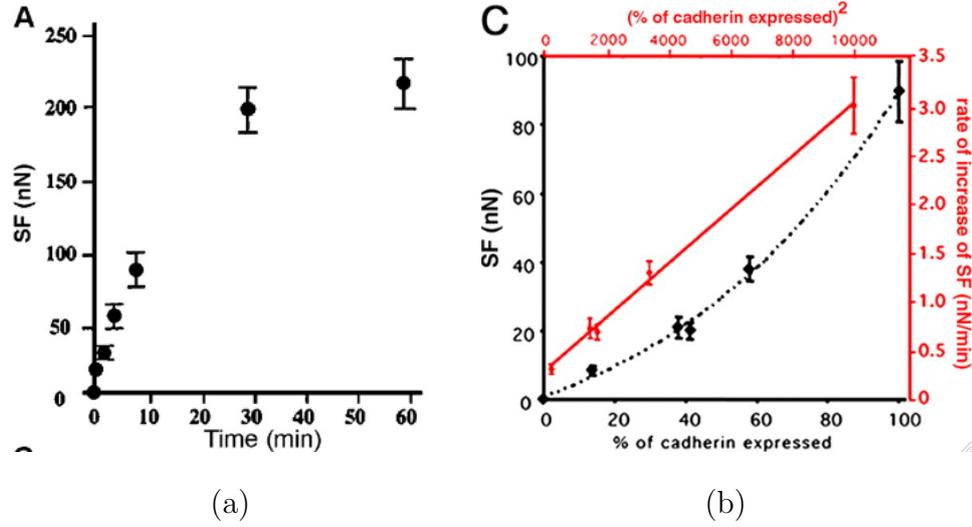
Figure 6.2: *Plots showing the separation force (SF) or adhesion force as measured by Chu et al. [2004] under different conditions. Plot (a) shows the force development after initial contact. The force reaches a steady-state after about 60 minutes. Plot (b) shows the dependency of cell-cell separation force on the percentage of expressed E-cadherin after 30 minutes of the initial cell-cell contact. Images reproduced from Chu et al. [2004] in accordance with RUP copyright policy.*

of data, we worked with the actual values of experiment (a) and scaled the results of experiment (b) accordingly so that the results for cells with 100% E-cadherin expression match those of experiment (a) after 30 minutes but the dependency of the separation force on E-cadherin expression levels is taken from experiment (b).

In order to fit Model 1 and Model 2 to the data, we wrote a MATLAB code that solves each system of ODEs using the MATLAB ODE-solver '*ode45*' and then calculates the adhesion force $\frac{d(\epsilon_{ij})}{d(d_{ij})}$ between two cells i and j from the number of bonds between them using the following equations.

Model 1:
$$\frac{d(\epsilon_{ij})}{d(d_{ij})} = min(E/\beta_{\text{c}}^{\text{i}}(t), E/\beta_{\text{c,i}}^{\text{i}}(t)) * \frac{210}{15}\text{nN} \qquad (6.9)$$

Model 2:
$$\frac{d(\epsilon_{ij})}{d(d_{ij})} = min(E/\beta_{\text{c}}^{\text{i}}(t), E/\beta_{\text{c,i}}^{\text{i}}(t)) * \frac{210}{80}\text{nN} \qquad (6.10)$$

The repulsion between the two cells is calculated using the Hertz model (see Chapter 3.2). The extended Hertz model (see Chapter 3.2) is then used to calculate the overall force the cells exert on each other and their movement is derived from this. The cells' resulting new position is then used in the next time step to calculate the new size of the contact area between the two cells and the change of the latter to feed into the ODEs. In order to simulate both models, the procedure was followed for 100 minutes of real time and the force between two cells was noted. The initial conditions were as follows:

$$E(0) \quad = \quad \text{E-cadherin expression level,} \qquad (6.11)$$

$$\beta(0) \quad = \quad 100, \qquad (6.12)$$

$$E/\beta(0) \quad = \quad 0, \qquad (6.13)$$

$$E/\beta_{\text{C}}^{\text{i}}(0) \quad = \quad 0 \ , \ \forall \text{i.} \qquad (6.14)$$

The first round of fitting was done assuming that only two cells came into contact. The radii of both cells was set to 5μm and the initial total contact area between the two cells was set to $1\mu\text{m}^2$. As mentioned before, for Model 1 we assumed that a maximum of 15% of all possible E-cadherin–β-catenin complexes can be trafficked to one cell-cell contact site. Thus we also assumed that the number of complexes at the cell-cell contact site in the neighbouring cell ($E/\beta_{\text{c,i}}^{\text{i}}$) is 15%. For Model 2 we set the number of complexes at the cell-cell contact site in the neighbouring cell equal to 80%. The other parameters needed were initially assigned numbers randomly generated by the inbuilt MATLAB random number generator '$rand$' from the log space between 10^{-6} and 10^6. As only the

105

attachment of two cells was considered in this round of fitting, only the parameters $\eta, \iota, \nu_n, \nu_p$ and ρ_c could be estimated for each model. As in Chapter 4.4, we used the inbuilt MATLAB function '*fminsearch*' to fit the cell adhesion data from the simulations to the experimentally found cell adhesion data in Figure 6.2. Figure 6.3 shows the results for both Model 1 (Figure 6.3(a)) and Model 2 (Figure 6.3(b)) using the set of parameters that gave the best fit. It can be seen

(a) Model 1   (b) Model 2
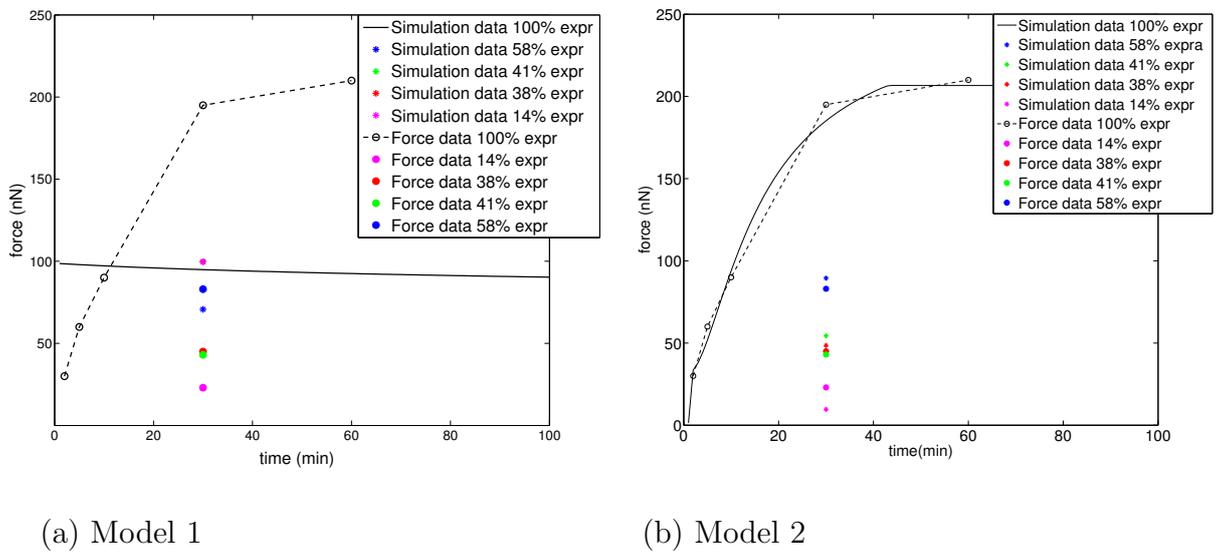
Figure 6.3: *Graphs showing the best fits of Model 1 and Model 2 to the data from Chu et al. [2004]. Plot (a) shows the best fit for Model 1 to the data. It can be seen that the simulation results for E-cadherin expression levels of 14% to 58% after 30 minutes are reasonably close to the data but the time course of the force for 100% E-cadherin expression does not fit the data. Plot (b) shows the best fit for Model 2 to the data. Here the force time course for 100% E-cadherin expression as well as the forces after 30 minutes at different expression levels fit the data well.*

very clearly that a good fit was found for Model 2 whereas the results of Model 1 produced a very poor fit to the data. This was surprising, as the main difference between the two models, the fact that E-cadherin–β-catenin complexes have to

be redistributed between cell-cell contact sites after the attachment of more cells in Model 2, is not of importance at this stage. However, given the clear result that the results of Model 2 provided a better fit to the data, we used this model as the intracellular pathway of cells in our multiscale model. Running simulations with two cells quickly showed however, that the model did not transfer well to the multiscale model. The two cells only showed initial adhesion but then the adhesion force displayed damped oscillatory behaviour until the cells separated. Therefore it was obvious that some of the assumptions must be wrong. In the multiscale simulations the intracellular dynamics of both cells started with the same initial conditions as shown in equations (6.11)-(6.14). This means that, in contrast to the assumptions made above, the percentage of E-cadhern–β-catenin complexes at the contact site in the neighbouring cell was variable and very low to start with. Thus this assumption had to be changed in the MATLAB code to find parameters that also produce a good fit to the data in the multiscale model. We therefore changed this part of the code so that instead of assigning $E/\beta_{\mathrm{c,i}}^{\mathrm{i}}$ a value of 15% in Model 1 and 80% in Model 2, it varied and was assigned the same value as was calculated for the percentage of E-cadherin–β-catenin complexes at the cell-cell contact site in the current cell of interest. This new MATLAB code was again run for both models as explained above, trying to find values for $\eta, \iota, \nu_n, \nu_p$ and ρ_c that give the best fit of the models to the data. Figure 6.4 shows the adhesion force between two cells produced by the new Model 1 (Figure 6.4(a)) and the new Model 2 (Figure 6.4(b)) using the sets of parameters that give the best fit to the data in each case. Although producing a better fit to the data than previously, it was clear that the results of Model 1 still did not fit the data well. The results of Model 2 gave a reasonable fit to the actual data points but it was obvious that the time course behaviour is very different to what one would expect and it is very unlikely that cells show this oscillatory pattern. Therefore
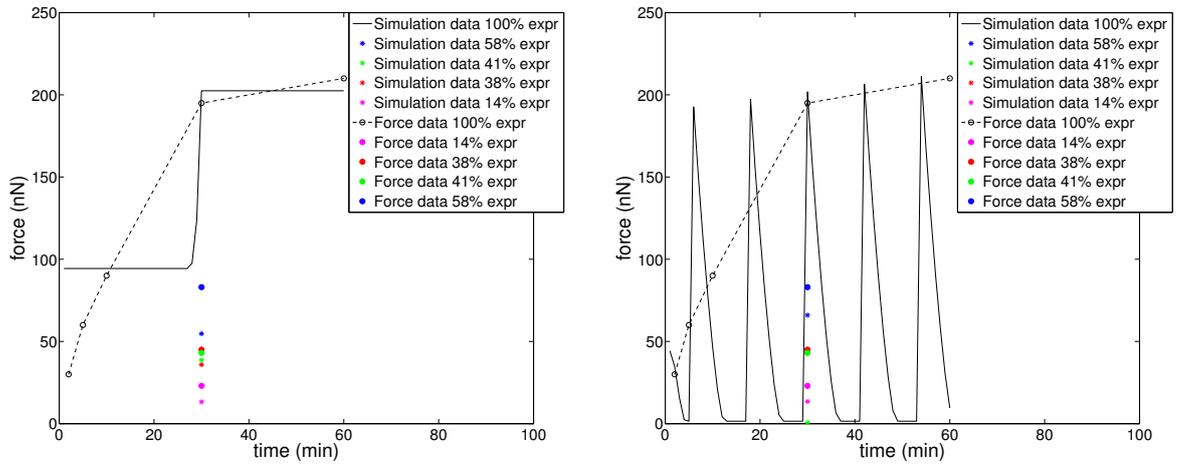
(a) Model 1   (b) Model 2

Figure 6.4: *Graphs showing the best fits of Model 1 and Model 2 without the assumption that 80% of E-cadherin are at the cell–cell contact site in the neighbouring cell to the data from Chu et al.. Plot (a) shows the best fit for Model 1 to the data and plot (b) shows the best fit for Model 2. In both cases the simulation results for E-cadherin expression levels of 14% to 58% after 30 minutes are reasonably close to the data but the time course of the force for 100% E-cadherin expression does not fit the data.*

it seemed as though another adjustment of the model had to be made.

The internalisation of E-cadherin–β-catenin complexes due to non-binding obviously has a strong impact on the model dynamics as the differences between the results in Figure 6.3 and Figure 6.4 are solely due to the fact that the value of $E/\beta^{\mathrm{i}}_{\mathrm{c,i}}$ in this term was changed. Although the term describes a process that is very likely to take place, its dynamics seem to be more complicated than captured here and are possibly coupled with a time delay. As not enough data are available to find the precise term necessary and in order to keep the model as simple as possible, we decided to remove that term from both models. Again using the MATLAB code from previous optimisations and only adjusting the models to

this newest version, the simulation results were fitted to the data to find optimal values for the parameters $\eta, \iota, \nu_n, \nu_p$ and ρ_c. Figure 6.5 shows the simulation results of both models using the optimal parameter set.

Similar to the previous optimisation and simulation results, Model 1 (Figure
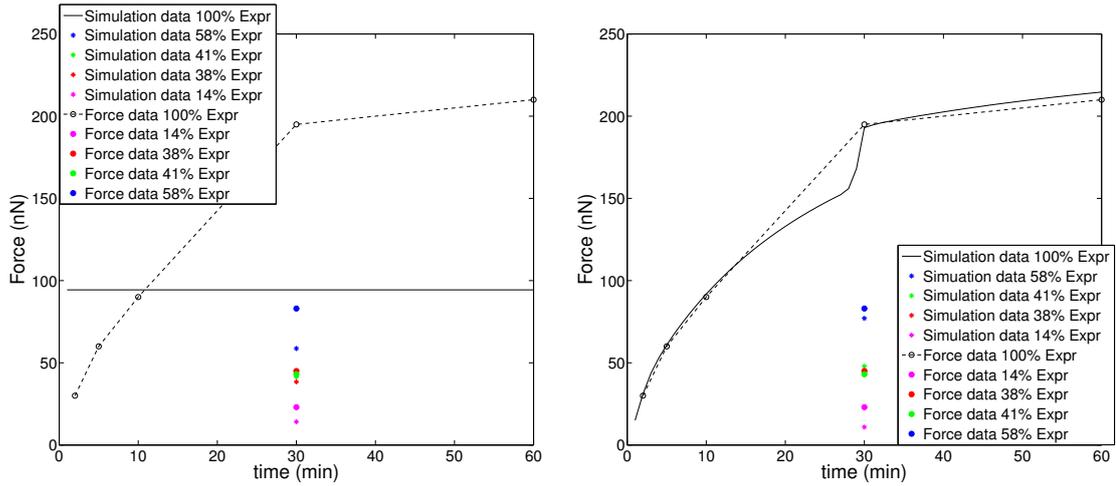
(a) Model 1   (b) Model 2

Figure 6.5: *Graphs showing the best fits of the new Models 1 and 2 to the data from Chu et al. [2004]. Plot (a) shows the best fit for Model 1 to the data. It can be seen that the simulation results for E-cadherin expression levels of 14% to 58% after 30 minutes are reasonably close to the data but the time course of the force for 100% E-cadherin expression does not fit the data. Plot (b) shows the best fit for Model 2 to the data. Here the tim course of the force for 100% E-cadherin expression as well as the forces after 30 minutes at different expression levels fit the data well.*

6.5(a)) produced a very poor fit to the force data. However, Model 2 (Figure 6.5(b)) produced cell-cell adhesion forces that fit the data well. The parameter

values found are:

$$\iota = 8.2/\text{min},$$

$$\nu_n = 0.6/\text{min},$$

$$\nu_p = 0.02/\text{min and}$$

$$\rho_c = 0.6/\text{min}.$$

In order to find an optimal value for the parameter γ as well, we ran another round of optimisations for Model 2, this time for three cells. As no data are published on the adhesion between cells in a group of three, we assumed that if all three cells come into contact at the same point in time, the adhesion force develops in the same way at both contact sites of a cell and that the adhesion force reached is half the force measured for two cells in contact by Chu et al. [2004]. Furthermore, we assumed that if initially only two cells were in contact and then a third cell came into contact with the cell of interest, the E-cadherin–β-catenin complexes would be redistributed between the two contact sites, such that after 30 minutes the adhesion forces would be equal between the cell of interest and both of its neighbouring cells.

Figure 6.6 shows the results of the simulations with three cells using the optimal value for γ which is given as:

$$\gamma = 0.16/\text{min}.$$

Figure 6.6(a) shows the adhesion force between a cell and both its neighbours which came into contact with the cell at the same point in time. It can be seen that the adhesion force generated by the cell in the simulations fits the assumed actual force well. Figure 6.6(b) shows the development of the adhesion force between a cell and both its neighbours where initially only one of the neighbouring cells is in contact with the cell of interest, and a strong adhesion force develops between the two cells. After 20 minutes the second cell comes into contact with
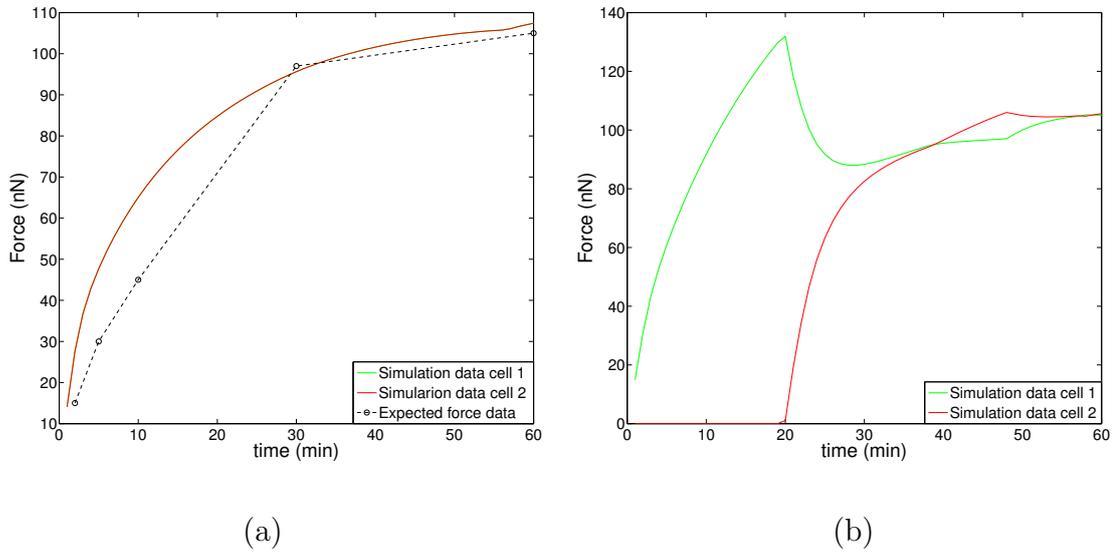
Figure 6.6: *Graphs showing the time course profiles of the forces at two contact sites with two different neighbours. Plot (a) shows the forces over time in the case where the two cells come into contact with the cell of interest at the same time. The expected force at both contact sites is taken to be half of the force measured by Chu et al. [2004]. The time course fits the expected data well. Plot (b) shows the forces over time in the case where initially only one cell contacts the cell of interest and after 20 minutes a second cell comes into contact with the cell. The E-cadherin–β-catenin complexes are redistributed quickly so that the force at both contact sites is very similar after 10 minutes and completely the same after 40 minutes.*

the cell of interest and the adhesion between the initial two cells decreases whereas the adhesion between the cell of interest and the new neighbour increases until the force at both adhesion sites is equal after about 40 minutes. It can be seen in 6.6(b) that the adhesion force between the cell of interest and the new neighbour continues to increase for another 10 minutes when the increase abruptly stops and the force levels out. The force between the cell of interest and the initial neighbour on the other hand increases slowly during these 10 minutes and

then has a sharp increase so the forces at the two contact sites are of the same strength. This behaviour is most likely due to the three processes (i.e. directed exocytosis, undirected exocytosis and redistribution) that lead to exocytosis of the E-cadherin–β-catenin complexes. The force at the contact site between the cell of interest and the new neighbour increases from the time point of contact onwards due to undirected exocytosis, directed exocytosis and the redistribution of complexes from the other contact site. 40 minutes into the simulation the forces at the two contact sites are the same and thus no more complexes are brought to the new contact site due to redistribution. Directed exocytosis however still takes place as the contact area is still growing. Thus the redistribution changes direction at this point so that complexes are shifted from the new cell-cell contact site to the old one. This happens at a very low rate as the density of complexes at both contact sites are presumably very similar. However, at some point the redistribution overcomes the directed exocytosis which stops the increase in contact area and therefore stops the directed exocytosis altogether. This is most likely the switching point in the time course profile of the forces which takes places roughly 50 minutes into the simulation. This then leads to the levelling out of the forces due to the redistribution of the E-cadherin–β-catenin complexes.

6.2.3 The E-cadherin–β-catenin pathway: the final model and its dynamics

The model, the simulation of which fit the data by Chu et al. [2004] (see Figure 6.2) best, is the E-cadherin–β-catenin model which will be used in the rest of this chapter. To recall, the equations governing the dynamics are:

$$\frac{dE}{dt} = -\nu_{\mathrm{p}} * E * \beta + \nu_{\mathrm{n}} * E/\beta + \sum_{\mathrm{i}=0}^{\#\mathrm{contacts}} d_{\mathrm{i}}(t) * E/\beta_{\mathrm{c}}^{\mathrm{i}},$$

$$\frac{d\beta}{dt} = -\nu_{\mathrm{p}} * E * \beta + \nu_{\mathrm{n}} * E/\beta + \sum_{\mathrm{i}=0}^{\#\mathrm{contacts}} d_{\mathrm{i}}(t) * E/\beta_{\mathrm{c}}^{\mathrm{i}},$$

$$\frac{dE/\beta}{dt} = \nu_{\mathrm{p}} * E * \beta - \nu_{\mathrm{n}} * E/\beta - \sum_{\mathrm{i}=0}^{\#\mathrm{contacts}} c_{\mathrm{i}}(t) * E/\beta,$$

$$\frac{dE/\beta_{\mathrm{c}}^{\mathrm{i}}}{dt} = c_{\mathrm{i}}(t) * E/\beta - d_{\mathrm{i}}(t) * E/\beta_{\mathrm{c}}^{\mathrm{i}} + C_+^{\mathrm{i}} - C_-^{\mathrm{i}} \ , \ \forall \mathrm{i},$$

with

$$d_{\mathrm{i}}(t) = \underbrace{a_{\mathrm{d,i}} * \rho_{\mathrm{d}}}_{\substack{\text{internalisation due to} \\ \text{junctional disassembly}}} \quad , \quad a_{\mathrm{d,i}} = \begin{cases} \|\frac{\partial}{\partial t}\hat{a}(t)_{\mathrm{i}}\|, & \text{if} \frac{\partial}{\partial t}\hat{a}(t)_{\mathrm{i}} < 0, \\ 0, & \text{otherwise}, \end{cases}$$

$$c_{\mathrm{i}}(t) = \underbrace{a_{\mathrm{c,i}} * \rho_{\mathrm{c}}}_{\text{directed exocytosis}} + \underbrace{\iota * \frac{\hat{a}(t)_{\mathrm{i}}}{4 * \pi * \mathrm{r}(t)^2}}_{\text{undirected exocytosis}} \quad , \quad a_{\mathrm{c,i}} = \begin{cases} \frac{\partial}{\partial t}\hat{a}(t)_{\mathrm{i}}, & \text{if} \frac{\partial}{\partial t}\hat{a}(t)_{\mathrm{i}} > 0, \\ 0, & \text{otherwise}, \end{cases}$$

and

$$C_+^{\mathrm{i}} = \gamma * \sum_{\mathrm{j}=1, \mathrm{j}\neq \mathrm{i}}^{\#\mathrm{contacts}} E/\beta_{\mathrm{c}}^{\mathrm{j}} * \left(1 - \frac{\frac{E/\beta_{\mathrm{c}}^{\mathrm{i}}}{\hat{a}(t)_{\mathrm{i}}}}{\frac{E/\beta_{\mathrm{c}}^{\mathrm{j}}}{\hat{a}(t)_{\mathrm{j}}}}\right),$$

$$C_-^{\mathrm{i}} = \gamma * E/\beta_{\mathrm{c}}^{\mathrm{i}} * \sum_{\mathrm{j}=1, \mathrm{j}\neq \mathrm{i}}^{\#\mathrm{contacts}} \left(1 - \frac{\frac{E/\beta_{\mathrm{c}}^{\mathrm{j}}}{\hat{a}(t)_{\mathrm{j}}}}{\frac{E/\beta_{\mathrm{c}}^{\mathrm{i}}}{\hat{a}(t)_{\mathrm{i}}}}\right).$$

The parameter values used in the following are given in Table 6.1.

The initial conditions are given in equations (6.11)-(6.14). ρ_{d} is estimated later in the multiscale simulations as the effect of this parameter can only be seen at the cell population level. In the following simulations it is initially set to 0. In order to understand the dynamics of the model completely, we ran simulation where 6 cells come into contact with the cell of interest at different points in time. We also tested scenarios where different numbers of cells attached themselves and

parameter		value
undirected E/β translocation rate	ι	8.2/min
E/β dissociation rate	ν_n	0.6/min
E/β binding rate	ν_p	0.02/min
directed E/β translocation rate	ρ_c	0.6/min
$E/\beta_\mathrm{c}^\mathrm{i}$ redistribution rate	γ	0.16/min

Table 6.1: *Table showing parameter values used in the final version of the intracellular E-cadherin–β-catenin model.*

detached themselves from the cell of interest. The simulation results are shown in Figure 6.7 in terms of the force between the cell of interest and its neighbours. Figure 6.7(a) shows the force at different contact sites during and after the consecutive attachment of six cells to the cell of interest. It can be seen that independent of the length of time between attachments, the E-cadherin–β-catenin complexes were redistributed successfully such that the forces were equal at all cell-cell contact sites at the end of the simulation. Figure 6.7(b) shows the forces at different contact sites during and after attachment and detachment processes as well as the detachment of two cells simultaneously. Initially all six neighbours were attached to the cell of interest. After 30 minutes we forced one of the cells to detach itself and it can be seen in the plots that the E-cadherin–β-catenin complexes that had been forming the bonds at this cell-cell contact site, were internalised and then recycled to enforce the other cell-cell attachments. This becomes clear as the graphs depicting the force at the different cell-cell contact sites show a step-like increase in their values. The same can be seen for further cell detachments after different time intervals. After the detachment of 4 cells (cells 3-6), 80 minutes into the simulation, we forced cell number 3 to reattach itself to the cell of interest and the graphs show the expected behaviour of the redistribution of force. At 110 minutes, we forced two cells to detach themselves
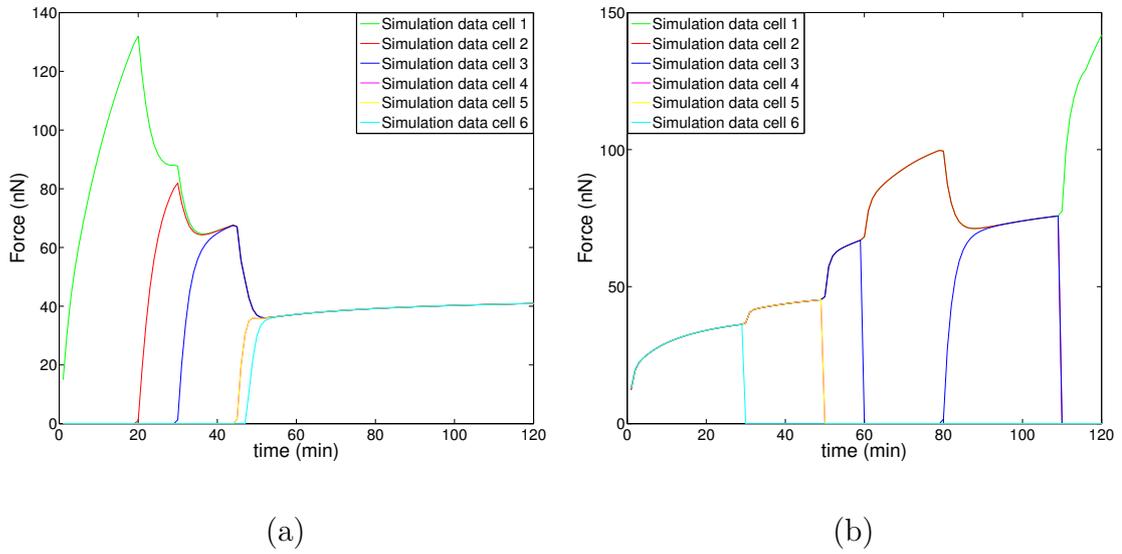
Figure 6.7: *Graphs showing the time course of the forces at different cell–cell contact sites after multiple cell–cell attachment and detachment processes. Plots in figure (a) show the forces at different contact sites during and after the consecutive attachment of six cells to the cell of interest. Plots in figure (b) show the forces at different contact sites during and after attachment and detachment processes as well as the detachment of two cells simultaneously.*

and again, the graph of the final neighbour left, shows a smooth increase in force at this cell-cell contact site.

These simulations show that the model exhibits the behaviour one would expect to see at cell-cell contact sites during and after attachment and detachment processes. It also fits the cell-cell adhesion data by Chu et al. [2004] well and is stable in a variety of scenarios. Thus we will use this model as the intracellular component of the multiscale model in the following.

6.3 The multiscale cell model

The model we introduce here combines the intracellular dynamics from the previous section and a cell-level model to produce a multiscale model of individual cells as well as cell-cell interactions. The cell-level model we will formulate is again similar to the one found in the work by Ramis-Conde et al. [2008b]. However the model of the cell division is modified so that the growth curve of a small cell population is a smooth exponential curve rather than a step-like function. The cell-cell interactions are then modelled using the extended Hertz model as explained in Chapter 3.2.

6.3.1 Modelling the cell

One biomarker of cells in an epithelial layer is the apical-basal polarity and the columnar shape as explained in Chapter 2. Epithelial cells in isolation are generally of a spherical shape [Galle et al., 2005]. As we intend to model epithelial cells specifically, a very detailed model would include a shape change of the cells from a spherical shape of cells in isolation to a columnar shape of adhering cells to a hemispherical shape of migrating cells which have emerged from the EMT. However, so that we do no over-complicate the model, we assume a spherical cell shape. Furthermore we assume the cells to be elastic as has been done in previous modelling work [Ramis-Conde et al., 2008b, 2009, Drasdo and Hoehme, 2005]. This allows for a near columnar shape for cells in epithelial layers. In order to be able to both use the migration model developed earlier and rely on its previous validation and results, we choose the radius of the cell to be 5μm. This ensures the volume to be roughly the same as the volume of the cells in the previous chapters. In addition we ran 2 sets of 15 simulations each of the single cell migration model in Chapter 4 over six hours of real time to investigate the effect of having a spherical rather than a hemispherical cell shape. In one of

these sets we used a hemispherical cell shape with a base radius of 15μm (as in Chapters 4 and 5) and in the other set we used a spherical cell shape with a radius of 5μm. All the model parameters were set to their default values which means that the 15,000 matrix fibres were randomly orientated and of a mean length of 75μm. The stiffness was calculated independently for each fibre. The maximum cell speed was set at 20μm/h, the standard deviation of $\mathbf{f}_j(t)$ was set at 0.01 and the standard deviation used in the calculation of the cell's polarity axis was set at 1. Figure 6.8 shows the results of these simulations. The box plot summarising
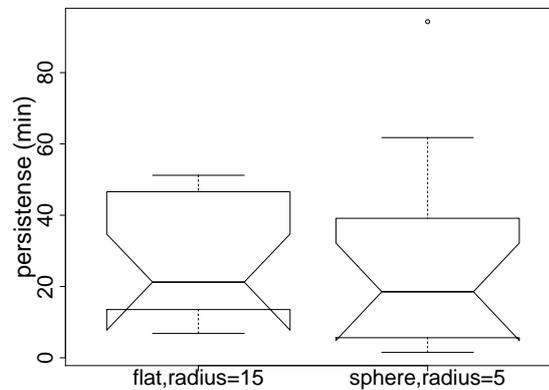
Figure 6.8: *Notched box plots showing the persistence times of single cells migrating on a 2D matrix. The left box plot summarises the results of simulations where a cell is modelled as being relatively flat and hemispherical with a base radius of 15μm. The box plot on the right shows the results of simulations where a cell is modelled as being spherical with a radius of 5μm. The notches indicate that there is no statistically significant difference between the medians of the two sets of simulations.*

the results of the simulations with a spherical cell with a radius of 5μm shows a slightly larger range of persistence times than the box plot representing the simulation results with a hemispherical cell. The median persistence time of the latter set of simulations is slightly higher than that of the former one, however,

117

the notches of both box plots are overlapping. This means that the difference of the median value is not statistically significant. Thus the simplification of the model, which provides the possibility of avoiding the change of shape of the cell during the EMT, does not have a significant impact on the cell migration characteristics.

6.3.2 Modelling cell-cell interactions

We model the cell-cell interactions by the extended Hertz model as explained in Chapter 3.2. The cell–cell adhesive forces are governed by the E-cadherin–β-catenin complexes on the surface of two neighbouring cells at the contact site. The smaller number of complexes on either cell's surface at the contact site determines how strong the adhesion is, as it limits the number of bonds that can be formed. This number of E-cadherin–β-catenin complexes involved in cell-cell adhesion bonds is given as a percentage of the maximum number of complexes that can theoretically be formed (= 100%). As explained earlier, the number of bonds between two cells is translated into a cell-cell adhesion force using the data by Chu et al. [2004]. We again assume that the force, resulting from 80% of E-cadherin–β-catenin complexes forming bonds with a neighbouring cell, is 210nN as was measured by Chu et al. [2004]. Thus the adhesion force between two cells is calculated by equation (6.10). The cell-cell repulsive forces, on the other hand, are calculated using the Hertz model. As explained in Chapter 3.2 the Hertz model calculates the repulsive force between two spherical objects dependent on their biophysical properties. The parameters used for these properties are given in Table 6.2. The resulting force between two cells is calculated by the extended Hertz model using equations (3.5) and (3.6).

parameter		value	reference
Radius of a cell	R	5μm	[Galle et al., 2005], [Ramis-Conde et al., 2008b]
Poisson ratio of cell i	σ_i	1/3	[Galle et al., 2005], [Ramis-Conde et al., 2008b]
Elastic modulus of cell i	E_i	1 kPa	[Galle et al., 2005], [Ramis-Conde et al., 2008b]

Table 6.2: *Table showing the parameter values used to calculate the repulsive force between two cells with the Hertz model.*

6.3.3 Inclusion of the intracellular dynamics in a two-cell model

As mentioned above, the cell-cell adhesion forces are governed by the number of E-cadherin–β-catenin bonds between cells which in turn is determined by the number of E-cadherin–β-catenin complexes on the surface of each of the two cells at the site of contact. This number of complexes at the contact site of two cells is governed by the intracellular E-cadherin–β-catenin dynamics derived in Section 6.2. Thus we included these intracellular dynamics into the cell model explained above and ran simulations with 2 of these cells. The movement of the cells was calculated using equation (3.4) and, as only cell-cell interactions were considered as present here, all terms on the right hand side except this one were ignored. This gave the force balance equation $\mathbf{F}_{drag} = \sum_{innj} \mathbf{F}_{ij}$. $\mathbf{F}_{drag}$ was calculated as explained in Chapter 3 but taking into account the spherical cell shape. $\sum_{innj} \mathbf{F}_{ij}$ results from the extended Hertz model which is calculated as explained in the previous section. In the simulations we tracked the time course development of the intracellular components as well as that of the two cells. The results are shown in Figure 6.9. The top row of plots shows the development over 40 minutes at the cellular level. In the figures it can be seen how the positions of the cells relative to each other changes until a steady-state configuration is reached where the repulsive
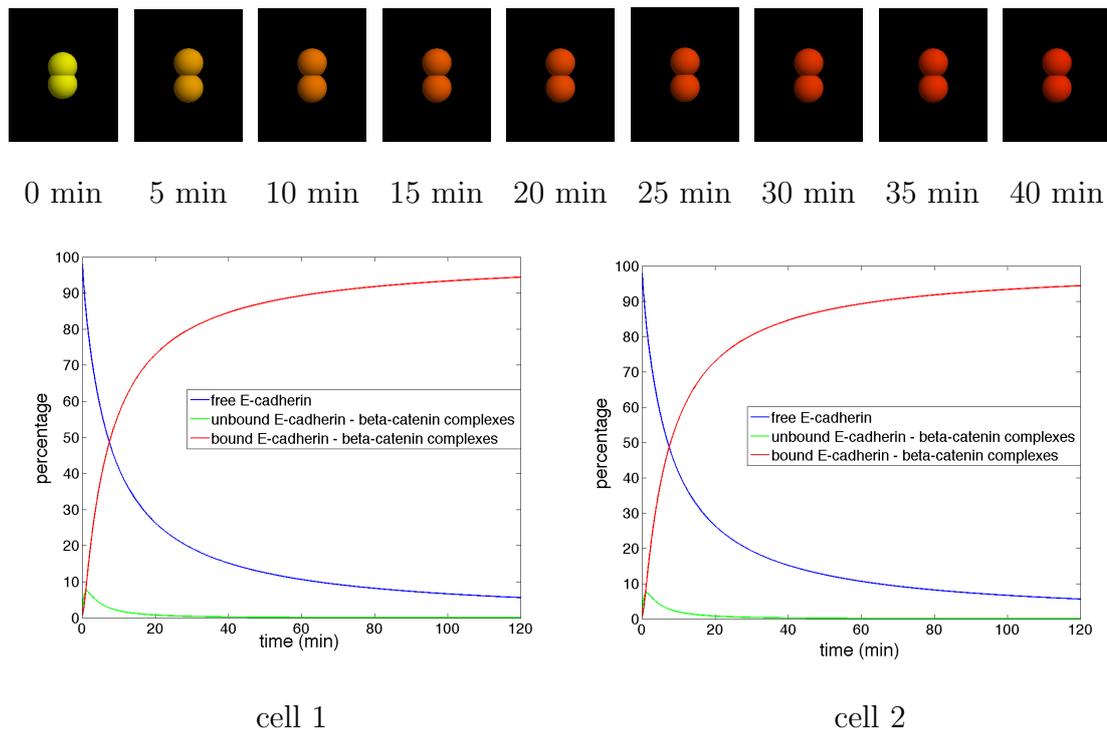
Figure 6.9: *Figures showing the time course development of the multiscale cell model at the cell as well as the intracellular level in a simulation with two cells. In the top row of plots, E-cadherin in the cytosol is shown by the intensity of the yellow colour. Thus the uptake of E-cadherin in cell-cell bonds can be followed over time through the colour change of the cells in the figure from yellow to red. At the same time it can be seen that the distance between the mid-points of the two cells changes until the two cells are at a steady-state configuration. The bottom row of plots shows the intracellular dynamics in both cells during these simulations.*

and adhesive forces balance each other out. At the same time these plots show the amount of E-cadherin that is taken up in bonds at the cell-cell contact site - the intensity of the yellow colour in the otherwise red cells is proportional to the amount of E-cadherin in the cytosol. Thus the redder the image of the cell is, the less E-cadherin is in the cytosol and the more E-cadherin is at the cell-cell contact

site. The time course of the E-cadherin dynamics is also shown in the plots in the bottom row of the figure. Here the dynamics of the intracellular components in both cells are shown over 120 minutes. It can be seen that the dynamics are the same in both cells. After initial complex formation and a fast increase of E-cadherin–β-catenin complexes at the contact site, the dynamics slow down to reach a steady-state. These simulations show that the intracellular dynamics are well-modelled in the cell level model and provide a multiscale model of cell-cell interaction dynamics.

6.3.4   Modelling cell division

In order to model whole cell populations, or colonies, it is important to not only take cell-cell interactions into account but also cell division. We model the cell cycle by assuming that the G1-phase has an average length of 7 hours, the M-phase an average length of 2 hours and G0, S and G2 together have a length uniformly distributed between 8 and 18 hours. This gives an overall cell cycle length of 17- 27 hours which agrees well with published cell cycle lengths [Bernard and Herzel, 2006, Drasdo and Hoehme, 2005]. Each time a cell enters the cell cycle a new G0-S-G2 time is calculated for this cell to take into account the randomness of biological systems. When a cell enters M-phase, it starts dividing into two daughter cells. The division occurs along the axis of highest pressure and at the end of M-phase, the division is complete and two new cells of radius R/$2^{\frac{1}{3}}$ exist. The value of this new radius is chosen for volume conservation. This is implemented by generating a new cell at the start of M-phase and by placing the centre of the new cell immediately next to the centre of the original mother cell. The original mother cell becomes the second daughter cell. Both daughter cells have initially the same radius as the mother cell had. The same is the case for the amount of intracellular proteins. During the 120 minutes of M-phase, the

centres of the two daughter cells move further apart by a distance of 1.6 * time step(min)/120 min * $R/2^{\frac{1}{3}}\mu$m per time step. This ensures that the centres have a distance of 1.6 * $R/2^{\frac{1}{3}}\mu$m from one another at the completion of M-phase. A cell's radius shrinks monotonously by ((R- $R/2^{\frac{1}{3}}$) * time step(min)/120min)μm in order to ensure a radius of $R/2^{\frac{1}{3}}\mu$m at the end of M-phase. This is done to capture the slow shape change during the division of the mother cell into two daughter cells. Since in reality the two daughter cells are not two completely separate cells until the end of the M-phase, no cell-cell interactions are assumed to take place. However, interactions with other cells can occur. The intracellular dynamics and the adhesive and repulsive forces are calculated for all contacts between one of the daughter cells and its neighbouring cells. Before calculating the net force at these cell-cell contact sites the neighbouring cells are however checked to see whether they are also in contact with the other daughter cell. If this is the case, the force is halved as it would otherwise be counted twice when considering the two daughter cells as the one mother cell. In addition, after the two daughter cells have been moved according to these forces, they are checked to ensure that they have not moved further apart than they should. If this is not the case they are moved closer together. This adjustment is done to each cell proportional to the movement they have just completed.

During G1 phase both daughter cells then grow up to their maximum radius R. When cells are taken to exist on an extracellular matrix, a further consideration has to be taken into account. It has been found for glioma cells that motile cells have a decreased proliferation rate or do not proliferate at all [Giese et al., 2003]. This has come to be known as the 'go or grow' mechanism or the migration/proliferation dichotomy. Although this has not been explicitly shown for epihelial cells, the underlying molecular reasons given by Giese et al. [2003], also hold for other cells than glioma cells. Therefore we also assume that cells with

front-rear polarity, and thus the motile cells, cannot enter M-phase.

6.3.5   Constraints of the modelling technique

The general way by which the force that is generated by bonds between two cells is measured, is by trying to separate the two cells and record the force it takes to do this. Chu et al. [2004] call this force a 'separation force' which is indeed exactly what it is. However, it is often also referred to as an 'adhesion force' as it is also seen as the force that cells generate to adhere to each other. Adams et al. [1996] showed the accumulation of E-cadherin at cell-cell contact sites over time and the 'zipping up' of these contact sites. The E-cadherin bonds were clearly exerting an adhesion force which led to the movement of the cells closer to each other and thus an increase in the cell-cell contact site. Adams et al. [1998] show images of the distribution of E-cadherin molecules tagged with green fluorescent protein (GFP) during the formation of a monolayer. Here it can be seen that although E-cadherin bonds form all along cell-cell contact sites, they seem to be represented in higher densities at the end-points of the contact site. These are the points where the cell-cell dynamics happen. The E-cadherin bonds here either prevent the 'unzipping' and separation of the cells or try to extend the contact site. If it is assumed that the natural state of epithelial cells is in a layer or sheet, as this is the way they occur in the body, then the E-cadherin bonds try to restore this natural state by preventing cell-cell separation. It can be assumed that cells with mature cell-cell contacts are in their natural state and either no force or little force is used in this case to keep the cells together. For that to be true, the repulsive force originating from the cell's cytoskeleton also has to be minimal in this case. This assumption can be made as otherwise the cell would have to produce enormous amounts of energy just to keep itself in its natural state which is very unlikely. Thus the so called 'adhesion force' is only exerted when two cells

are separated and their natural state is perturbed. Therefore calling it 'separation force' may be more precise. The extended-Hertz modelling technique however, models adhesion rather than separation forces. Here it is assumed that the cells constantly exert adhesion and repulsion forces on one another. The assumption of the model is that adhesive and repulsive forces balance each other out in cells at an optimal distance from one another (see Figure 3.3). However, this is not necessarily the case. One can also assume that with an increase in the cell-cell contact site, the number of cell-cell adhesion bonds increases which in turn leads to stronger adhesion and thus a further increase in the cell-cell contact site and so on. Although the repulsion also increases with cell-cell closeness, the adhesion can be larger until the two cells engulf each other. This is exactly what happened in some test-simulations we ran. One could argue that it is due to a wrong choice of parameters, but given that all the parameter values are either taken directly from the literature available or are derived from fitting the intracellular dynamics to the data, it rather highlights the problems arising from using potential functions to model cell-cell interactions. In order to overcome these problems and adjust the model to be more in line with the underlying biology, we assume that the force between two cells is zero if the cells are in their natural state. Given that a cell in a monolayer has on average six neighbours, we assume that the force between two cells is zero if the cell-cell contact area has a diameter larger than or equal to one sixth of the cells circumference. However as soon as the contact area decreases, the separation force/adhesion force and repulsion force are calculated as explained above in Section 6.3.2.

6.3.6 The estimation of ρ_d

The only part missing in the multiscale cell model that we develop here, is the value of the parameter ρ_d. This parameter describes the rate at which E-cadherin–β-catenin complexes are endocytosed after bond breakage due to the separation of two cells. Its effect can therefore only be seen in a multi-cell simulation where dynamic detachment processes take place. For this reason we ran simulations varying this parameter by five orders of magnitude to get some idea of the value of this parameter before starting a more refined parameter search. We examined the simulation results after three days and after seven days. To evaluate the results we took note of the number of cells at these two time points, the number of neighbours each cell had and the average adhesion force at the cell-cell contact sites. Unfortunately we did not have any specific cell-colony data to compare the results with. However, some basic assumptions helped to compare the simulation results and decide on a good parameter value.

Firstly, without enforcing a particular growth law, we assumed that the cell colony should grow considerably between day three and day seven. Secondly, we assumed that the cell colony should have a near-circular shape. Thirdly, we used histograms which show the distribution of the number of neighbours of cells in a proliferating epithelial layer [Gibson et al., 2006, Sanderius et al., 2011], to compare the results with. The histograms are shown in Figure 6.10. It can be seen in both images that the distributions are very similar for this large variety of species right across the range of the Metazoa. This means that it is a very stable pattern and we assumed, therefore, that it is very similar in humans as well. Thus we used this as the main marker to distinguish between a "good" and a "not-so-good" parameter value. In addition to varying the value of ρ_d, we also varied some constraints that influence the model dynamics. In the model
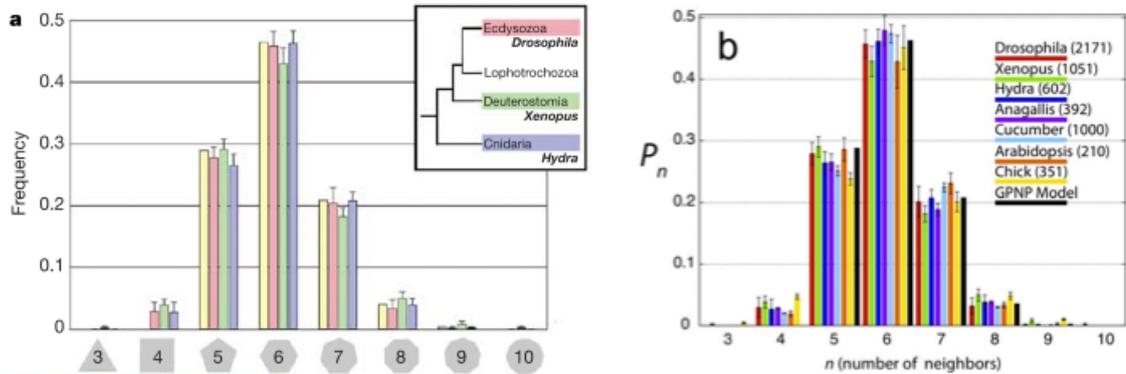
Figure 6.10: *Graphs showing the distribution of the number of neighbours of cells in growing epithelial layers. (a) The distribution of the number of neighbours in the proliferating metazoan epithelia of Drosophila, Xenopus and Hydra. (b) The distribution of the number of neighbours of cells in a much broader range of proliferating epithelia. In addition to the epithelia of Drosophila, Xenopus and Hydra it also shows the distribution for epithelia of Anagallis, cucumber, Arabidopsis, and chick (also the results of a model presented in that paper are shown in black). Image (a) is reproduced with copyright permission from Gibson et al. [2006]. Image (b) is reproduced from Sanderius et al. [2011] in accordance with the Creative Commons Attribution License.*

described so far, the cells grow and divide independently of one another. However, as explained in the introduction, the cell cycle can be arrested, or paused, if the conditions for division are not right, e.g. if the cell is under stress. Such stress could be cell compression. Thus, in one set of the simulations, we assumed that, if the cell is exerting a repulsive force above a certain threshold, the cell's division cycle is paused and it does not enter M-phase. The threshold force chosen is 13000pN which translates into a cell having six neighbours each with a distance of $\approx 8.5\mu$m from cell-midpoint to cell-midpoint. The second change we made in one of the simulation sets, was to change the force between two cells

which are considered to be in their natural state. Instead of assuming that the force is zero between these cells, we assumed that the adhesion force is zero and thus the overall force equalled the repulsive force in this case unless the two cells were positioned relative to each other so that the diameter of the contact area was exactly one sixth of the cell's circumference. As well as having sets of simulations with these two changes separately, we also ran a set of simulations with both changes. Thus, in total we ran four sets of simulations, in each of which we varied the value of ρ_d between 500 and 0.005 by one order of magnitude at a time. The first set we called *'original'* and in this we used no constraints on the cell division and a force of zero between cells that are in their natural state or closer than that. The second set we called *'division constraints'*, as here the division of a cell was constrained by the pressure it was under. The third set has the name *'repulsion btw cells closer than natural'* to note that in these simulations the force was set to be equal to the repulsion in cells closer than their natural state. And finally the fourth set is called *'repulsion btw cells closer than natural and division constraints'* to note that here both changes were made. The results, first after three days and then after seven days, are shown in Table 6.3. Whenever a row does not have any entries, that means that those simulations failed due to the cells getting too close. As in Figure 6.9, the colours of the cells in the plots in the second column of the table, are related to the amount of free E-cadherin which is shown in yellow.

Table 6.3: *Table showing the results of varying ρ_d, the constraints under which cells can enter M-phase and the force that cells generate between them if they are closer than their natural state.*

ρ_d	Image	#cells	#neighbours	average force
Behaviour after 3 days				
	original			
500		18		0pN
50		23		13502.69pN
5		22		4781.9pN
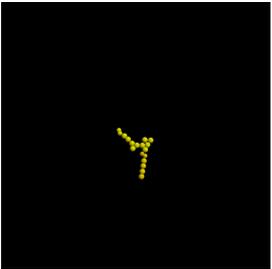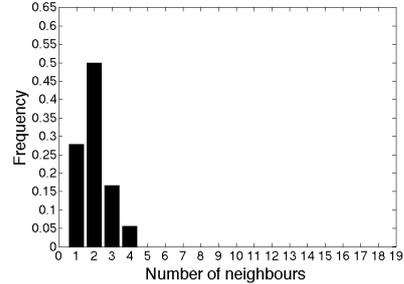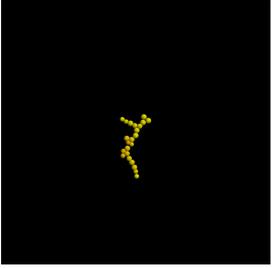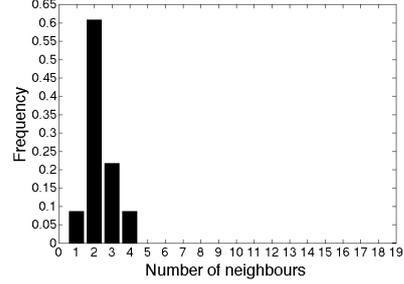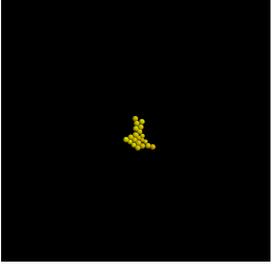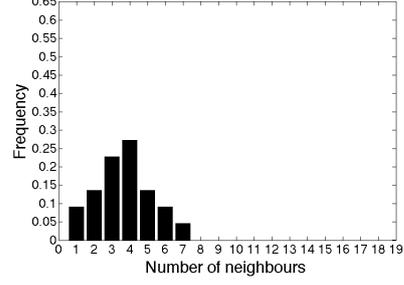
Table continues on next page...
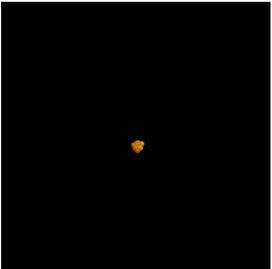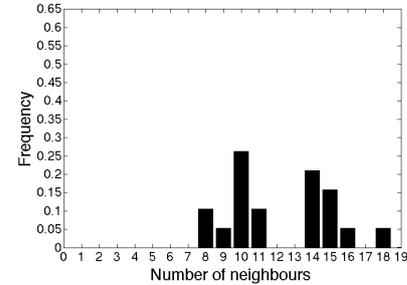
ρ_d	Image	#cells	#neighbours	average force
0.5		19		7895.95pN
0.05	–	–	–	–
0.005	–	–	–	–
	division constraints			
500		18		0pN
50		23		13502.69pN
5		23		9857.29pN
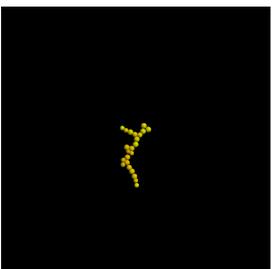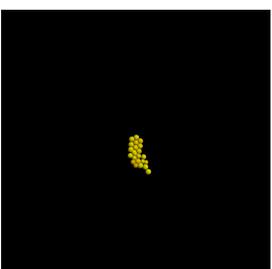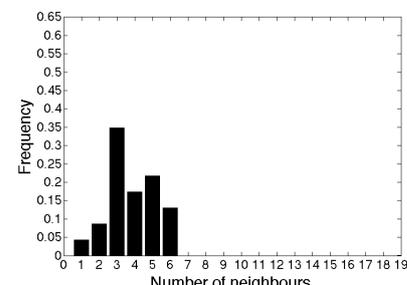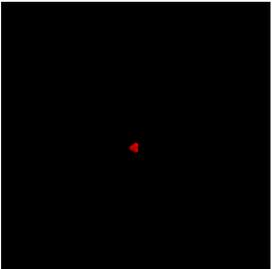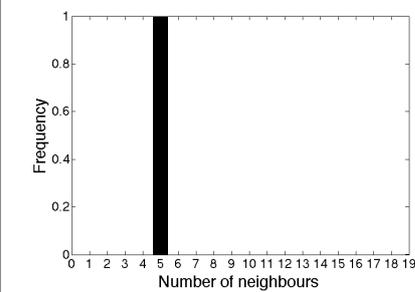
ρ_d	Image	#cells	#neighbours	average force
0.5		6		52154.77pN
0.05		5		65414.28pN
0.005	–	–	–	–
	repulsion btw cells closer than natural			
500		22		3079.32pN
50		24		0pN
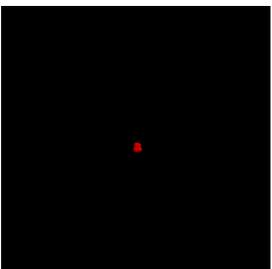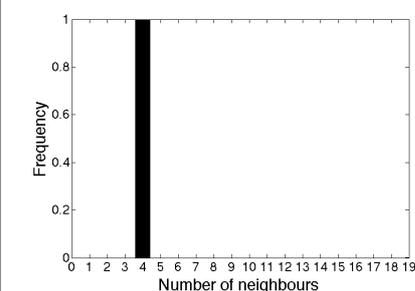
ρ_d	Image	#cells	#neighbours	average force
5		21		323.56pN
0.5		24		5071.39pN
0.05	–	–	–	
0.005	–	–	–	–
	repulsion btw cells closer than natural and division contraints			
500		22		3079.32pN
50		24		6296.73pN
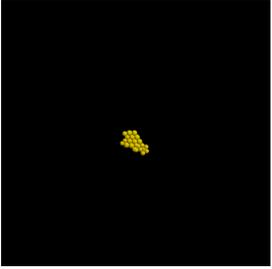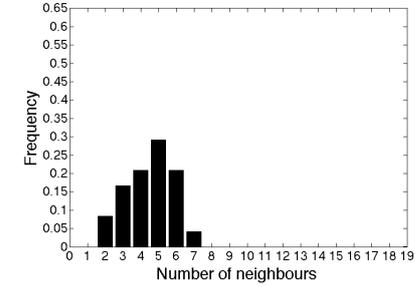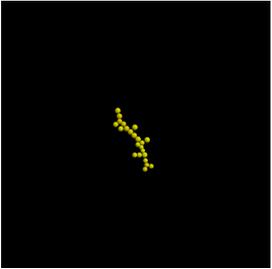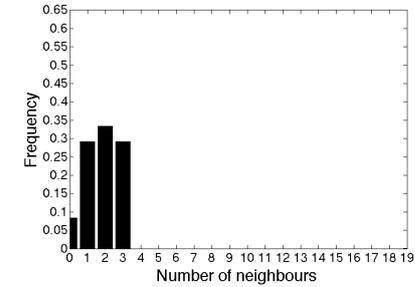
Table continues on next page...

131

ρ_d	Image	#cells	#neighbours	average force
5		21		3782.98pN
0.5		22		4525.54pN
0.05		11		20528.58pN
0.005		7		37182.90pN

ρ_d	Image	#cells	#neighbours	average force
Behaviour after 7 days				
	original			
500		552		4531.67pN
50		528		4808.33pN
5	–	–	–	–
0.5	–	–	–	–
0.05	–	–	–	–
0.005	–	–	–	–
	division constraints			
500		560		5080.04pN
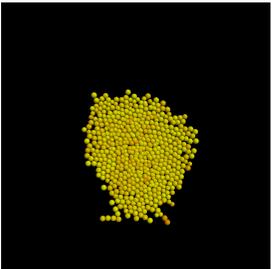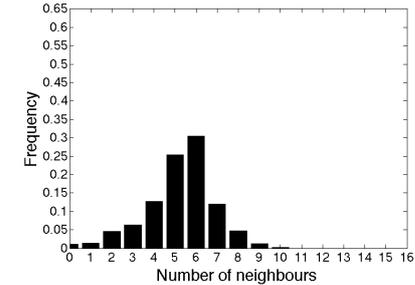
133

ρ_d	Image	#cells	#neighbours	average force
50		648		5167.65pN
5		613		4611.23pN
0.5		6		52109.05pN
0.05		5		65454.69pN
0.005	—	—	—	—
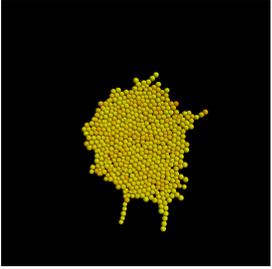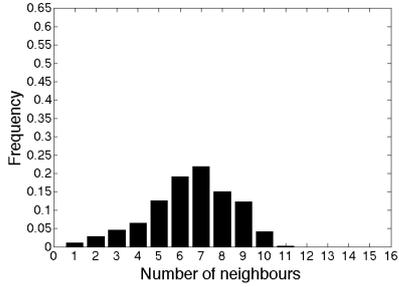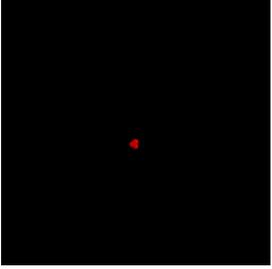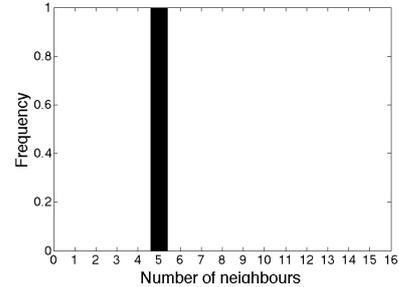
ρ_d	Image	#cells	#neighbours	average force
	repulsion btw cells closer than natural			
500		659		1770.12pN
50		650		1723.86pN
5		631		1956.04pN
0.5		640		6680.74pN
0.05	—	—	—	
0.005	—	—	—	—
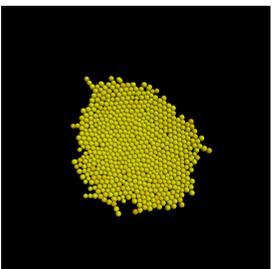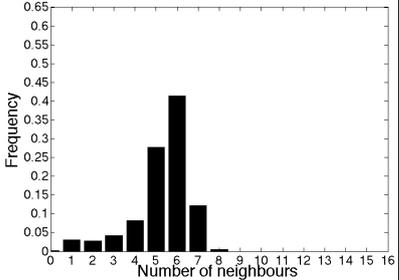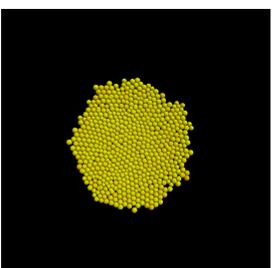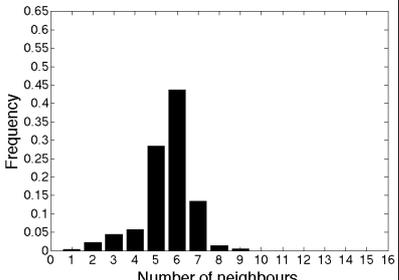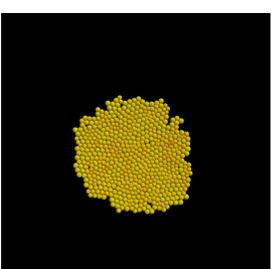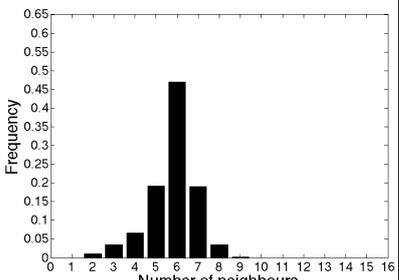
Table continues on next page...

ρ_d	Image	#cells	#neighbours	average force
	repulsion btw cells closer than natural and division contraints			
500		659		1783.20pN
50		639		1937.0pN
5		597		1972.87pN
0.5		581		6392.6pN
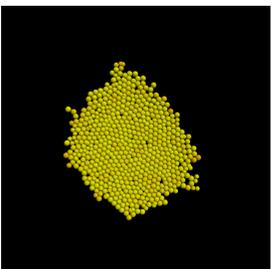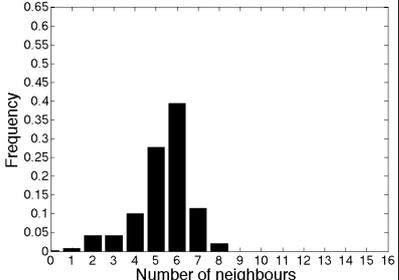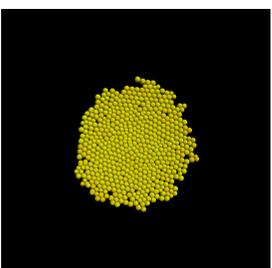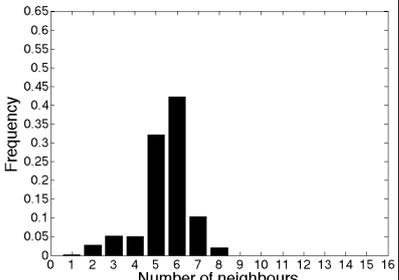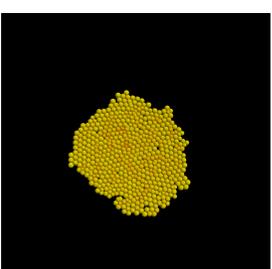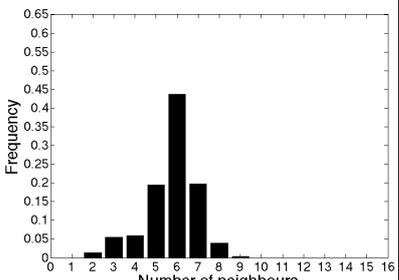
Table continues on next page...
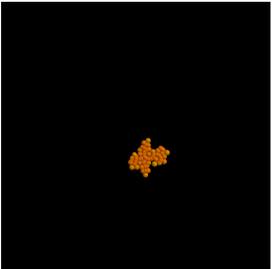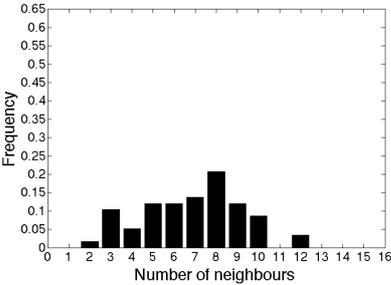
ρ_d	Image	#cells	#neighbours	average force
0.05		58		18931.96pN
0.005		7		37380.47pN
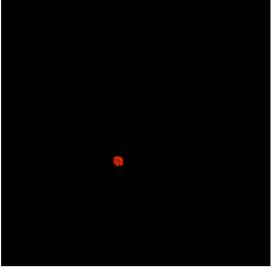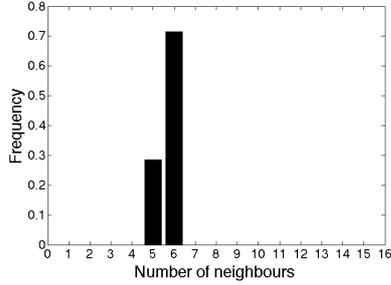
The results in Table 6.3 show that a huge variation of behaviour can be observed in the simulations and even in each individual set of simulations by changing just this one parameter ρ_d. The most obvious overall result is that when assuming the active force is just the repulsive force in cells closer than their natural state and, even more so, constraining the division of cells dependent on the pressure they are under, one can observe the behaviour for the whole range of parameter values, whereas otherwise, the cells get too close and cause the code to fail when ρ_d takes on small values. The colour-coding used shows that in all the sets of simulations the amount of free E-cadherin is much lower for small ρ_d values than it is for larger ones.

When considering the behaviour after three days, values of 500 and 50 for ρ_d lead to very extended shapes of the cell colonies and to only a small number of neighbours, between 1 and 4 for each cell. For small parameter values of 0.5, 0.05 and 0.005 the behaviour is very much dependent on the set of simulations. Whenever

the simulations could finish though, the colonies look to be circular, generally with a relatively low number of cells. For parameter values of 5, the shapes of the colony are somewhere between circular and extended. In the original set of simulations, the simulations failed for parameter values of 0.005 and 0.05. For a value of 0.5 it can already be seen that the cells had unnaturally high numbers of neighbours which can only result from them being squashed together. The image of the cell colony for a ρ_d value of 5 already looks very irregular and for values of 50 and 500 the colonies are very extended.

When division constraints were introduced in the second set of simulations, again very extended colonies were the result of the simulations with parameter values 500 and 50. Also the colony resulting from a parameter value of 5 once again had a very irregular shape. However, the simulations using parameter values of 0.5 and 0.05 could both finish in this set and they both showed a stable circle of 6 and 5 cells which all had 5 and 4 neighbours, respectively.

In the third set of simulations, extended and irregular shapes of the cell colonies could again be observed for parameter values of 5, 50 and 500. Also the image of the simulation results using a value of 0.5 for ρ_d shows a slightly irregular shape although the irregularity decreases and the number of neighbours the cells have increases with decreasing parameter value. The simulations for parameter values of 0.05 and 0.005 did not complete.

The fourth set of simulations, which had both, a constraint on the cell division and a force set to the repulsive force for cells closer than their natural state should allow, was the only set of simulations where the results after three days could be gathered for all parameter values. Again extended and irregular shapes of the cell colonies were found for parameter values of 5, 50 and 500. For values of 0.5, 0.05 and 0.005 the cell colonies were near circular. However, as previously seen in simulations with values of 0.5 and 0.05, the number of cells was very small in

the simulations with parameter values of 0.05 and 0.005.

Based not only on the colony shape and the number of cells but also excluding simulations that have an extreme distribution of the number of neighbours, the most favourable combination of parameter value and constraints after three days is the one used in the fourth simulation of the last set of simulations. Here, a value of 0.5 is used for ρ_d, a constraint is put on the cell division and the active force between two cells that are closer than their assumed natural state is just the repulsive force.

For the original set, only the simulations using parameter values of 500 and 50 finished running over seven days. The colonies in both simulations had grown considerably between day three and day seven. Both generally look to be near circular, but they do have areas of extended rows of individual cells. This is mirrored in the distribution of neighbours with about 5% of cells having no, or only one, neighbour.

In the next set of simulations, which included the constraints on cell division, the colonies generated using the parameter values 5, 50 and 500 have even more irregular shapes than the results from the first set of simulations. Especially the colony resulting from using a parameter value of 5 in the simulations appears almost to be an invasive colony. There was no growth at all between day three and day seven in the simulations that used ρ_d values of 0.5 and 0.05. In the third set of simulations all the colonies look near circular. For the higher parameter values of 500 and 50, a few irregularities can be seen and there are still cells that are not in contact with any other cell. However, the simulations resulting from using parameter values of 5 or 0.5, produce good shapes and the distribution of neighbours is also similar to those shown in Figure 6.10.

The final set of simulations shows a very clear decrease in free E-cadherin for decreasing parameter values. Similar to the third set of simulations, the overall

geometry of the colonies is almost circular for parameter values 0.005, 0.5, 5, 50 and 500. The only real irregular shape can be seen for a value of 0.05. However, in both the simulations with a ρ_d value of 0.005 and 0.05, the amount of free E-cadherin is very low and the colonies have hardly grown or not grown at all since day three. In addition, the distribution of the number of neighbours the cells have is very different to the ones shown in Figure 6.10. For high parameter values of 50 and 500, the circumferences of the cell colonies have slight irregularities and again some cells exist that are not in contact with any other cell. The results of the simulations using the parameter values 5 and 0.5 both show relatively circular cell colonies as well as a distribution of the number of neighbours that is comparable to the ones shown in Figure 6.10. The distribution resulting from simulations with ρ_d equal to 0.5 however gives a slightly better fit to the distributions found from experiments, as the difference between the frequency with which 5 and 7 neighbours are observed is not so drastic. There is also less free E-cadherin present in the cells which is what one would expect.

Given that this combination of parameter value and changes to the original model was also the most favourable in the results of the simulations after three days, we examined parameter values around 0.5 more closely again, using the constraint on the cell division as well as assuming a force equal to the repulsive force if cells get closer than their assumed natural state. The results of varying ρ_d between 0.1 and 0.9 using increments of 0.1 are given in Table 6.4. In contrast to the results presented in Table 6.3, unsurprisingly only little variation can be seen between the results of the simulations with the different parameter values.

Table 6.4: *Table showing the results of varying ρ_d between 0.1 and 0.9 with increments of 0.1 and using both, constraints on cell division and a force equal to the repulsive force if cells get closer than their assumed natural state*

ρ_d	Image	#cells	#neighbours	average force
		Behaviour after 3 days		
0.9		19		3553.33pN
0.8		22		3970.7pN
0.7		24		4000.27pN
0.6		22		4375.13pN
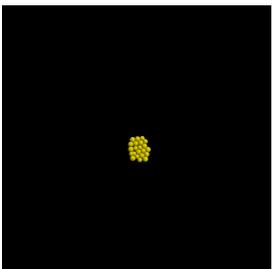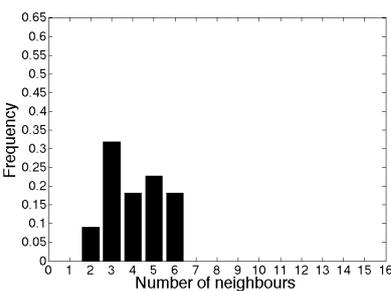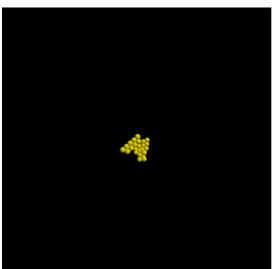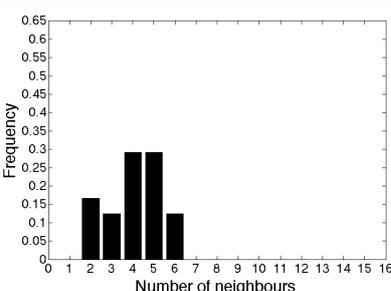
ρ_d	Image	#cells	#neighbours	average force
0.5		22		4525.54pN
0.4		19		5500.54pN
0.3		19		5920.43pN
0.2		19		8250.23pN
0.1		12		15786.24pN

142

ρ_d	Image	#cells	#neighbours	average force
			Behaviour after 7 days	
0.9		594		3898.86pN
0.8		647		4503.34pN
0.7		662		4905.28pN
0.6		661		5714.5pN
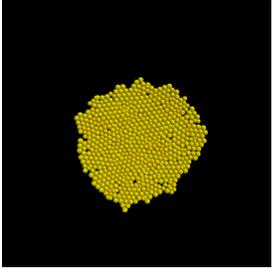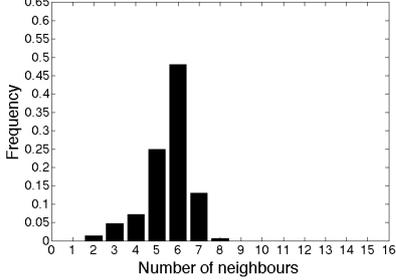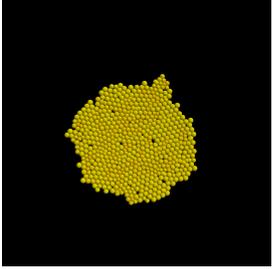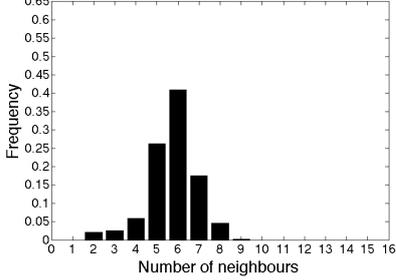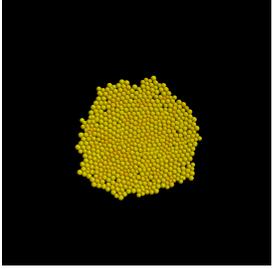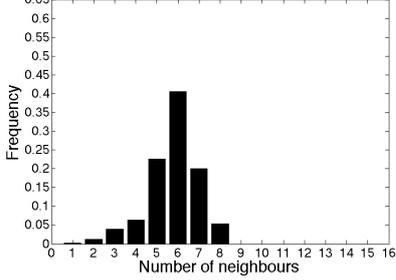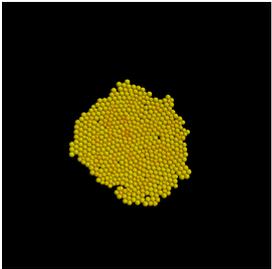
ρ_d	Image	#cells	#neighbours	average force
0.5		581		6392.59pN
0.4		569		7413.07pN
0.3		543		8672.28pN
0.2	–	–	–	–
0.1		93		14794.69pN
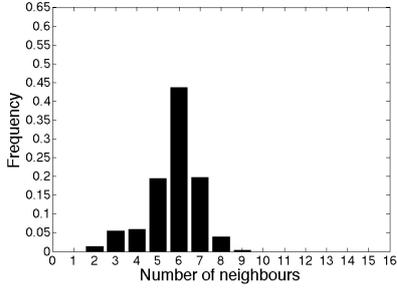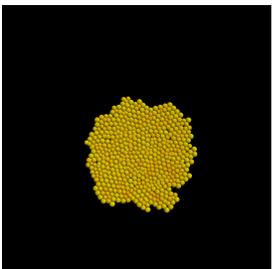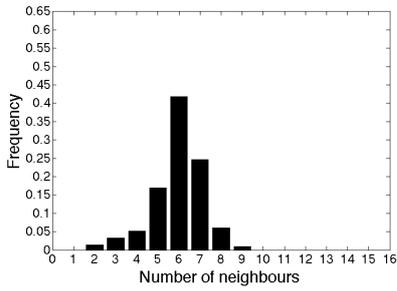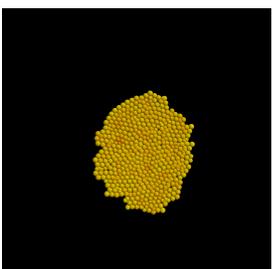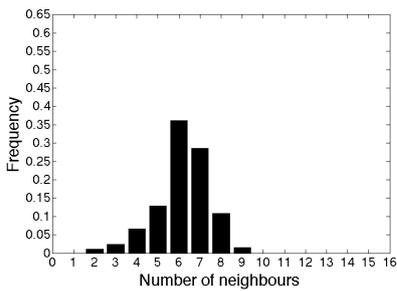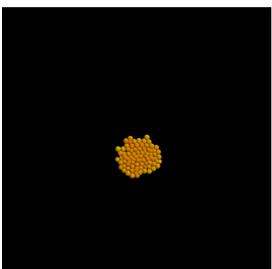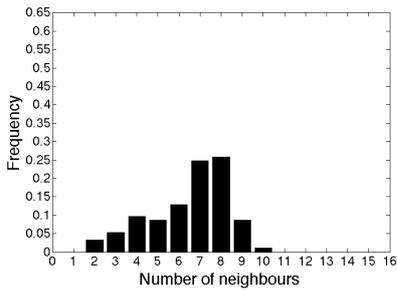
The results for the parameter value 0.1 show unexpected distributions in the number of neighbours cells have. The results of the simulation where a parameter value of 0.2 was used, do not exist because the simulations failed. Apart from these two cases, the results show nearly circular colonies both after three days and

after seven days. However, a careful comparison of the distributions of the number of neighbours the cells have after seven days shows that the distribution resulting from the simulation with a ρ_d value of 0.6 is closest to the distribution found through experiments in Figure 6.10. The only difference between the distributions is that a small number of cells in the simulation have between 1 and 3 neighbours which is lower than what can be seen in the experimental findings. This is most likely due to the fact that in the simulation all the cells are counted whereas in the experiments only a segment of the inner part of the proliferating epithelial layer is considered. Thus the low number of neighbours arises most likely from cells at the colony's circumference. Other than that, the results of the simulation fit the experimental finding both qualitatively and quantitatively.

As mentioned above, we assumed that the cell colony should grow considerably between day three and day seven. In this particular simulation this was clearly the case. In order to get a more precise idea of the growth law underlying these results, we plotted the number of cells against time. The results can be seen in Figure 6.11. The results of the simulation are shown in red. It can be seen that the curve they form looks like an exponential growth curve. When trying to fit an exponential curve to the results by varying the growth constant, we found that the curve of the function

$$y(t) = 2 \times e^{0.83t}, \ t \text{ in days} \tag{6.15}$$

fits the results well. This can be seen in Figure 6.11 where the curve of equation (6.15) is shown in black. This agrees with the well-established assumption that cell monolayers grow exponentially on short time scales [Freshney, 1992, Bru et al., 2003, Drasdo and Hoehme, 2005]. Given the description in Freshney [1992] and Drasdo and Hoehme [2005], seven days are seen to be a short time-scale and within the exponential growth regime. Thus concerning multiple different measures, this parameter choice and the choice of model changes, show results that

145
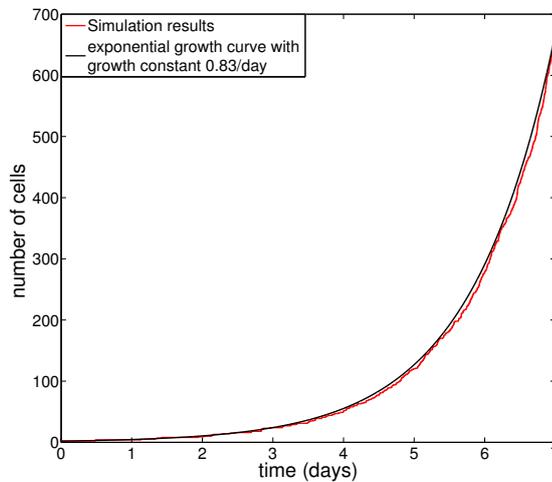
Figure 6.11: *Graph showing the growth of the cell colony arising from the simulation where ρ_d is set to 0.6, cell division is constrained and the force is set to equal the repulsive force between two cells closer than their natural state should allow. The growth curve of the simulated cell colony is plotted in red. This is very close to an exponential growth curve with growth rate 0.83/day which is shown in black.*

fit the data well. It is surprising, however, to see that in order to obtain these fits, the parameters of a simulation have to be chosen where the amount of free E-cadherin is relatively high and the average force at cell-cell contact sites is only just under 6000pN. Given the single-cell adhesion/separation experiments by Chu et al. [2004] and extrapolating to a cell that has more than one neighbour, one would expect a much higher cell-cell adhesion/separation force and hardly any free E-cadherin. To our knowledge there are unfortunately no data in the available literature that we can compare this with as, although experiments that use fluorescence staining of E-cadherin exist, they are difficult to quantify. Thus this comparatively low separation force in cells in the epithelial monolayer compared to the separation force of two individual epithelial cells, can only be hypothesised to exist and it would be interesting to find a way of setting up experiments that

can test this. For further simulations we will use the parameter value of 0.6 for ρ_d as well as the constraints on the cell division and the adjustment of the force between two cells that are closer than their natural state is assumed to allow.

6.4 Modelling the extracellular matrix and cell-matrix interactions

To complete the model, the extracellular matrix and the cell-matrix interactions are modelled as in previous chapters. The individual matrix fibres are again represented by thin cylinders, the lengths of which are normally distributed with mean 75μm and standard deviation 5μm and the widths are 200nm. Cells also reorientate the matrix fibres as before and the fibre rotation is again calculated using equation (4.13). Here the stiffness of the matrix fibres is always calculated dependent on fibre inter-connectedness. All the parameters are chosen as in Chapters 4 and 5, i.e. the number of fibre cross-links at which the maximum stiffness is reached is given the value of 15, the factor that reduces the reorientation is given the value of 0.1, the standard deviation used to generate $\mathbf{f}_j(t)$ is 0.01, the standard deviation used to generate χ is 1 and s_{max} is $20\mu m/h$. Similar to the model in the previous chapter, the matrix reorientation by one cell can be constrained by the presence of other cells. If multiple cells are in contact with the same fibre, we assume again that a cell should not be able to reorient that fibre in such a way that it loses contact with the other cells. Thus, after calculating the reorientation angle, the other cells are checked to see whether or not they are still in contact with the fibre. If they are in contact, the reorientation takes place as before. Otherwise, the reorientation is reduced such that the other cells stay in contact with the fibre.

6.5 Modelling the cell movement

Just as before, the cell movement is governed by the total force acting on the individual cell. Here, in addition to the force the cell generates to move itself forward on the matrix, the drag force and the random movement, the cell-cell interactions have to be taken into account. Thus the equation of motion from previous chapters has to be extended to

$$\mathbf{F}_{drag} = \sum_{innj} \mathbf{F}_{ij} + \sum_{fnnj} \mathbf{F}_{fj} + \mathbf{f}_j(t),\tag{6.16}$$

where $\mathbf{F}_{ij}$ is the force generated between two cells i and j with the sum taken over all cells that cell j is in contact with, $\mathbf{F}_{fj}$ is the force generated by an individual cell through contact with an individual matrix fibre, with the sum taken over the fibres which are in contact with the cell, and $\mathbf{f}_j(t)$ is the term accounting for noise.

6.6 Computational simulation algorithm

Similar to the previous simulations, the time step was chosen to be three seconds and the procedure during each time step can be summarised as follows:

Step 1:

Each fibre is examined to see whether cells were in contact with it during the previous time step and whether those cells are polarised. If both of these conditions are true for at least one cell, then it means that the cell(s) has exerted force on the fibre during that time step. All the fibres for which this is the case are reorientated by one cell at a time in the way explained in equation (4.13). After each reorientation, the fibre is checked to see that it has not lost contact with another cell and, if necessary, the reorientation is decreased.

Step 2:

All the fibres that the cell is in contact with are found for each cell that is not in the M-phase and it is determined whether the cell is in contact with any other cell.

Option 1: If the cell is not in contact with any other cell and the cell in question is in contact with at least one fibre and has established front-rear polarity, either through previous polarisation or through the new formation of focal contacts over 10 minutes, the new polarity axis is calculated using equation (4.5).

Option 2: If the cell is in contact with at least one other cell and the cell in question has front-rear polarity, defining the rear to be where the contacts with all the other cells are, and the cell is in contact with at least one fibre, the new polarity axis is calculated using equation (4.5).

Option 3: If the cell is in contact with at least one other cell and the cell in question has front-rear polarity, defining the front to be where the contact with at least one other cell is, the cell loses its polarity. If those cells at the front of the cell were in contact with the cell in question in the previous time step and the cell in question has already lost its polarity and is re-polarising, then the direction of the fibres w_i in equation (4.5) is chosen to be pointing away from the contacting cell(s).

Step 3:

For each cell all its neighbouring cells are found and their distances are noted. Then the intracellular dynamics are run and the adhesive force is calculated for each cell-cell contact site using equation (6.10). The repulsive force is also calculated for each cell-cell contact site and, using equations (3.5) and (3.6), the cell-cell interaction force is calculated. If cells are closer than their natural state should allow, this force is set equal to the repulsive force alone.

If the cell has front-rear polarity, the net force that it generates for its movement

on the matrix is also calculated. This is done using the knowledge from the previous step of how many fibres the cell is in contact with. Together with the given parameter determining the maximum speed the cell can reach, the calculation of how fast the cell should be moving (see Figure 4.2(a)) and the magnitude of the force the cell must therefore be generating (see Figure 4.2(b)) can be made. Equation (4.6) is then used to calculate the net force.

Step 4:

Each cell is moved according to the forces calculated in Step 3. This is done by first solving equation (6.16) for the cell velocity and then applying the Forward Euler method. This gives the new position of the cells at the end of this time step.

Step 5:

For each cell that is in the M-phase, the distance that the two daughter cells' centres should have from the each other in order to give a smooth cell division, is calculated and their movement is adjusted accordingly.

Step 6:

Each cell is checked to see whether it should be entering M-phase during this time step and if this is the case, the cell is marked as being in this phase and the daughter cell is generated.

Each cell that is already in M-phase is checked to see whether this cell cycle phase should have finished in this time step. If this is the case the cell is reset to not being in M-phase and a new G0-S-G2 time is generated for it. If this is not the case, the cell's radius is decreased accordingly.

Each cell that is in G1 phase is checked to see whether its radius is smaller than R. If this is the case, the radius is increased by $(R-R/2^{\frac{1}{3}})/8400$.

6.7　Data analysis

The idea behind carrying out the simulations explained above, is to find individual parameters or parameter combinations that can transform a normal epithelial cell colony into an invasive colony that reflects the behaviour of cancer cells. Just as in the previous chapters, 15 simulations are run for each combination of parameters. In each of these 15 simulations a different seed is used for the random number generator for the cell movement. The data generated by the simulations is a list of all the cells and their positions in each time step. In order to be able to compare these data of the different sets of simulations objectively, it is important to introduce ways of measuring the results. Probably the most distinctive characteristics of a cell colony are how fast it grows and whether it grows as one cohesive colony or whether it splits into small groups or individual cells just as cancer cells do when they invade the tissue. Thus, here we introduce two algorithms which we use to analyse the data concerning those two characteristics.

6.7.1　Colony growth analysis

Figure 6.11 shows the exponential growth of a simulated cell colony. In that simulation, the colony started with 2 cells and grew up to a number of 661 cells. This is still within the range of a colony size which has also experimentally been shown to follow an exponential growth law. However, as a colony grows larger, it has been shown that it switches from an exponential to a linear growth law [Freshney, 1992, Bru et al., 2003, Drasdo and Hoehme, 2005]. In our simulations we start with a colony of 50 cells in order to be able to study a cell colony of considerable size. This means that it is possible that also in the simulations the colony does not grow exponentially the entire time but switches to a linear growth regime. In order to take this into account, we try to fit an exponential growth rate, a switching time, a linear growth rate and a y-intercept to the growth of the colony

in each simulation. The MATLAB code used starts off by reading in the number of cells that exists in each time step. Then we generate random numbers as starting points for the fitting process. For the exponential growth rate we generate a random number between 0 and 1, for the switching point a number between 3 and 7 days, for the linear growth rate a number between 1,000 and 10,000 and for the y-intercept a number between -30,000 and 0. As mentioned previously, the cell cycle times are uniformly distributed between 17 and 27 hours, or 0.7083 and 1.125 days. Thus if the entire cell population had the fastest cell cycle possible, the population would double in 0.7083 days and therefore the exponential growth constant would be $\frac{\ln(2)}{0.7083}$/day which is equal to 0.9786/day. Given that the cell cycle times of most of the population will be longer and also cells might enter G0 for longer periods of time, the exponential growth rate will be smaller than 0.9786 and thus approximately smaller than 1. As the population is growing the growth constant will be greater than 0. This gives the upper and lower limit of 1 and 0, respectively, for the exponential growth constant.

Similarly, estimates for the upper and lower limits of the linear growth constant and the y-intercept can be found. Assuming that the smallest number of cells at which exponential growth switches to linear growth is 1,000 and that the entire cell population has the fastest cell cycle time possible of 17 hours, the switching point is reached after approximately 3 days. If the cell population was to grow exponentially at that rate over the whole 7 days it would reach a cell number of approximately 47,000. This is of course a large overestimation. However, the fitting of a linear growth function to the population between day 3 (when there are 1,000 cells) and day 7 (when there are 47,000 cells), gives a linear growth constant of approximately 11,000 and a y-intercept of -32,000. Thus the value used as an upper limit for the linear growth constant is 10,000 and the values used as a lower limit for the y-intercept it is -30,000. The lower bound for the

linear growth constant is chosen to be 1,000 because the cell number reached after 7 days is more likely to be an order of magnitude smaller than it would be in an exponentially growing population. In addition, the time it will take the population to grow up to 1000 cells will most likely be longer and the switching point might happen at a larger number of cells which in turn would increase the linear growth constant. Thus we can assume that it will be between 1,000 and 10,000.

Unlike in previous chapters we use a constrained optimisation routine. Thus the estimated upper and lower limits of the parameters are not only used for the generation of initial guesses for the fitting process but are also used as upper and lower bounds in the optimisation routine. After generating starting points, the optimisation routine *'fminsearchbnd'* (by John d'Errico, MATLAB file exchange) is used to find optimal parameter values. Here, the set of optimal parameter values is defined to minimise the largest residual between the simulated population size over time and the population size calculated by the generated growth function. This procedure is carried out for 30 initial guesses and then the parameter set, which gives the overall minimal largest residual in the 30 optimisations, is chosen as the optimal parameter set. This is done to eliminate the possibility of choosing a parameter set that leads to a local rather than a global minimum.

In order to be able to to compare the results for the different simulation sets, the average and standard deviation are calculated for all four parameters found for each set of 15 simulations. We ensure that the average and the standard deviation represent the data well by generating a log-log plot for each set of simulations in which the simulated number of cells is plotted against the number of cells calculated by the growth function resulting from the average values of the exponential growth constant, the switching time, the linear growth constant and the y-intercept. The number of cells is plotted at five uniformly distributed

points in time, generated by the MATLAB random number generator *'rand'*, for each of the 15 simulations in the set.

6.7.2 Colony dispersal and spread analysis

For the analysis of the colony's dispersal and its spread, only the data from the final time step are considered. Thus the MATLAB code used reads in the cell position and radius data from that time step. Then clusters of cells are identified using the idea of hierarchical clustering. Initially each cell is seen as a cluster. Then a while-loop is started which, in each run, loops over the number of clusters and fuses two clusters together if they have at least one cell-cell contact site. This is done until all the clusters in the list are spatially separate and thus no more fusions can be made. The number of the resulting clusters gives a measurement of the dispersal of the colony. During the clustering process the number of cells per cluster is also saved which gives further information of the cell colony's characteristics. In order to classify the clusters by their number of cells, the code also calculates the frequency of clusters with one cell, clusters of less than or exactly 10 cells and so on for each order of magnitude. Furthermore the minimum distance from cell centre to cell centre can be calculated between each pair of clusters. From this the code calculates the maximum distance and the average distance between clusters which gives a measurement of the spread of the cell colony.

Once again all of these calculations are carried out for each simulation in each set of 15. Then the average and the standard deviation are calculated for each set regarding the number of clusters, the largest distance between clusters and the average distance between clusters. The distribution of the number of cells per cluster (i.e. cluster size) is calculated from all 15 simulations in one set.

6.8 Computational simulation results

For all the simulations, a group of 50 non-polarised cells and 15000 fibres were initially placed in the 1000x1000μm domain. This initial setup is displayed in Figure 6.12 where the cells are shown in red and the matrix fibres are shown in yellow. The cells were left to polarise, migrate, divide and interact over 7 days of real time
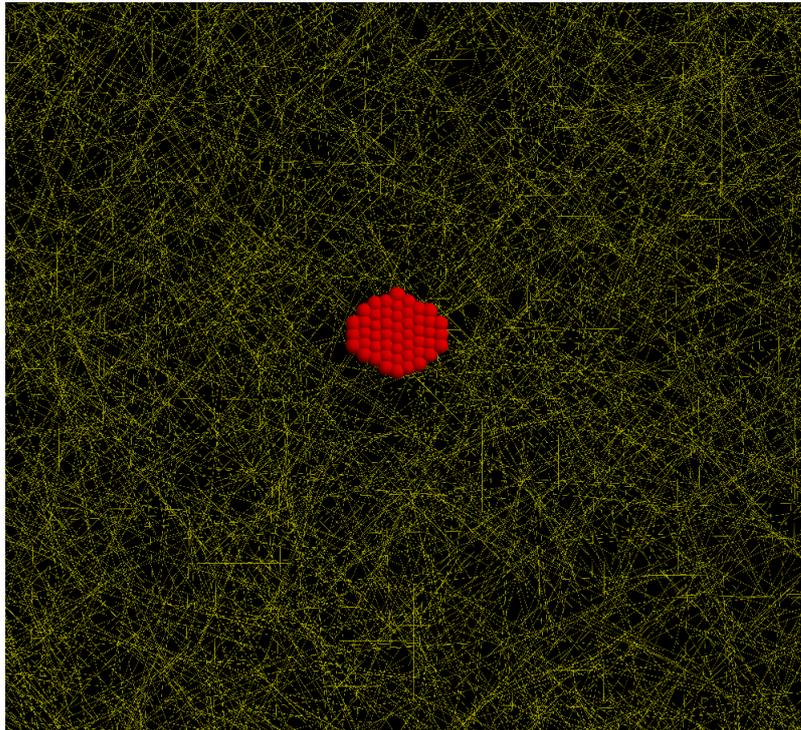
Figure 6.12: *Image of the central region of the initial setup for the simulations. 50 cells are positioned in the middle of the domain and 15000 fibres are placed throughout the domain to make up a layer of matrix.*

according to the equations of the model and the simulation-specific parameters. The simulations were run on two 2.26 GHz quadcore Intel Xeon processors. 15 simulations could be run simultaneously. Depending on the simulation-specific parameters these simulations took between four days and two weeks. Due to this computational complexity and the time scope of this thesis, we were only able to look at a subset of all the possible combinations of the parameters. We

155

considered cell populations with a homogeneous E-cadherin expression level of 75% and 50% as well as cell populations with heterogeneous E-cadherin expression levels between 100% and 0%. For each of the E-cadherin expression profiles, we increased the pressure threshold under which the cells can still divide from 13000pN to 20125pN to 36966pN to 56916pN. A pressure threshold of 13000pN is equivalent to a cell with a radius of 5μm having six neighbours of the same size each at a cell-centre to cell-centre distance of 8.5μm. A pressure threshold of 20125pN is equivalent to a cell having six neighbours each at a distance of 8μm. A threshold of 36966pN is equivalent to a cell having six neighbours each at a distance of 7μm and a pressure threshold of 56916pN is equivalent to a cell having six neighbours each at a distance of 6μm. Thus the increase in the pressure threshold allows cells to proliferate under increasing stress resulting from a lack of space. One of the characteristics of cancer cells is that they proliferate under conditions that normal, healthy cells would not proliferate under, e.g. space constraints. Changing this parameter therefore allows us to change the cell phenotype to a more aggressive, cancer-like phenotype. Combining this with another process that occurs during the EMT in cancer development, the down-regulation of E-cadherin, gave us the opportunity to study not only the impact of the individual factors but also their combined effect on cell colony development. Finally, for all of the different combinations of E-cadherin expression profiles and the pressure threshold for division, we ran a set of 15 simulations each with a 10% integrin expression and a 50% integrin expression level. We considered this variation in integrin expression as, again, it is well known that up-regulation of integrins plays a role in cancer invasion. For all of these simulations the analysis was done as described in Section 6.7.

6.8.1 The development of cell colonies with a homogeneous E-cadherin expression level of 75%

First we considered cell populations in which all cells express E-cadherin at a 75% level (compared to the expression levels in Section 6.1). In these populations we studied the effect of increased ability of the cells to proliferate under pressure and combined this with different levels of integrin expression. For each combination of parameters we again ran 15 simulations each of which used a different seed for the random number generator that controls the cell movement. All the simulations of the first set, which used a threshold on the pressure of 56916pN under which the cells could still divide and an integrin expression level of 10%, failed due to the cells coming too close. This shows that in a setting where the cells are free to move in any direction, these cells would most likely move out of the layer. Depending on whether or not they could survive without anchorage to the matrix, the development could lead to hyperplastic phenomena. However, as our focus lies on two dimensional cell colony development, we do not study this here. In the following we can therefore not include this combination of parameters. For the remaining 7 sets we observed the colony development over a period of 7 days and the configuration of the colony at the end of the 7th day. Figure 6.13 shows images of the cell colonies for the first simulation of each set after 7 days. The images show the entire domain. The matrix fibres are not shown as they are too small in size to be properly detectable in the images. As in Sections 6.3.3 and 6.3.6, the intensity of the yellow colour in the otherwise red cells depicts the amount of E-cadherin in the cytosol. Only very little difference can be seen between the images of the different simulations. In the following we thus first study the colony growth and then the spread and dispersal for the different sets of simulations.

10% integrin expression   **50% integrin expression**

56916pN   (a) ——   (b)

36966pN   (c)   (d)

20124pN   (e)   (f)

13000pN   (g)   (h)
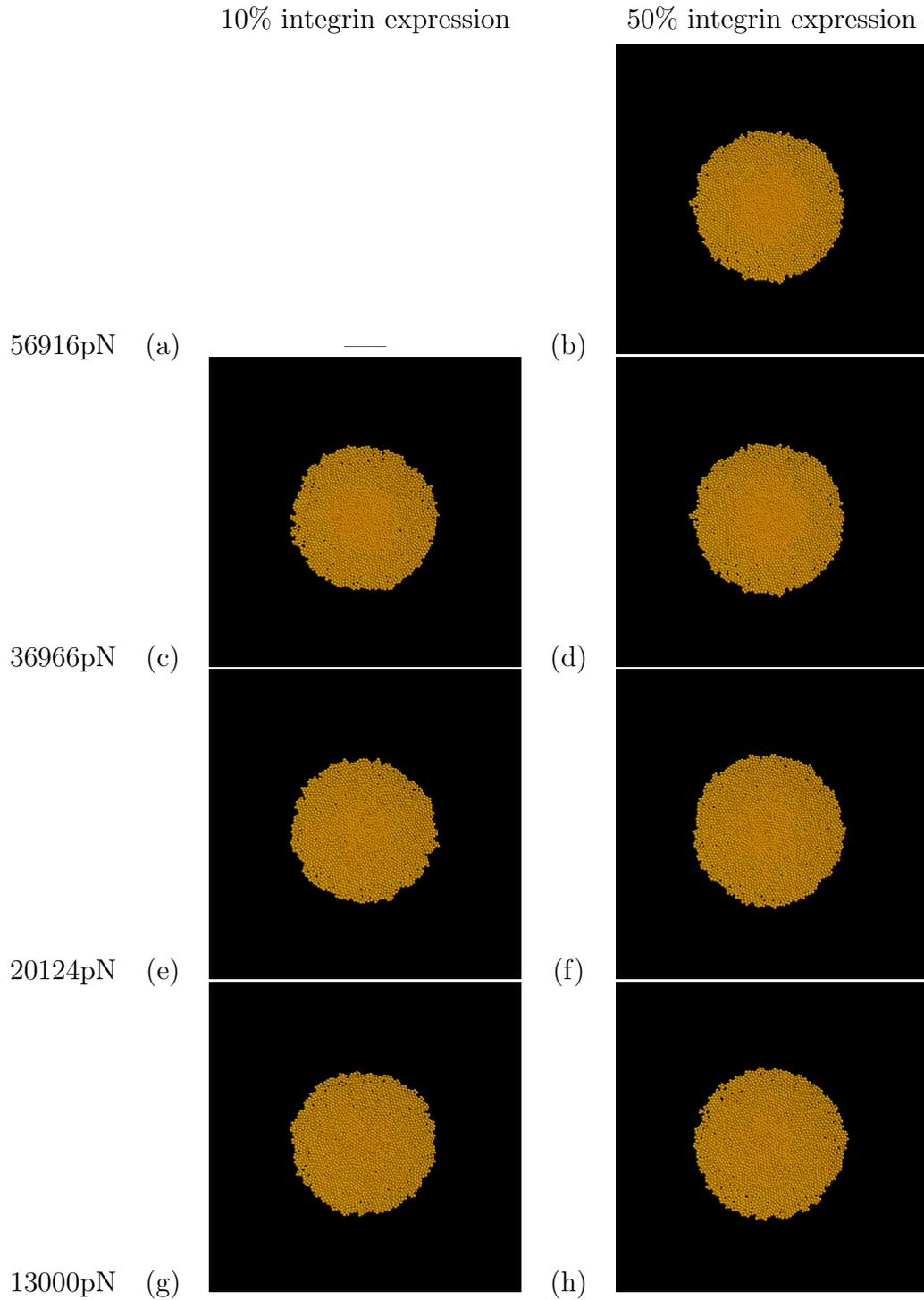
Figure 6.13: *Images of the cell colonies of the first simulation in each set with a 75% E-cadherin expression level after 7 days. The matrix fibres are not shown here as they are too small in size to be properly detectable in the images.*

Colony Growth Results

Here we consider the growth of the cell colony over 7 days. For each of the 15 simulations in each set, an exponential growth rate, a switching point between exponential and linear growth regimes and a linear growth rate were estimated as described in Section 6.7.1. The average was calculated for each set and plotted as a bar chart. Error bars were plotted to show the standard deviation within the sets. Figure 6.14 shows the results. Figure 6.14(a) shows the average value and standard deviation of the exponential growth rates per day. It can be seen that these growth rates are very much the same for the different pressure thresholds for the cell division. There is a slight difference in the growth rates between the

(a)

(b)

(c)
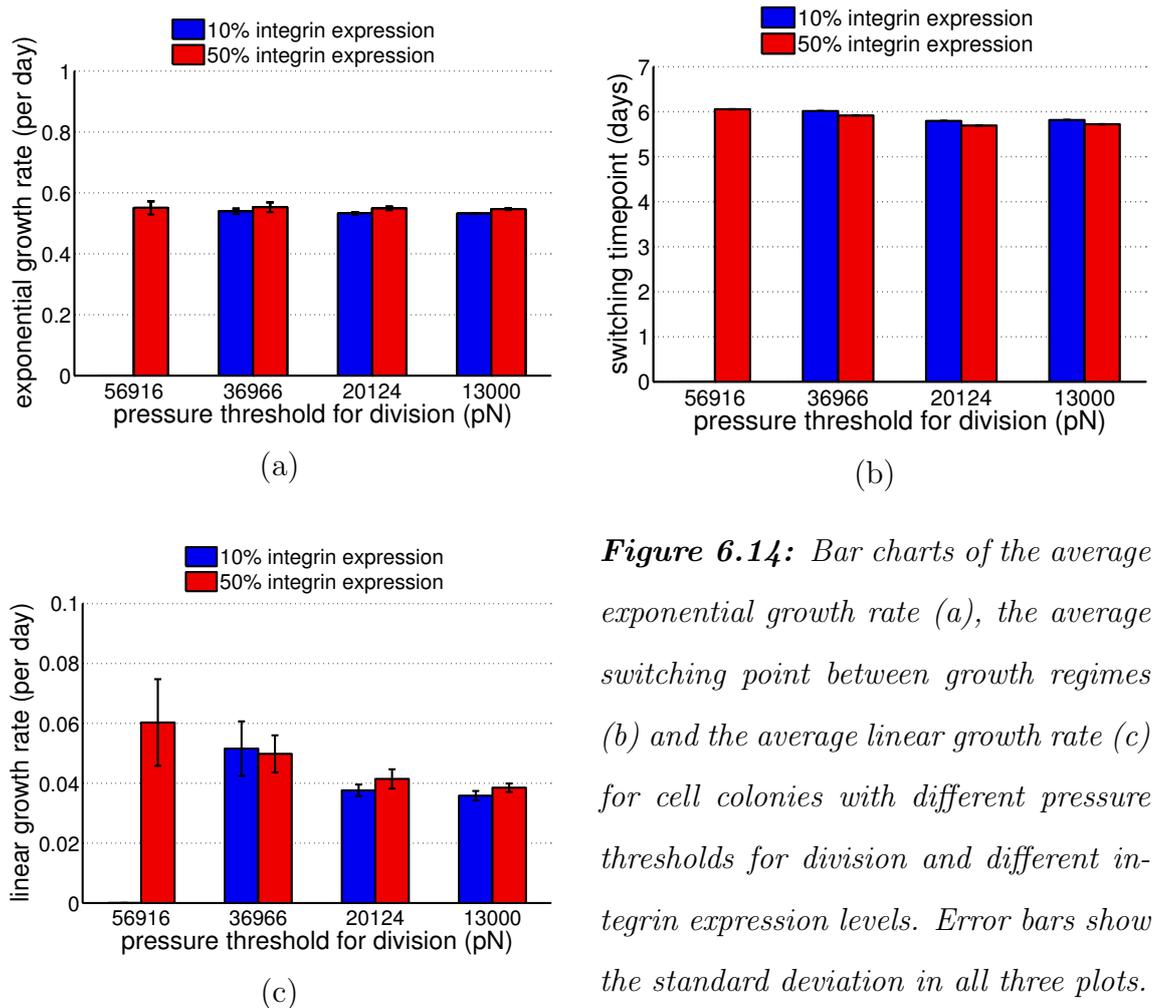
Figure 6.14: *Bar charts of the average exponential growth rate (a), the average switching point between growth regimes (b) and the average linear growth rate (c) for cell colonies with different pressure thresholds for division and different integrin expression levels. Error bars show the standard deviation in all three plots.*

simulations with 10% integrin expression level and those with 50% integrin expression level for all the different pressure thresholds. A 10% integrin expression level leaves the cells less mobile than a 50% expression level and thus the cells cluster closer together which leads to higher pressure and therefore less proliferation. However, it can be seen (Figure 6.14(b)) that the switching between growth regimes happens slightly earlier in those populations with a 50% integrin expression level than in those with a 10% integrin expression level. Furthermore, the switching between growth regimes shifts to earlier points in time with decreasing pressure threshold. Similarly there seems to be an overall decrease in the linear growth rates with decrease in the pressure threshold (see Figure 6.14(c)). However, these results are statistically not significant as the error bars overlap here. It is also not entirely clear in this case how the integrin expression level affects the results as in the case of a high pressure threshold of 36966pN, the average linear growth rate in cells with a 10% integrin expression level is higher than it is in cells with a 50% integrin expression level. In the simulations with lower pressure thresholds of 20124pN and 13000pN the reverse is the case. However, in all of these results the error bars are overlapping thus showing that from these simulations no statistically significant difference can be found between the linear growth rates of populations with 10% and 50% integrin expression level, independent of the pressure threshold.

Overall we can say, that these results indicate that in cell colonies with a 75% E-cadherin expression level, a high integrin expression level of 50% may lead to faster growth of the cell colony than a lower integrin expression level of 10%. Unsurprisingly, a higher threshold on the pressure under which cells still divide, leads, on average, to faster growth that a lower one. However, the noise in the system regarding cell migration affects the linear growth rate more in these cases of higher pressure thresholds that in the cases of lower ones and thus a larger

variability in the growth of these colonies can be observed.

In order to ensure that these results are truly representative of the simulations, we generated log-log plots of the simulation data against the data resulting from the growth function that combines the average exponential growth rate, the average time point for switching between an exponential growth regime and a linear growth regime, the average linear growth rate and the matching y-intercept. The results are shown for all seven sets of simulations in Figure 6.15. As explained in Section 6.7.1, 5 time points were randomly selected from a uniform distribution and the number of cells that existed in the simulation at these time points were plotted against the number of cells that the constructed growth function would predict for these time points. This was done for each simulation in each set. In the plots in Figure 6.15 these points are shown as squares. It can be seen that in all cases all the squares are very close to the diagonal which means that the constructed function gives a good fit to all the simulations in a set. Therefore the averages and standard deviations that we considered in Figure 6.14 represent the data very well.

Colony Dispersal and Spread Results

We also investigated the dispersal and spread of the cell colonies using the procedure explained in Section 6.7.2. We first calculated the number of cell clusters that the cell colony formed after the simulation of 7 days of real time. The results can be seen in Figure 6.16(a). The figure shows that in all simulations with all the different parameter combinations, the colony formed one contiguous cluster. This shows that the 'normal' development of epithelial cell colonies as cohesive cell clusters is robust to some fluctuations in the E-cadherin expression level. An

161

10% integrin expression   50% integrin expression

56916pN   (a)   (b)
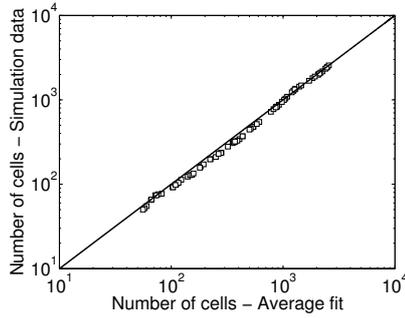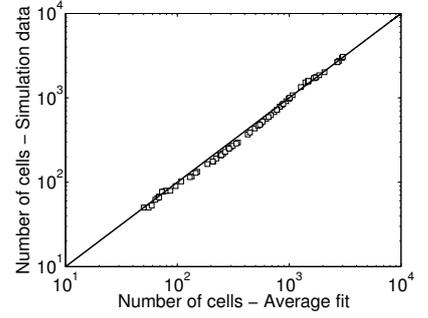
36966pN   (c)   (d)
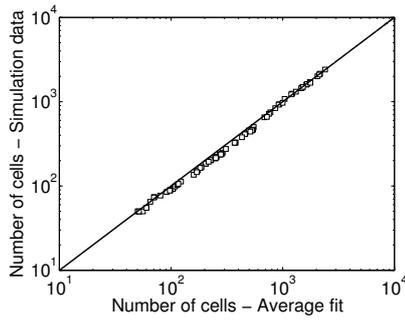
20124pN   (e)   (f)

13000pN   (g)   (h)

Figure 6.15: *Log-log plots showing the goodness of fit of the simulated data to the growth function constructed from the average exponential growth rate, the average switching point, the average linear growth rate and the matching y-intercept. For each simulation in each set 5 uniformly distributed time points were randomly chosen and the number of cells at each of these time points in the simulation were plotted against the number of cells predicted by the constructed growth function.*

162

Figure 6.16: *Figure showing the number of clusters (a) and the frequency of clusters of a certain size (b) for cell colonies with the different thresholds on the pressure for division and a 10% and 50% integrin expression level.*

expression level of 75% still leads to this 'normal' behaviour independent of integrin up regulation and increased proliferation. As there was only one cell cluster in all cases there was no need to calculate any distances between clusters. We did however calculate the distribution of the cluster sizes in the different sets of simulations. Figure 6.16(b) shows the results. The plot shows that in all the simulations of each set the cells formed a cluster of between 10^3 and 10^4 cells.

6.8.2 The development of cell colonies with a homogeneous E-cadherin expression level of 50%

Secondly we considered cell populations in which all cells express E-cadherin at a 50% level. Again we studied the effect of increased ability of the cells to proliferate under pressure and combined this with different levels of integrin expression. Just as before, we ran 15 simulations for each of the different parameter combinations. We observed the development of the colony over time and the configuration

of the colony after 7 days. Figure 6.17 shows images of the cell colonies for the first simulation of each set after 7 days. The main mass of the growing cell colony looks vey similar in all the simulations. However, close examination shows that there is a difference between the simulations of colonies which express integrins at a 10% levels and those which express integrins at a 50% level. In the simulations with low integrin expression the cell colonies grow contiguously whereas the colonies in the simulations with high integrin expression, seem to have some individual cells or small cell clusters invading the environment. We can also see in some of the images, especially image 6.17(e), that the circumference of the cell mass is quite rough with small extensions into the environment.

In order to shed more light on the differences between the different simulations, we again study first the growth of the colony and then the spread and dispersal for the different sets of simulations.

Colony Growth Results

Here we again considered the growth of the cell colony over 7 days. We estimated an exponential growth rate, a switching point between exponential and linear growth regimes and a linear growth rate for each of the 15 simulations in each set as described in Section 6.7.1. The average was calculated for each set and plotted as a bar chart. Error bars were plotted to show the standard deviation within the sets. Figure 6.18 shows the results. Figure 6.18(a) shows the average value and standard deviation of the exponential growth rates per day. Similar to previous results, the figure shows that in the cell colonies considered here the exponential growth rate is again independent of the pressure threshold for proliferation. There is, however, again a dependence of the growth rate on the integrin expression level. Cell colonies which express integrins at a 10% level, have a decreased exponential growth rate compared to cell colonies with a 50%

10% integrin expression   **50% integrin expression**

56916pN   (a)   (b)

36966pN   (c)   (d)

20124pN   (e)   (f)

13000pN   (g)   (h)

Figure 6.17: *Images of the cell colonies of the first simulation in each set with a 50% E-cadherin expression level after 7 days. The matrix fibres are not shown here as they are too small in size to be properly detectable in the images.*
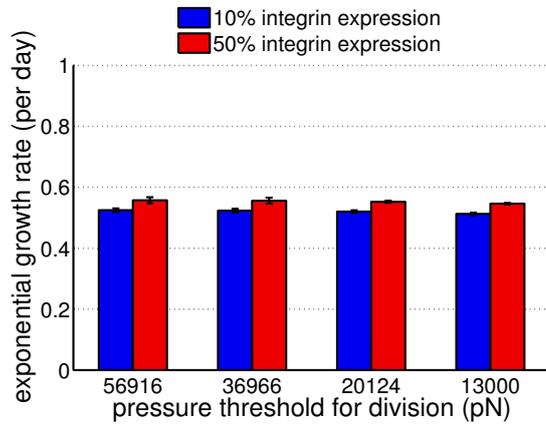
(a)
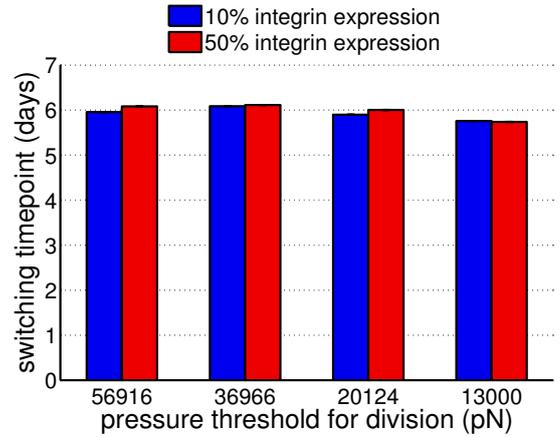
(b)
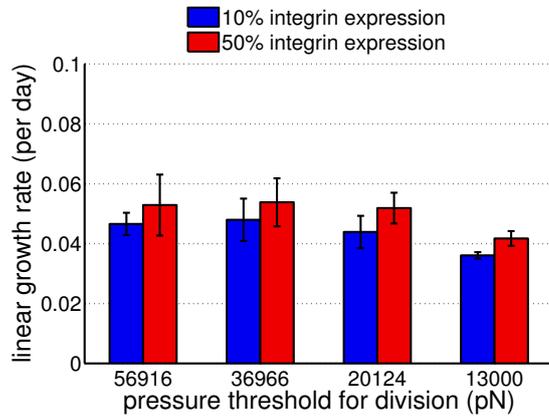
(c)

Figure 6.18: *Bar charts of the average exponential growth rate (a), the average switching point between growth regimes (b) and the average linear growth rate (c) for cell colonies with different pressure thresholds for division and different integrin expression levels. Error bars show the standard deviation in all three plots.*

integrin expression level.

The point in time at which the switching between growth regimes takes place is shown in Figure 6.18(b). These time points are very similar for high pressure thresholds but seem to be shifted slightly to earlier points in time for pressure thresholds of 20124pN and 13000pN. In addition, the switching between growth regimes takes place slightly earlier in cell colonies with low integrin expression compared to those with high integrin expression except in colonies with a pressure threshold of 13000pN where the switching time point is independent of integrin expression. The linear growth rates show a similar pattern in Figure 6.18(c). Here the average of the growth rates is also roughly the same in colonies with

166

high pressure thresholds of 56916pN and 36966pN whereas it decreases with further decreasing pressure threshold. Furthermore, the averages of these growth rates are also lower for colonies with low integrin expression levels compared to those with high integrin expression levels. However, the standard deviations are relatively large for the linear growth rates and thus the pattern found is statistically not significant. Figure 6.19 does show, though, that the growth function which combines the average exponential growth rate, the average time point for switching growth regimes, the average linear growth rate and the matching y-intercept fits the simulated data well.

Overall this indicates again that a high integrin expression level may lead to faster growth of the cell colony than low integrin expression levels.

Colony Dispersal and Spread Results

As in the previous section, we again investigated the spread and dispersal of the cell colonies according to the algorithm explained in Section 6.7.2. First we calculated the number of cell clusters that the cell colony formed after 7 days. The results are shown in Figure 6.20(a). The figure shows a definite difference between the number of cell clusters that arise in simulations with different integrin expression levels. Independent of the pressure threshold for proliferation, the cell colonies, which express low levels of integrin, grow as contiguous clusters. Cell colonies, which express high levels of integrin, on the other hand, can form more than one cluster after 7 days. The average of the number of clusters formed by these colonies decreases with decreasing pressure threshold. This is however a difference that is statistically not significant. The results highlight the impact of the stochasticity on the simulations, since it is clear that in some of the 15 simulations in each set, only one cell cluster arises whereas in others the cells form more than one cluster after 7 days. Figure 6.20(b) shows that whenever

167

10% integrin expression   50% integrin expression
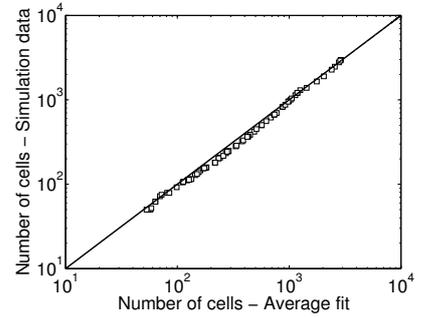
56916pN   (a)   (b)
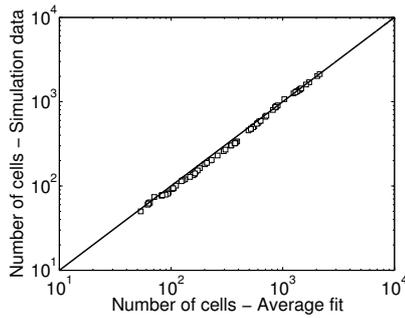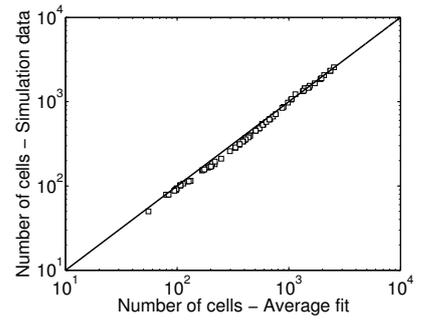
36966pN   (c)   (d)
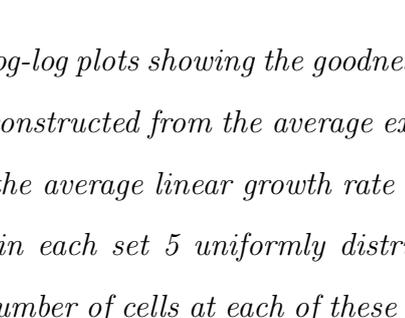
20124pN   (e)   (f)
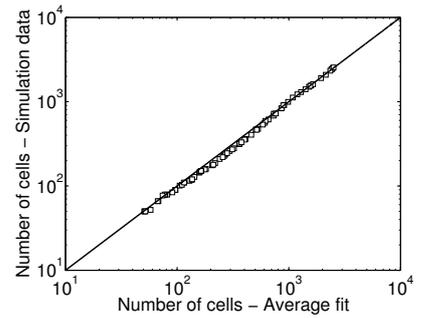
13000pN   (g)   (h)

Figure 6.19: *Log-log plots showing the goodness of fit of the simulated data to the growth function constructed from the average exponential growth rate, the average switching point, the average linear growth rate and the matching y-intercept. For each simulation in each set 5 uniformly distributed time points were randomly chosen and the number of cells at each of these time points in the simulation were plotted against the number of cells predicted by the constructed growth function.*
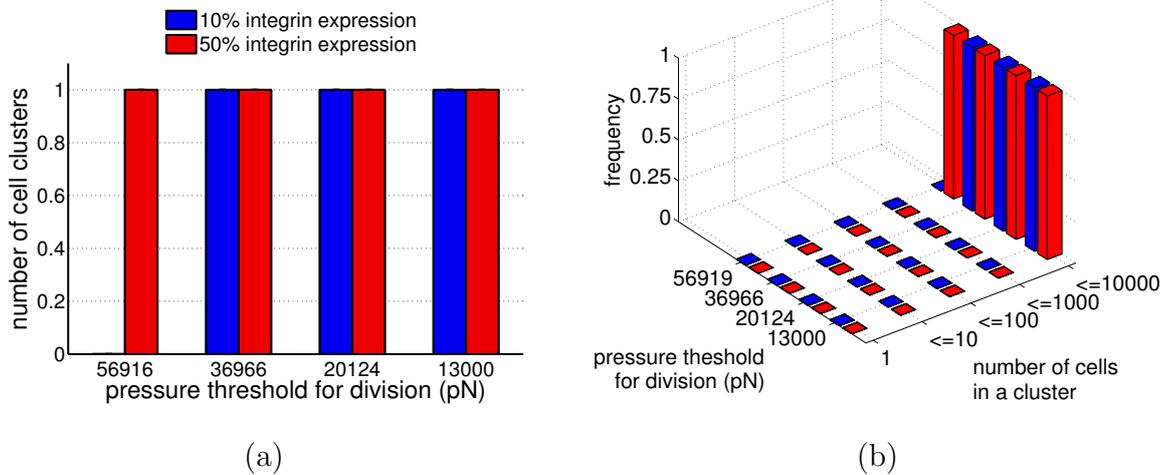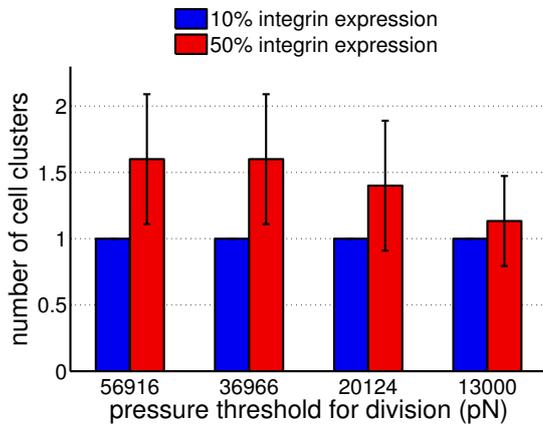
(a)
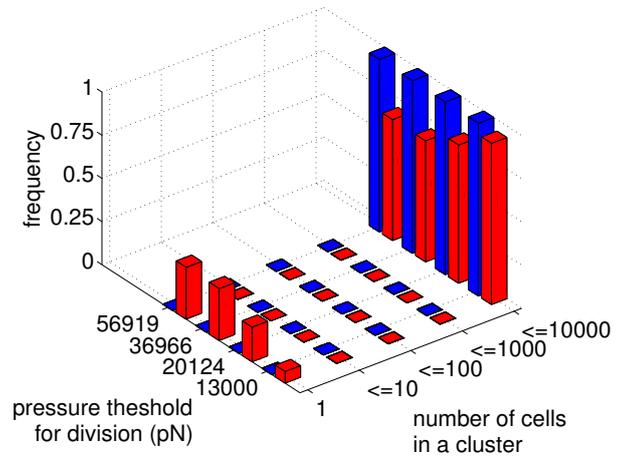
(b)
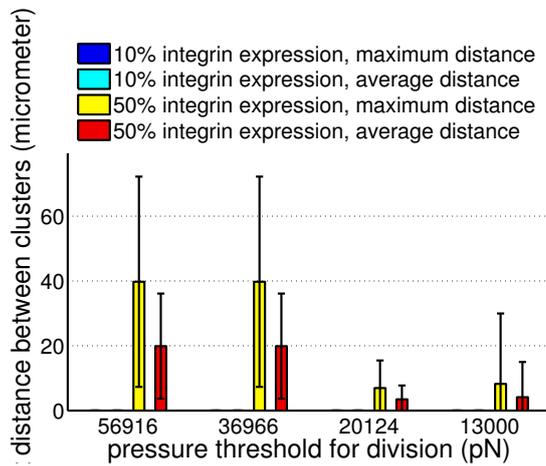
(c)

Figure 6.20: *Plots showing the number of clusters (a), the frequency of clusters of a certain size (b) and the average as well as maximum distances between clusters (c) for cell colonies with the different thresholds on the pressure for division and a 10% and 50% integrin expression level.*

more than one cell cluster arises, these additional clusters are single migrating cells. Their maximum and average distance from the main cell mass is shown in Figure 6.20(c). We found very large standard deviations for these values, most likely due to the small number of cases in which additional clusters arose and the fact that they only arose in very small numbers. When considering the average values, the individual cells seem to have been less motile over the course of the 7 days in cell colonies with lower pressure thresholds. However, due to the large and overlapping standard deviations, we cannot interpret them as significant results.

6.8.3   The development of cell colonies with heterogeneous E-cadherin expression levels

Lastly we considered cell populations with heterogeneous E-cadherin expression profiles. We assumed the level of heterogeneity to be as high as possible and thus each cell in the population could express between 100% and 0% of E-cadherin. The cell-specific levels were randomly generated when cells were initialised. The levels of the 50 cells that were initially placed in the domain, were generated at the start of each simulation and stayed fixed throughout the simulation. For each cell that was generated during a simulation, the E-cadherin level was computed immediately at the start of the cell division process and was independent of the expression level of the 'mother cell'.

Again we studied the effect of increased ability of the cells to proliferate under pressure and combined this with different levels of integrin expression. Just as before, we ran 15 simulations for each of the different parameter combinations. The development of the colony was observed over a period of 7 days and the configuration of the colony at the end of the 7th day. Figure 6.21 shows images of the cell colonies for the first simulation of each set after 7 days. The intensity of the yellow colour again shows the amount of E-cadherin in the cytosol. In Section 6.3.3 the cells changed their colour from yellow to red as the E-cadherin was sequestered in adhesive bonds. However, here it seems that the cells, which express a high amount of E-cadherin, show a higher intensity of the yellow colour than those which express low levels of E-cadherin. This is due to the fact that, because of the heterogeneous E-cadherin expression, less E-cadherin can be taken up in bonds in a cell that has a high expression level if some of the neighbouring cells have low expression levels, than if all cells had a homogenous high expression level. Thus there is more E-cadherin left in the cytosol in cells with a high expression level. The images in Figure 6.21 therefore show the E-cadherin expression

10% integrin expression   **50% integrin expression**

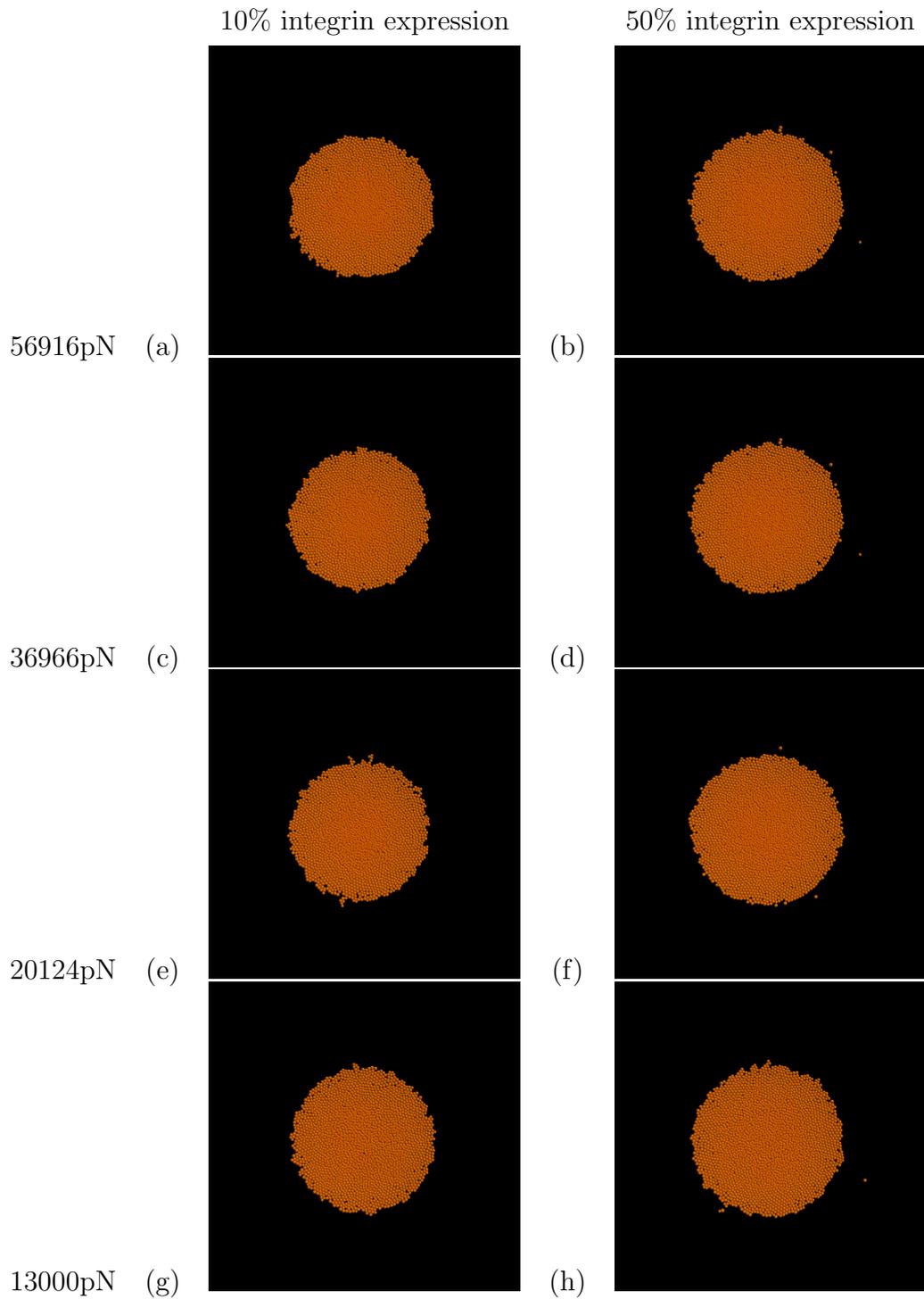56916pN (a) (b)

36966pN (c) (d)

20124pN (e) (f)

13000pN (g) (h)
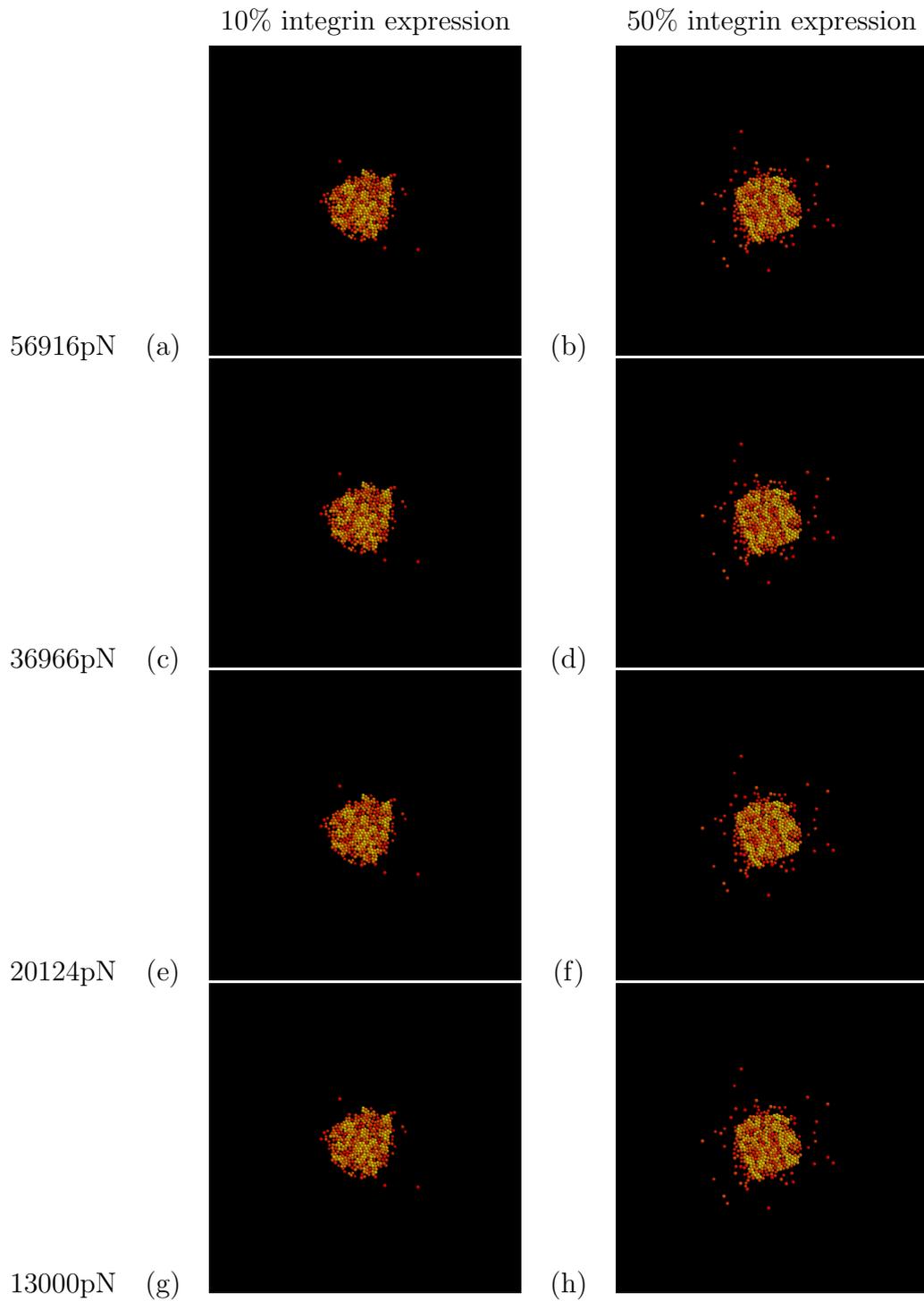
Figure 6.21: *Images of the cell colonies of the first simulation in each set after 7 days. The matrix fibres are not shown here as they are too small in size to be properly detectable in the images. The colour of the cells is proportional to the amount of E-cadherin in the cytosol.*

profile and its effects on the E-cadherin dynamics well. They, furthermore, show the dispersal of individual cells from the main cell cluster. This behaviour can be observed to a much higher extend in colonies with a 50% integrin expression level compared to those with a 10% expression level.

Colony Growth Results

In order to analyse the growth characteristics of the colonies over 7 days, we again calculated the average and standard deviation of the exponential growth rate, the switching time between growth regimes and the linear growth rate from all 15 simulations in a set. The results are shown in Figure 6.22. Figure 6.22(a)

(a)

(b)

(c)
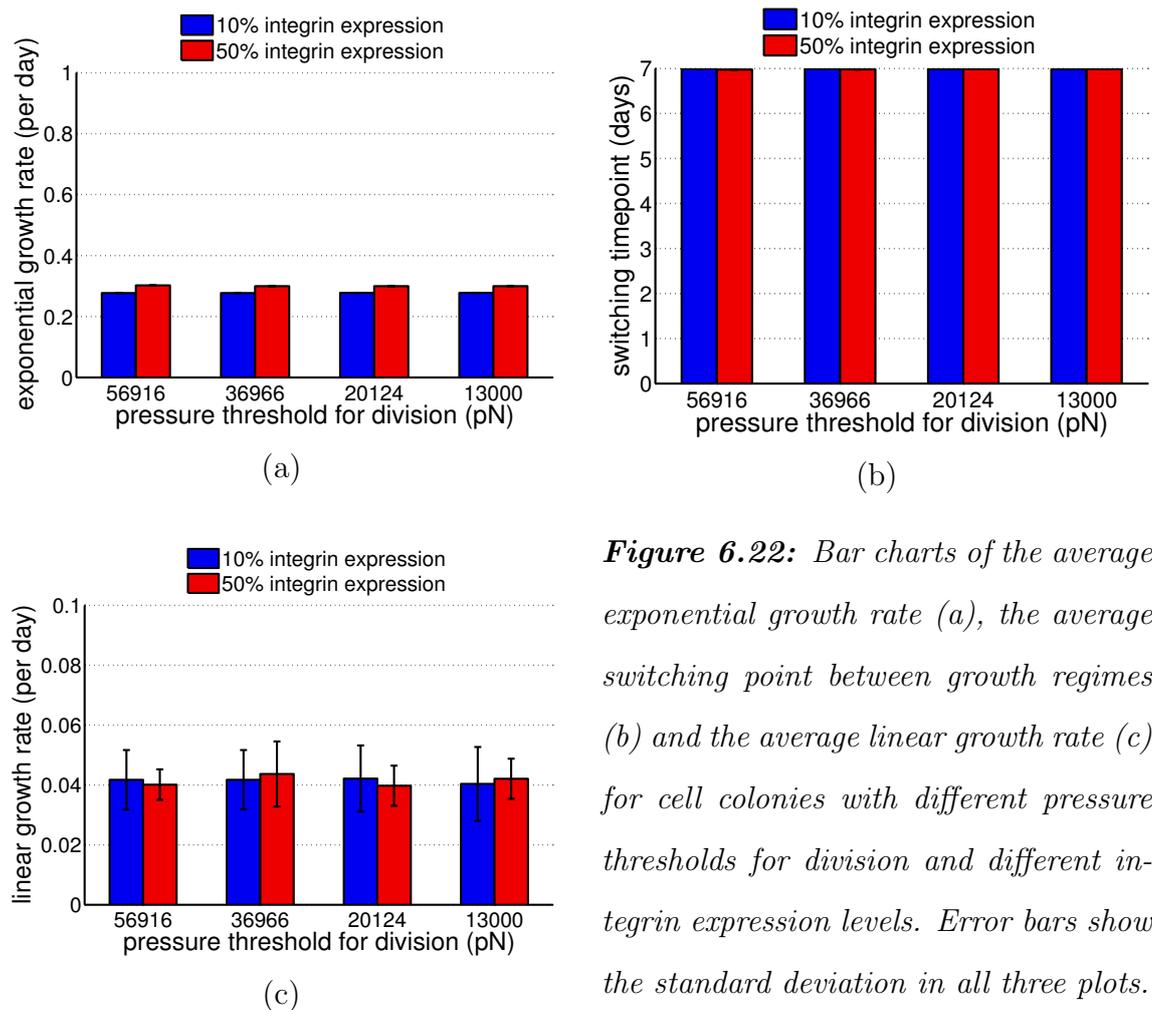
Figure 6.22: *Bar charts of the average exponential growth rate (a), the average switching point between growth regimes (b) and the average linear growth rate (c) for cell colonies with different pressure thresholds for division and different integrin expression levels. Error bars show the standard deviation in all three plots.*

shows the average value and standard deviation of the exponential growth rates per day. It can be seen that the growth rates are quite low, at roughly 0.3/day. Just as in previous sections there is a slight difference between the exponential growth rates in colonies that express low levels of integrin and those that express high levels. Lower integrin expression levels lead to lower exponential growth rates. However, the pressure threshold, under which cell division can occur, does not seem to influence these growth rates. Figure 6.22(b) shows the average value and standard deviation of the point in time when the colonies switch from an exponential to a linear growth regime. In all the simulations with the different parameter combinations the switching point is very close to the end of the 7th day. It is possible that longer simulations would show that these colonies can grow exponentially for even longer than 7 days. Therefore, and also due to the large standard deviations, the linear growth rates shown in Figure 6.22(c) can not be seen to hold any meaningful information.

We can conclude by saying that cell populations with extremely heterogeneous E-cadherin expression levels grow exponentially over periods of at least 7 days. Therefore they follow an exponential growth regime longer than populations with medium to high homogenous E-cadherin expression levels. The exponential growth rates are however relatively low and are not influenced by pressure thresholds on the proliferative activity. Low integrin expression, however, slows growth down slightly.

As in the previous sections, we produced a log-log plot of the simulation data against the data resulting from the growth function that combines the average exponential growth rate, the average time point for switching growth regimes, the average linear growth rate and the matching y-intercept. Figure 6.23 shows these plots. It can be seen in all the plots that the growth function fits the data well. Therefore the growth characteristics displayed in Figure 6.22 give a good

10% integrin expression   50% integrin expression

56916pN   (a)   (b)

36966pN   (c)   (d)
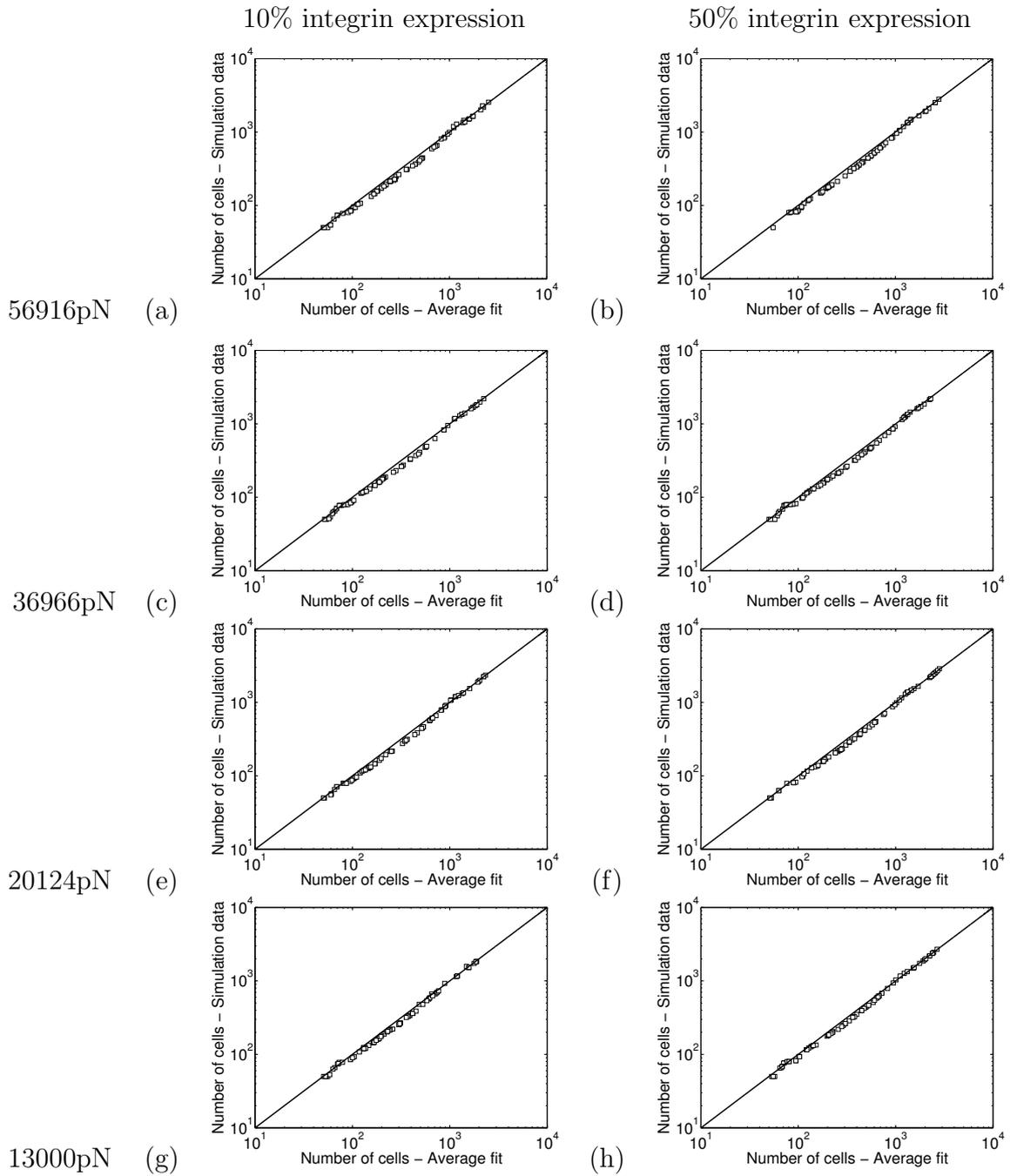
20124pN   (e)   (f)

13000pN   (g)   (h)
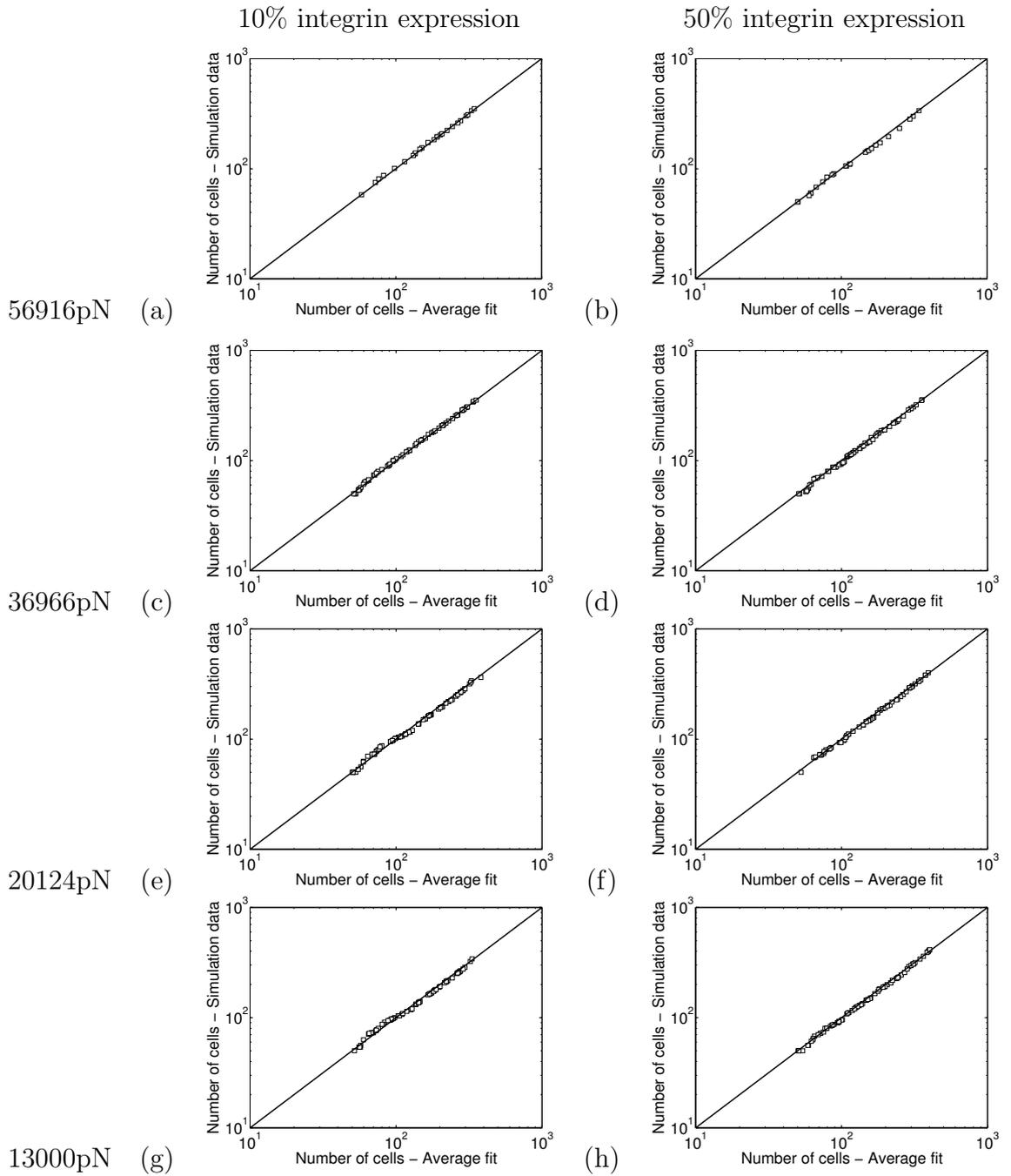
Figure 6.23: *Log-log plots showing the goodness of fit of the simulated data to the growth function constructed from the average exponential growth rate, the average switching point, the average linear growth rate and the matching y-intercept. For each simulation in each set 5 uniformly distributed time points were randomly chosen and the number of cells at each of these time points in the simulation were plotted against the number of cells predicted by the constructed growth function.*

description of the colonies' growth dynamics.

Colony Dispersal an Spread Results

We again investigated the dispersal and spread of the cell colonies using the procedure explained in Section 6.7.2. Doing so, we first calculated the number of cell clusters that the colonies formed after the simulations of 7 days of real time. The results are shown in Figure 6.24(a). There is a striking difference between the

(a)

(b)

(c)
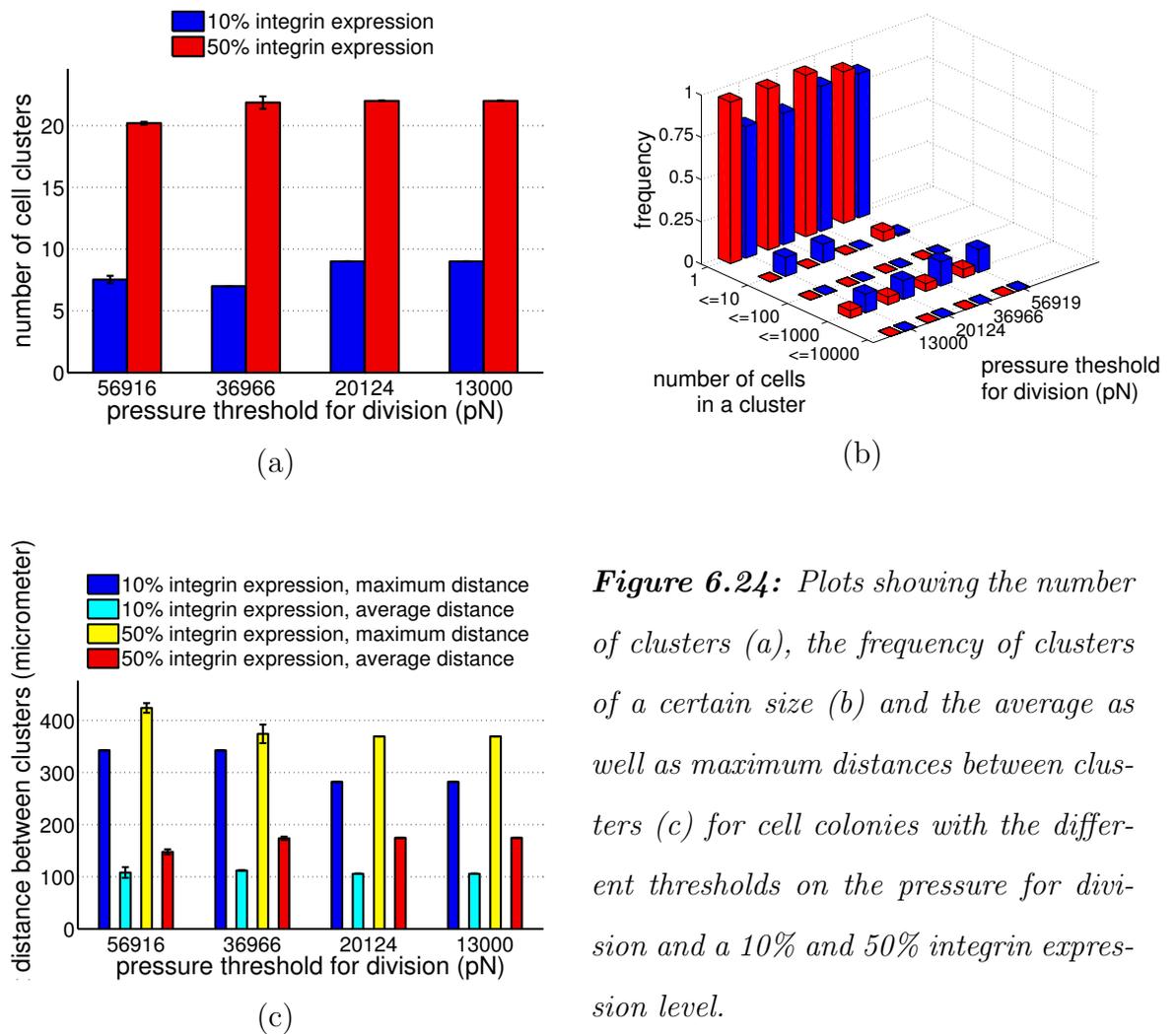
Figure 6.24: *Plots showing the number of clusters (a), the frequency of clusters of a certain size (b) and the average as well as maximum distances between clusters (c) for cell colonies with the different thresholds on the pressure for division and a 10% and 50% integrin expression level.*

number of cell clusters that colonies with a 10% integrin expression level form and the number of cell clusters that colonies with a 50% integrin expression level form.

Unsurprisingly, the number of clusters is higher for colonies with a 50% integrin expression level. The pressure threshold for cell proliferation does not affect these results very much. The only influence of this parameter that can be seen is that colonies with higher pressure thresholds of 56916pN and 36966pN seem to form slightly fewer cluster than those with lower pressure thresholds. This is the case for both, high and low integrin expression levels. What is even more interesting than the number of clusters that the colonies form is the distribution of the cluster size. Figure 6.24(b) shows this. It can be seen that, for low pressure thresholds, the colonies with a 50% integrin expression level form either clusters of a single cell, or of 10^2 to 10^3 cells. Given the images in Figure 6.21, these colonies form one large cluster and have a number of single cells migrating away from the cluster. Colonies with a 10% integrin expression profile, on the other hand, seem to form small clusters with 2 to 10 cells as well as clusters of individual cells and those of 10^2 to10^3 cells. For a pressure threshold of 36966pN both, the colonies with high and those with low integrin expression levels, disperse into one cluster of 10^2 to 10^3 cells and individually migrating cells. The number of individually migrating cells is however higher for colonies with higher integrin expression levels than for those with low integrin expression levels. In the case of a pressure threshold of 56919pN both, high and low integrin expression levels, lead to colonies consisting of one large cell cluster, individually migrating cells and small clusters of 2-10 cells. Figure 6.24(c) shows the average and maximum distance between these different clusters in the different colonies. It is unsurprising to see that both, the maximum and the average distance between cell clusters of a colony are greater in colonies that express high levels of integrin compared to those that express low levels of integrin. The maximum distance achieved between clusters of colonies with a 10% integrin expression levels is, however, still considerably greater than the average distance between clusters in population with a 50% integrin expression

levels.

Overall, we can say that colony dispersal and spread is generally independent of any pressure threshold for proliferation in cell populations with heterogeneous E-cadherin expression profiles. Integrin expression however plays a major role in this. Low integrin expression leads to fewer clusters and thus less dispersal as well as less spread of these cluster but it does allow the formation of large and small cell clusters as well as individually migrating cells at low pressure thresholds for division. Higher integrin expression leads to more dispersal and spread but only allows the growth of the original large cluster and individually migrating cells in most cases.

6.8.4 A comparison of the cell colonies with different E-cadherin expressions profiles

In the previous sections we studied the impact of different integrin expression levels and different pressure thresholds on proliferation, on the development of cell colonies with a homogenous 75% E-cadherin expression level, a homogenous 50% integrin expression level and a heterogeneous E-cadherin expression level. The images of the cell colonies after 7 days (see Figures 6.13, 6.17 and 6.21) already show that there are big differences between the development of the colonies with the different E-cadherin expression profiles. All the colonies with a 75% E-cadherin expression level form contiguous cell clusters. In the case where we examined colonies with a 50% expression level, only those colonies with a low integrin expression level consistently formed one contiguous cluster. Colonies with high integrin expression level showed the appearance of some individually migrating cells in addition to the main cell cluster. When comparing the colour of the cells between the simulations of colonies with a 75% expression level and those with a 50% expression level, it is obvious that the cells with lower E-cadherin

expression level are darker and thus have less E-cadherin in the cytosol. This is most likely due to the lower E-cadherin expression level but a similar uptake of E-cadherin in the cell-cell bonds.

Unsurprisingly, the cell colonies had a very different appearance after 7 days in the case when heterogeneous E-cadherin levels were considered compared to when homogenous E-cadherin levels were considered. In both cases, high and low integrin expression levels, many individually migrating cells arose in addition to a main cell cluster in these simulations with heterogeneous E-cadherin expression profiles.

Although the physical structure of the colonies varies highly with E-cadherin expression levels, the overall growth patterns are similar. The exponential growth rates of the cell colonies are smaller when the cells express integrins at a low level in comparison to when they express integrins at a high level. The pressure threshold had no significant influence on this behaviour in any of the sets of simulations. In addition, in both of the cases where we considered homogenous cell populations regarding the E-cadherin expression, the point of time, at which the the switch between growth regimes occured, seems to be shifted to earlier time points for decreasing pressure threshold. This is unsurprising as a lower pressure threshold decreases the possibility of cell division in the centre of the colony, leading to a switch of the growth regime from exponential to linear. This switching time point does, however, seem to be influenced differently by the integrin expression level in the two cases. In the populations with a 75% integrin expression level switching between growth regimes happens slightly later in cells expressing low levels of integrin whereas in populations with a 50% integrin expression levels switching occurs slightly earlier in cells expressing low levels of integrin compared to those, which express high levels of integrin. The differences are very small so it is possible that they should not be taken too closely into consideration.

However given that the results are consistent throughout the simulations, there might be a fundamental difference between the influence of integrin expression levels on the switching between growth regimes in cell populations with different levels of E-cadherin expression. The linear growth rates show similar patterns again for the cell populations with 75% and 50% E-cadherin expression levels. The average is in both cases lower for low integrin expression for all the different pressure thresholds. While a clear decrease in the average growth rates with the pressure threshold is seen in the cases of 75% E-cadherin expression levels, only slight differences can be noticed in cell populations with a 50% expression level. The standard deviations are high in all the cases so that most of the results are statistically not significant. This is also the case for the linear growth rates of the heterogeneous cell populations.

The results concerning the colony spread and dispersal show the greatest differences between the colonies with the different E-cadherin expression profiles. They confirm the observation from the images of the colonies that a 75% E-cadherin expression always leads to the development of one contiguous cell cluster. They also confirm that, dependent on the integrin expression level and the noise in the system, a 50% E-cadherin expression level can lead to either single contiguous clusters or the additional development of individually migrating cells. In the case of heterogeneous E-cadherin expression profiles, the cell populations develop more than one cell cluster in all cases. Low integrin expression levels lead to fewer clusters than high integrin expression levels, but in addition to one large cluster and single migrating cells, they allow the development of small cell clusters in the case of low pressure thresholds for division and also at a very high pressure threshold. In colonies with high integrin expression levels this could only be observed in colonies with high pressure thresholds for division. The spread of the entire colony is much larger in the case of a heterogeneous cell populations than

in the case of the cell populations with a homogenous E-cadherin profile of 50%. Overall the results show that the phenotype of the cell populations does not change for some level of perturbation in the E-cadherin expression level. However, further down-regulation of E-cadherin allows individual cells to leave the main mass of cells and invade the environment. This process is also highly dependent on the integrin expression level. Heterogeneous E-cadherin expression levels can, furthermore, lead to the invasion of the environment by not just single cells but also by small groups of cells. This is however again dependent on the integrin expression level. In addition, the integrin expression level also seems to have an influence on the growth of a cell colony where high integrin expression levels lead to faster colony growth.

6.9 Discussion

In this chapter we have formulated a multiscale model of cell colonies and used it to study the transition from healthy cell behaviour to invasive, cancer-like behaviour by introducing parameter changes that are in accordance with changes happening during the epithelial to mesenchymal transition.

We developed the model on the basis of the cell migration model from previous chapters by including cell division as well as intra- and intercellular cell-cell interaction dynamics. To this end, we first studied the intracellular E-cadherin–β-catenin dynamics and developed an ordinary differential equation model to describe the E-cadherin–β-catenin complex formation and disruption as well as the transition of these complexes between the cell's interior and the different cell-cell contact sites. The model parameters were estimated by fitting the number of E-cadherin–β-catenin bonds between two cells to cell-cell separation force data. Two possible models were considered. In the first one we assumed that the number of E-cadherin bonds that can form at one cell-cell contact site is limited and

thus only a certain percentage of all possible E-cadherin–β-catenin complexes in a cell will be translocated to this site. In the second model we assumed that there is no limit on the number of bonds that can form between two cells, however, in order to ensure that more cells can bind to a pair of cells, the E-cadherin–β-catenin complexes are redistributed upon attachment of further cells to either cell. This leads to the assumption that the forces that are measured between a pair of cells in isolation is higher than the forces between any two cells in an epithelial layer. Interestingly much better fits to the data could be found for this second model. It would be very fascinating to investigate this experimentally and see whether this really is the case. The search for a model that fits the data best also showed that there might be interesting dynamics occurring in cells concerning the internalisation of E-cadherin–β-catenin complexes that cannot bind to a complex on the neighbouring cell. It made a big difference whether or not this process was included in the models tested here. The simple description of the process as a function proportional to a smooth approximation to a Heaviside function which is zero when there are more E-cadherin–β-catenin complexes on the neighbouring cell at the contact site than at the contact site in the cell of interest and a certain proportionality factor otherwise, lead to bad fits to the data for the models. Thus this process is likely to be more complex and might involve time delays in such a way that E-cadherin–β-catenin complexes stay available on the surface of one cell for short periods of time to enable binding to complexes on a neighbouring cell which are exocytosed slightly later. However, in order to keep the model as simple as possible and given how little data there are available to help estimate parameter values, we ignored this process in the final version of the intracellular model and obtained good fits to the data. It is possible that this is due to the redistribution of complexes between cell-cell contact sites which automatically also redistributes the unbound complexes. Fitting this final version of the model

of the intracellular dynamics to the data produced a set of parameters that gave the best fit. Surprisingly, the undirected E-cadherin–β-catenin complex translocation rate ι has the highest value of all the fitted rate constants. Thus it is also higher than the directed translocation rate. This might be the case because it allows a much faster binding-reaction when cells come into contact initially. However, it is possible that this is not the case in live cells and that it is due to the simplifications of the model or the specific data we are fitting to. In any case, the rate constants themselves are not of very great importance here, as it is the overall dynamics at a cell level that we are aiming to capture and reproduce.

In addition to the intracellular dynamics and resulting adhesion, we included a simple cell cycle in each cell, which allows the cells to divide when they are in M-phase, and cell-cell repulsion governed by the Hertz model. Thus we could consider growing cell colonies. We ran simulations of these growing colonies and varied the value of the endocytosis rate of E-cadherin–β-catenin complexes. We also introduced a threshold for the level of pressure, i.e. the magnitude of the repulsive force, under which a cell could still divide, and studied its effects on colony growth. Thirdly, we worked around possible discrepancies between the modelling technique and the biology by assuming that the active force between two cells whose contact site had a diameter larger than one sixth of the cell's circumference was either zero or equal to the repulsive force alone. These three variations of the model were combined in simulations and the colonies were studied after three and after seven days. The results were compared to experimental findings regarding the distribution of the number of neighbours of cells in an epithelial layer. They were also examined as regards the expectation that the colonies should be growing considerably between day three and day seven and that they should be growing in a near-circular fashion. These considerations led to the observation that there indeed has to be a threshold on the pressure under

which cells divide. Also they showed that the active force between two cells has to be just the repulsive force when cells get too close and zero when they are at a distance that we consider to be their 'natural state' in epithelia. This means that the forces exerted by the E-cadherin–β-catenin bonds are only active when the cells are further apart than they would be in naturally occurring epithelial sheets and thus the forces prevent separation rather than leading to closer adhesion. From a biological point of view this is a sensible conclusion as a constant tugging by the adhesion and repulsion forces to keep a balance would be wasting the cell's energy. In future work a new model could be formulated to replace the extended Hertz model which takes these thoughts into account. However, for the rest of this work we enforced both, a threshold for the pressure under which the cell can divide and an active force equal to the repulsive force between cells closer than their natural state, by additional rules. From the simulations of the colony growth we furthermore saw that in order for the simulations to fit the expectations and the data best, we had to choose a parameter value for ρ_d that led to a much lower average adhesion force between any two cells in the colony than one would expect by extrapolation of the adhesion force data between two cells. This highlights again that it would be very interesting to devise experiments that allow the measurement of adhesion forces between cells in a layer rather than just a pair of cells in isolation.

When running the simulations of growing cell colonies and analysing their results, we were mainly interested in finding a good estimate for ρ_d as well as determining final model modifications. However from the simulations we could also get some ideas regarding the development of invasive colonies. In the results of the 'original' set and the 'division constraints' set we could see that strands of invasive cells could form from the outer cells of the growing colony. This happened at medium

to high values for ρ_d and even more so when a threshold was put on the cell division. This was especially the case for a ρ_d value of 5 in the 'division constraints' set. Thus it seems as though there are ways that invasion can be initiated where down regulation of E-cadherin and up regulation of cell-matrix interactions are not necessary. A relatively fast internalisation of E-cadherin–β-catenin complexes after cell-cell bond disruption and a highly deformable cytoskeleton seem to be the main requirements. As mentioned in Chapter 2, cancer cells have multiple ways of invading the tissue and one of them is using the amoeboid migration mode which indicates that the cells become highly deformable. This is generally only seen as being important for squeezing through holes in a three-dimensional matrix, however, the simulations here indicate that the deformability of cells might also have an impact on other stages of the invasive process and together with slight changes in the E-cadherin–β-catenin dynamics, i.e. an increased rate of internalisation of complexes, this might lead to the fingering patterns seen in invasive tumours. Due to the lack of cell-matrix interactions in these simulations, this behaviour here is most likely regulated by cell division, however it would be very interesting to undertake a more detailed study of this phenomenon in the future.

Finally we conducted simulations in which we combined the multiscale model of cell-cell interactions with the model of cell-matrix interactions. In these simulations we varied the E-cadherin expression level and profile, the integrin expression level as well as the pressure threshold for proliferation. In the first set of simulations we considered cell colonies with homogenous E-cadherin expression levels of 75%. We ran 15 simulations for each combination of integrin expression levels of 10% and 50% with pressure thresholds on proliferation of 13000pN, 20124pN, 36966pN and 56916pN. The simulations of cell colonies with a 10% integrin expression level and a pressure threshold of 56916pN failed due to cells coming

too close. This indicates that this combination of parameters would lead to a multilayered structure, with cells moving up out of the layer due to a lack of space, if a three dimensional domain was considered in these simulations. Thus we could simulate the development of hyperplastic tissue simply by creating this insensitivity to space constraints in cells of an epithelial layer. The simulations of cell colonies with the other combinations of parameters were all successful and we analysed the growth of the colony as well as its dispersal and spread. The results indicated that, as expected, a decreasing threshold on proliferation, which allows cells to divide only if the overall repulsive force they generate is smaller than the threshold, may lead to a slower colony growth. In addition they showed that high integrin expression levels seem to speed up this process slightly. This is unsurprising, as a higher integrin expression level allows the cells to migrate out from the centre of the colony thus relieving the pressure and allowing more cells to proliferate. The analysis of the colony dispersal and spread showed that the cell colonies formed one contiguous cluster in all simulations. This highlighted the robustness of cells to fluctuations in their characteristic parameters. Even with the decrease of the E-cadherin expression level to 75% the simulated colonies did not show any invasive behaviour independent of aberrant proliferation activity and integrin up-regulation.

Next we considered cell populations with a homogenous E-cadherin expression level of 50%. The pressure threshold on proliferation showed very little impact on the colony developments. A slight shift to earlier switching points and a slight decrease in the linear growth rates with decreasing thresholds were the only effects that could be seen. The integrin expression level, however, had a much greater impact on the results. Not only did a higher integrin expression level speed up the growth of these populations, but it also influenced their physical structure. Cell colonies with a 10% integrin expression level showed the development of one

contiguous cell cluster whereas a high integrin expression level of 50% allowed the additional formation of some individually migrating cells. This process was not only influenced by the integrin expression level but also by the added noise as the development of single cell clusters was not observable in all of the 15 simulations in each set with a high integrin expression level after 7 days. This shows that the down-regulation of E-cadherin to a 50% level moves the cell populations to the verge of the epithelial to mesenchymal transition where an additional up-regulation of integrin expression levels can lead to single cells splitting off of the main cell mass and invading the surrounding environment. Furthermore, we observed in especially one simulation (see Figure 6.17(e)), that in cell colonies with a low integrin expression level, very rough circumferences develop which seem to have small invading strands of cells sticking out into the environment which is a pattern seen in cancer cells.

Lastly we considered cell populations with a heterogeneous E-cadherin expression profile. The results of these simulations showed that colonies with heterogeneous E-cadhein expression levels grow with a low exponential growth rate but in turn continue growing exponentially for at least 7 days. The low exponential growth rates are most likely due to a high number of motile cells which express low levels of E-cadherin. This could be the case as the model assumes that cells with established front-rear polarity cannot divide. The slow growth over time most likely leads to the longer period of exponential growth. The pressure thresholds on proliferation did not seem have any influence on this behaviour. Interestingly, a lower integin expression level still led to even further decreased growth. The study of the colony dispersal and spread highlighted further differences between cell colonies with different integrin expression levels. Again, there was little difference between the results for cell colonies with different pressure threshold for division, however, cell colonies with a low integrin expression levels dispersed

into fewer clusters than those with high integrin expression levels. Interestingly, the colonies with a lower integrin expression level showed more variability in the sizes of the emerging clusters, generating clusters of single cells, small clusters of two to ten cells and large clusters (most likely just one large cluster) in most simulations, compared to those with high expression levels. Only a large cluster and single migrating cells emerged in most cases from the simulations of colonies with a high integrin expression level. Only for very high pressure thresholds on division could additional small clusters be observed. Unsurprisingly, the average and maximum distance between the clusters was higher in those colonies with a higher integrin expression level than in those with a low intergin expression level. The simulations of the multiscale model revealed a large variety of behaviours from the development of large, contiguous cell clusters to the invasion of the environment by small fingering patterns, small cell clusters and single cells. It is clear from these results that, unsurprisingly, especially the combination of E-cadherin and integrin expression levels control the integrity or invasiveness of the cell colonies. In future we shall explore the parameter space more thoroughly to get an even better idea of the combinations of the different parameters and characteristics that lead to the different invasive patterns. In order to do so it would also be very useful to include an explicit integrin pathway in the model and take into account the integrin–E-cadherin crosstalk. Furthermore it would be very interesting to study the additional effects of the different matrix characteristics that have shown to influence cell migration in Chapters 4 and 5.

Chapter 7

Modelling the Integrin Pathway

7.1   Introduction

In the previous chapter we examined the E-cadherin–β-catenin dynamics and their impact on cell migration and invasion. However, equally as important in regulating cell migration is integrin expression and activation as mentioned in Chapter 2 and as we also saw in Chapter 6. Not only are integrins up-regulated during the EMT and control the speed cells move at [Palecek et al., 1997], but they are most likely also highly important during the intravasation and extravasation processes which lead to metastases. In all of these processes cancer cells behave very similarly to leukocytes (white blood cells). Leukocytes are transported through the body in the blood vessels. At sites of inflammation the surface of the endothelial cells lining the blood vessel is altered and this can be recognised by proteins on the surface of the leukocytes. A signalling cascade is started inside the leukocytes which leads to the activation of integrins through inside-out signalling. Thus the integrins start adhering to the vessel wall, slowing down the leukocyte movement from flowing with the blood to rolling along the endothelial layer. Through engagement with their ligands on the endothelial cells and through the external forces originating from the blood flow, the integrins are

activated even further. This leads to the arrest of the leukocytes on the vessel wall where they engage with the endothelial cells, squeeze through the cell layer and break down the surrounding membrane and are then able to migrate through the tissue to the site of inflammation [Alon and Ley, 2008, Sahai and Madsen, 2010].

The group of Dr Susanna Fagerholm at the Division of Cancer Research, University of Dundee is focussing its research on the adhesion of B-cells, a specific type of leukocyte. They have carried out multiple studies on B-cell adhesion under static and under flow conditions as well as Western Blot assays of possibly important proteins in the integrin pathway under different conditions. Figure 7.1 shows the signalling pathway of how the transmembrane B-cell receptor (BCR), located on the outer surface of B-cells, triggers integrin activation. This figure was given to us by Hwee San Lek (a PhD student in Dr. Fagerholm's group) and Dr Fagerholm to build a model for the integrin pathway which fits their data and can be used to study certain scenarios to help to understand their experimental results. Furthermore such a model could also be used as a starting point for an integrin pathway which can be included in the multi-cell invasion model of the previous chapter.

7.2 The experiments and the available data

The first part of the BCR pathway up to Akt and PKC activation is well-established and can be found in this form in most representations of the pathway (see for example `cellsignal.com`). Therefore the interest of Dr Fagerholm's group is focussed on Akt and PKC and the pathway downstream from there. Thus Hwee San Lek has carried out a series of Western Blot assays linked with adhesion assays. A Western Blot assay is an analytical technique to detect specific proteins in a cell extract. Here it was used to be able to compare the amount
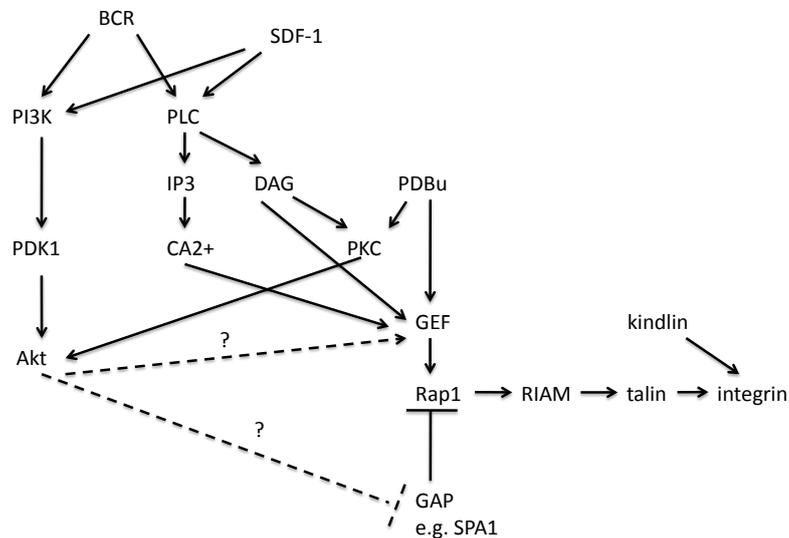
Figure 7.1: *Schematic diagram of the currently known integrin pathway in B cells. The B cell receptor (BCR) is the naturally occurring receptor the stimulation of which initiates the signalling. In experiments it can be substituted by SDF-1 or, if only the lower part of the pathway is studied, by Phobolester (PDBu).*

of activated Akt and PKC under different conditions. In Western Blotting the cells are first lysed and homogenised. Then they are generally covered in a negatively charged buffer. The samples are then placed into wells at the top of a plate covered in gel. When a voltage is applied to the gel, the proteins move towards the positive electrode at the bottom of the gel. The smaller the proteins are, the faster they can move. Thus, this gives a separation of the proteins by size (measured in kDa). Several steps make the detection of the proteins of interest in the different bands, defined by protein size, possible (see images 7.3 and 7.4 as examples of Western Blotting results).

In the experiments by Hwee San Lek, B-cells were exposed to different stimulants and inhibitors for two minutes before the Western Blotting. Three inhibitors

were used: AktVIII, Gö6976 and Ro-31-8220. AktVIII is an Akt inhibitor which selectively, allosterically and reversibly inhibits Akt1 and Akt2. In the present pathway it is Akt1 that is of interest and AktVIII has been shown to potently inhibit it. The other two inhibitors used are PKC inhibitors. Gö6976 inhibits only the classical PKC isoforms PKCα and PKCβ. Ro-31-8220 on the other hand also inhibits the novel isoforms PKCγ and PKCϵ.

Cells were exposed to these inhibitors at different concentration and also to a combination of AktVIII and either one of the two PKC inhibitors. After two minutes cells were homogenised and Western Blots were run for phospho-Akt (activated Akt) and total Akt as well as phospho- and total PKC and Rap-1. For the adhesion assays, cells were first exposed to the same stimulants and concentrations and combinations of the inhibitors as for the Western Blots. In the static assays the cells were then plated onto surfaces covered with integrin ligands (ICAM and fibronectin). After 30 minutes the cells that had not bound to the surfaces were washed off and the number of cells that were still stuck noted. In the flow assays, the leukocytes were transported over a surface covered with integrin ligands in similar conditions as in blood vessels. Images were taken at certain time intervals and the number of leukocytes that were rolling or had come to arrest on the surface were counted.

A lot of the experiments are still on-going so we only have access to a limited amount of data. The data available are Western Blot data of PKC and Akt (see Figures 7.3 and 7.4) in unstimulated cells, under BCR stimulation and after PDBu treatment. There are also some Western Blot data for Rap-1. However, the Rap-1 data are not complete and the results were different for cells from a mouse cancer B cell line as apposed to wild type mouse cells and thus not quite as easily usable for the modelling. Also adhesion data from static adhesion assays of unstimulated cells, BCR stimulated cells and PDBu treated cells are available.

As there are no data available upstream of Akt and PKC, the pathway can be simplified for modelling as shown in Figure 7.2. Also, no data are available for
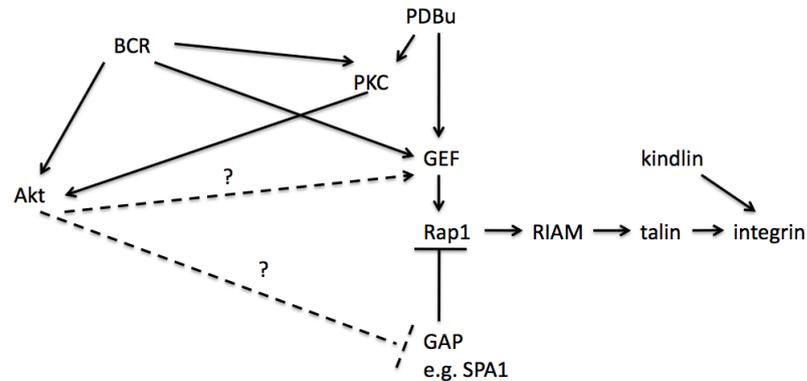
Figure 7.2: *Schematic diagram of the first step towards a simplified model of the integrin pathway. Due to the lack of raw data from Dr. Fagerholm's group for any of the elements in the signalling cascade between BCR and PKC / Akt, those elements can be neglected to decrease the degrees of freedom when trying to parameterise the model.*

species between Rap-1 and integrin. Thus this part can also be simplified.

7.3 Western Blot analysis and formulation of a mathematical model

Before starting to model the integrin pathway, the data had to be analysed and understood. So we decided to start by just considering the Western Blot data for Akt and PKC and understanding the underlying dynamics.

7.3.1 Western Blots analysis

The Western Blot data for PKC and Akt are shown in Figures 7.3 and 7.4 respectively. Initially only the control data, i.e. the data from cells that were not
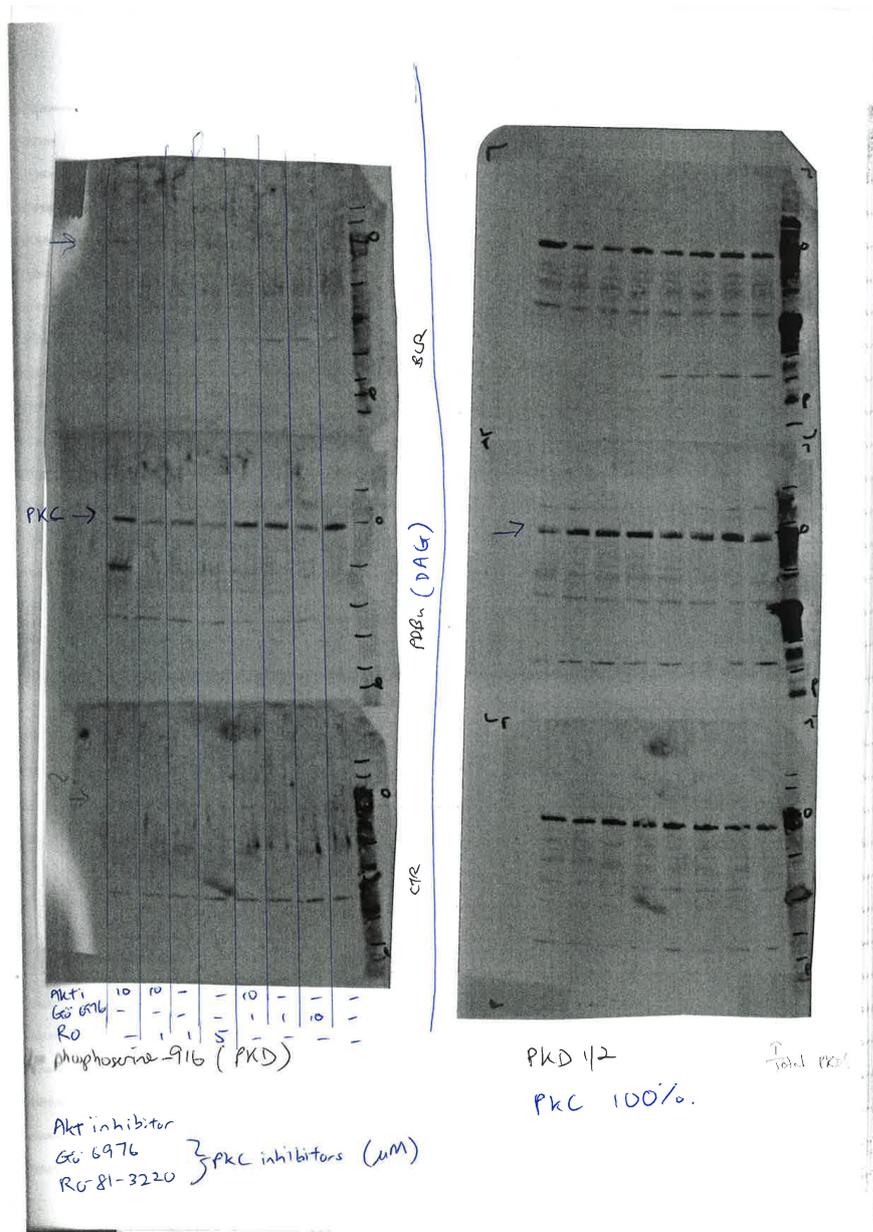
Figure 7.3: *Western Blots showing the influence of a specific PKC inhibitor (Gö 6976), a less specific PKC inhibitor (Ro-31-8220) and an Akt inhibitor (Akt VIII) on PKC activity. The bands marked with arrows on the left represent phosphorylated PKC in unstimulated B cells (bottom) , PDBu treated cells (middle) and BCR (top) stimulated cells after treatment with different combinations of the inhibitors (see bottom of panel). The bands on the right show the total PKC under the same conditions. Image provided by Hwee San Lek from Dr Fagerholm's group at the Division of Cancer Research, University of Dundee.*
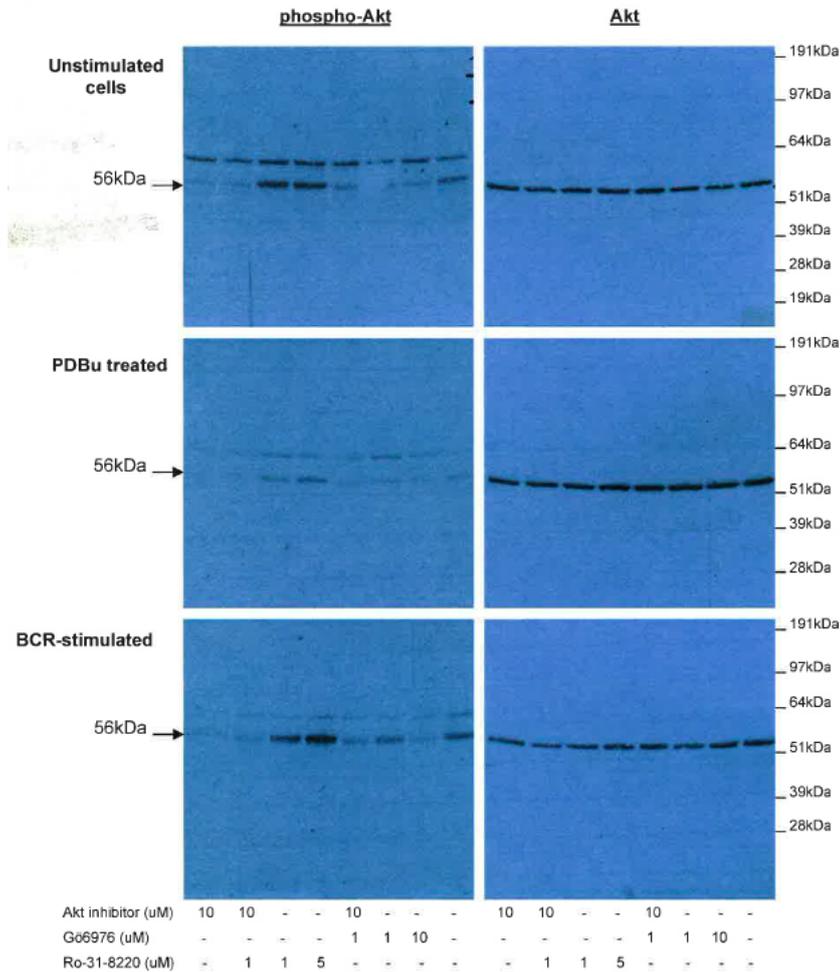
Figure 7.4: *Western Blots showing the influence of a specific PKC inhibitor (Gö 6976), a less specific PKC inhibitor (Ro-31-8220) and an Akt inhibitor (Akt VIII) on Akt activity. The bands in the left panel at 56kDa represent phosphorylated Akt in unstimulated cells (top), cells treated with PDBu (middle) and BCR stimulated cells (bottom) after treatment with different combinations of the inhibitors (see bottom of the panel). The bands on the right at a little less than 56kDa show the total Akt under the same conditions. Image provided by Hwee San Lek from Dr Fagerholm's group at the Division of Cancer Research, University of Dundee.*

exposed to inhibitors before the blotting, are taken into account. These data are shown in the right-hand band of the left panel in both figures. When looking at the PKC data, it can be seen that in unstimulated cells as well as in BCR stimulated cells, very little PKC is activated. However, the amount of activated PKC is increased and is very similar to the total PKC amount (right-hand band on right panel) after treatment with PDBu. The activation of Akt follows a different pattern. In unstimulated cells and BCR stimulated cells, about 50% of Akt seems to be activated (when compared to the total Akt). However after treatment with PDBu the activity decreases dramatically. Western Blots are very difficult to quantify. Thus in the following we will only consider three levels: no activation–low activation, medium activation, high activation. Using this, the data can be summarised as in Table 7.1. The surprising thing about these data

Stimulation/Treatment	phosphoAkt	phosphoPKC
unstimulated	medium	low
BCR	medium	low
PDBu	low	high

Table 7.1: *Summary of the Western Blots in Figures 7.3 and 7.4 under the condition that the cells are not exposed to inhibitors before the blotting.*

is that one would expect from the pathway (Figures 7.1, 7.2), that the amount of phospho-Akt would increase or stay the same after PDBu treatment. This is because Akt is activated in unstimulated cells to about the same extend as it is in BCR stimulated cells. Thus during PDBu treatment, which does not directly effect Akt, Akt activity should be at the same level as in unstimulated cells. Given that PDBu however leads to high activity of PKC and PKC activates Akt, higher activity of Akt would be plausible. However, Akt activity decreases.

To understand this behaviour, the blots from the experiments with inhibitors

have to be considered. Just as before, the results of the blots can be summarised in tables (see Tables 7.2(a),(b),(c)).

Stimulation/Treatment	phosphoAkt	phosphoPKC
unstimulated + AktVIII	low	low
BCR + AktVIII	low	low
PDBu + AktVIII	low	high

(a)

Stimulation/Treatment	phosphoAkt	phosphoPKC
unstimulated + Gö6976	low	low
BCR + Gö6976	low	low
PDBu + Gö6976	low	medium

(b)

Stimulation/Treatment	phosphoAkt	phosphoPKC
unstimulated + Ro-31-8220	high	low
BCR + Ro-31-8220	high	low
PDBu + Ro-31-8220	medium	medium

(c)

Table 7.2: *Summary of the Western Blots in Figures 7.3 and 7.4 under the condition that the cells are exposed to (a) AktVIII, (b) to Gö6976 and (c) Ro-31-8220 before the blotting.*

Table 7.2(a) shows the inhibition of Akt by the Akt inhibitor and that AktVIII has no effect on PKC activity. This is what one would expect and does not help understand the dynamics but only shows the specifity of the inhibitor. Table 7.2(b) shows that the exposure of cells to Gö6976 lowers not only PKC activity

but also Akt activity. This is easily explained by the pathway in Figures 7.1 and 7.2. Here it is shown that PKC activates Akt and thus the inhibition of PKC can decrease Akt activity. Therefore those blots confirm the pathway structure. However, the results shown in Table 7.2(c) are surprising. They show that Ro-31-8220 decreases PKC activity as it should but it also increases Akt activity. The difference between the two PKC inhibitors Gö6976 and Ro-31-8220 is that Gö6976 only inhibits the classical PKC isoforms whereas Ro-31-8220 also inhibits the novel isoforms PKCγ and PKCϵ. Thus the easiest explanation for this is that the classical PKC isoforms activate Akt as is shown in the pathways above but the novel isoforms inhibit Akt activation. This also explains why Akt decreases with PKC activation as in Table 7.1. Thus the BCR-PKC-Akt section of the pathway should be drawn as shown in Figure 7.5.
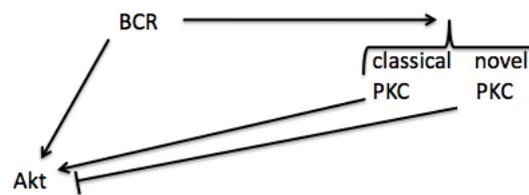
Figure 7.5: *Schematic diagram of the BCR-PKC-Akt section of the integrin pathway. The activation of Akt and atypical PKC by BCR and the inhibition of Akt by atypical PKC forms a feed-forward-loop.*

Something else can be learnt about this pathway from Table 7.1. As far as it is known, PDBu acts on classical, as well as novel, PKC isoforms. However activation of both together leads to a decrease in Akt activity which means that either more novel PKC isoforms are activated, or these novel isoforms have a higher affinity towards Akt leading to a higher inhibition rate than can be balanced out by the activation of Akt through the classical PKCs. Ignoring the weak activation of Akt leaves the signalling network of BCR, novel PKC and Akt. These three species form an interesting signalling network motif, the feed-forward loop

[Alon, 2007]. Given that the rest of the pathway is probably mainly linear, the modelling, simulation and analysis of this feed-forward loop might already give some insight into the integrin pathway dynamics.

7.3.2   Modelling the BCR-PKC-Akt feed-forward loop

The feed forward loop can be modelled using ordinary differential equations and the law of mass action. Because no quantifications of any species in the network is given, non-dimensional values between zero and one are used for the amount of protein. As specific stimulation of the BCR does not lead to an increase in Akt or PKC activity, it can be assumed that the receptor is fully activated, even when the cell is exposed to serum only. Thus it is assigned a constant value of one. Furthermore a constantly active phosphatase ($Phos$) is added to the network that dephosphorylises and thus inactivated PKC as this exists for all proteins that are activated by phosphorylation to ensure that activation only occurs after stimulation. The other species in the network are Akt, phospho-Akt ($pAkt$), novel PKC ($nPKC$) and phospho-novel-PKC ($pnPKC$) and their dynamics are governed by the following equations.

$$
\begin{aligned}
\frac{\mathrm{d}Akt}{\mathrm{dt}} &= -k_1 * BCR * Akt + k_2 * pnPKC * pAkt, \\
\frac{\mathrm{d}pAkt}{\mathrm{dt}} &= k_1 * BCR * Akt - k_2 * pnPKC * pAkt, \\
\frac{\mathrm{d}nPKC}{\mathrm{dt}} &= -k_3 * BCR * nPKC + k_4 * Phos * pnPKC, \\
\frac{\mathrm{d}pnPKC}{\mathrm{dt}} &= k_3 * BCR * nPKC - k_4 * Phos * pnPKC.
\end{aligned}
\tag{7.1}
$$

As neither production nor degradation of PKC and Akt are considered, the following assumption can be made:

$$
\begin{aligned}
nPKC_{\text{total}} &= nPKC + pnPKC &&= 1 \\
Akt_{\text{total}} &= Akt + pAkt &&= 1
\end{aligned}
\tag{7.2}
$$

With the equations (7.2), the system of four ordinary differential equations (7.1) can be simplified to a systems of two equations:

$$
\begin{aligned}
\frac{\mathrm{d}pAkt}{\mathrm{dt}} &= k_1 * BCR * (1 - pAkt) - k_2 * pnPKC * pAkt, \\
\frac{\mathrm{d}pnPKC}{\mathrm{dt}} &= k_3 * BCR * (1 - pnPKC) - k_4 * Phos * pnPKC
\end{aligned}
\tag{7.3}
$$

The Western Blots were started after two minutes as the system had supposedly reached its steady-state at that time. Thus we will also assume that a steady-state is reached after two minutes and we will solve the steady-state equations to find conditions for the parameters k_1 to k_4 under which phosphoAkt and phospho-novel-PKC reach the levels that were found in the Western Blots (see Table 7.1). In order to do so the activity levels have to be quantified. This is shown in Table 7.3.

level	non-dimensional values
no activity – low activity	$0 - 0.33$
medium activity	$0.33 - 0.66$
high activity	$0.66 - 1$

Table 7.3: *Quantification of the levels that are used to summarise the results of the Western Blots*

The steady-state system of (7.3) is given by:

$$
\begin{aligned}
0 &= k_1 * BCR * (1 - pAkt) - k_2 * pnPKC * pAkt, \\
0 &= k_3 * BCR * (1 - pnPKC) - k_4 * Phos * pnPKC.
\end{aligned}
\tag{7.4}
$$

First the second equation of the system (7.4) is solved for phospho-novel-PKC. This gives

$$
pnPKC = \frac{k_3 * BCR * nPKC_{\text{total}}}{k_3 * BCR + k_4 * Phos}.
\tag{7.5}
$$

Given that $BCR = nPKC_{\text{total}} = Phos = 1$,

$$
pnPKC = \frac{k_3}{k_3 + k_4}.
\tag{7.6}
$$

From the Western Blots it is known that at steady-state phospho-novel-PKC should exist at low levels, thus $pnPKC \in [0, 0.33]$. In the following we will consider only the ends of the interval. If $pnPKC = 0$, either k_3 has to be zero or k_4 has to tend towards infinity, both of which seem unlikely. Assigning $pnPKC$ the upper value of the interval gives

$$k_4 = \frac{0.66 * k_3}{0.33} = 2 * k_3. \tag{7.7}$$

Using this condition (7.7) on k_4 ensures that phospho-novel-PKC reaches a steady-state of 0.33. How fast this happens depends on the value of k_3. A similar condition can be found for k_1 and k_2 by solving the first equation of (7.4) for $pAkt$.

$$pAkt = \frac{k_1 * BCR * Akt_{\text{total}}}{k_1 * BCR + k_2 * pnPKC}. \tag{7.8}$$

Again using $BCR = Akt_{\text{total}} = 1$, equation (7.8) can be simplified to give

$$pAkt = \frac{k_1}{k_1 + k_2 * pnPKC}. \tag{7.9}$$

Thus, the steady-state value of phospho-Akt depends on the steady-state level of phospho-novel-PKC as well as k_1 and k_2. Assuming again that $pnPKC = 0.33$ at steady-state, and that phospho-Akt also reaches the highest possible steady-state value: $pAkt = 0.66$, the relationship between k_1 and k_2 is

$$k_2 = \frac{\frac{k_1}{0.66} - k_1}{0.33}. \tag{7.10}$$

Taking into account these two conditions (7.7) and (7.10), the solutions of (7.3) can be simulated over two minutes of real time. This can be done using a range of values for k_1 and k_3 to investigate the network dynamics leading to the required steady-states.

7.4 Computational simulation results

The system of equations (7.3) and conditions (7.7) and (7.10) were coded up in MATLAB and solved using the inbuilt function *'ode45'* which is based on an explicit Runge-Kutta algorithm. Both the parameters k_1 and k_3 were varied and it was apparent that it was not the values of these two parameters themselves but the ratio $\frac{k_1}{k_3}$ that is decisive for the network dynamics.

7.4.1 The time course profiles of phospho-Akt and phospho-novel-PKC

Figure 7.6 shows the time course of the solutions of (7.3) for $k_3 = 1$ and $k_1 = 2$ (Figure 7.6(a)) and $k_3 = 1$ and $k_1 = 5$ (Figure 7.6(b)). Although the steady-

(a)   (b)
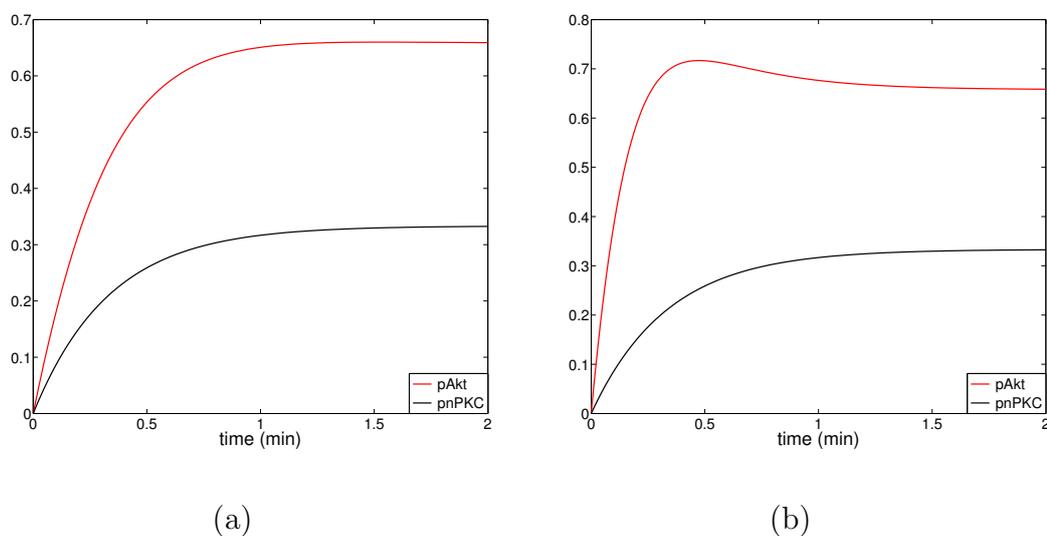
Figure 7.6: *Time course plots of the solutions of equations (7.3) using conditions (7.7) and (7.10). (a) Plot of the solution of 7.3 with $k_1 = 2$ and $k_3 = 1$. (b) Plot of the solution of 7.3 with $k_1 = 5$ and $k_3 = 1$.*

states reached after two minutes of real time are the same in both cases, the time course of the solutions for phospho-Akt is very different. For $k_1 = 2$, the amount

of active Akt slowly increases until the steady-state value is reached. However, for $k_1 = 5$ the amount of active Akt grows beyond its steady-state value initially before dropping again to the steady-state. In Figure 7.7 the solutions of (7.3) are plotted for k_1 ranging from 1 to 10 in 60 steps (step size≈ 0.167) (red curves) and $k_3 = 1$ (black curve). Here it can be clearly seen that for high values of k_1 the
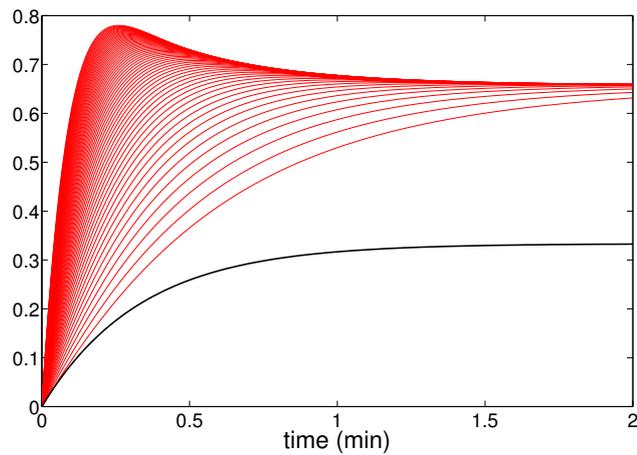
Figure 7.7: *Time course plots of the solutions of equations (7.3) using conditions (7.7) and (7.10) with k_1 ranging from 1 to 10 in 60 steps (step size ≈ 0.167) (red curves) and $k_3 = 1$ (black curve).*

amount of active Akt always grows beyond its steady-state value initially before dropping again. This 'hump' increases with increasing k_1. The occurrence of it depends on the speed at which Akt is activated compared to the speed at which novel PKC is activated and can inhibit phospho-Akt. Thus it is dependent on the ratio $\frac{k_1}{k_3}$.

7.4.2 Varying BCR activity profiles

In *in vivo* situations, the stimulation of BCR does, most likely, not happen at a constant level over a prolonged length of time. Due to the need for a tight control of signalling networks, inhibitors very often turn off the signal at the receptor level as soon as it has been passed on in order to prevent constant signalling and keep the cell sensitive to environmental cues. Thus the solution of the system
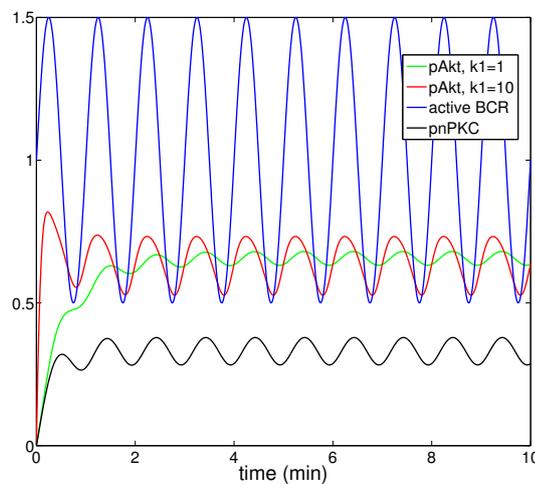
Figure 7.8: *Time course plots of the solutions of equations (7.3) where the profile of the amount of activated BCR follows a sine curve. It can be seen that the oscillatory behaviour of the amount of activated BCR is mirrored more strongly in the profile of phospho-Akt where $k_1 = 10$ than in the case where $k_1 = 1$. k_3 is again chosen to be one.*

of equations (7.3) was simulated using a sine curve, $y(t) = 0.5 * \sin(2\pi t) + 1$, to represent BCR activity. This was done using a parameter for k_1 that led to a 'hump' in the phospho-Akt profile as well as one that did not. The parameters were $k_1 = 10$ and $k_1 = 1$, respectively. The results are shown in Figure 7.8. It can be seen in this figure that phospho-Akt and phospho-novel-PKC follow the oscillatory behaviour of active BCR. However, dependent on the value of k_1 the

oscillations of phospho-Akt are more or less pronounced. In the case of $k_1 = 1$ the amplitude of the oscillations is very small, whereas in the case of $k_1 = 10$, the amplitude of the oscillations is about a quarter of the amplitude of the oscillations of active BCR and they show a hump again at early points in time.

A pulse of BCR activity is probably even more realistic than an oscillatory behaviour of active BCR. Figure 7.9 shows the results of simulations where this is the case. Here BCR is fully activated for one minute and then completely
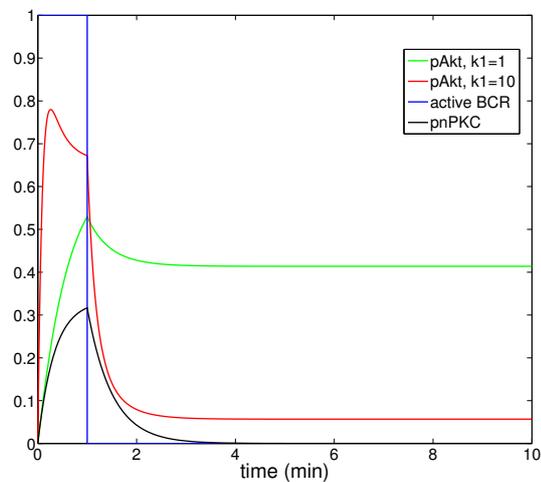
Figure 7.9: *Time course plots of the solutions of equations (7.3) where the profile of the amount of activated BCR follows a step function. This is a representation of what happens* in vivo *when short pulses of stimulation are given. It can be seen that the step-like behaviour of the amount of activated BCR is mirrored in the profile of phospho-Akt where $k_1 = 10$, however for $k_1 = 1$ the initial increase of activity is not as strong as in the former case but there is also hardly any decrease in activity after BCR is turned off.*

turned off. In the case where $k_1 = 10$, Akt activity increases very quickly before dropping slightly due to the influence of active PKC. However, as soon as BCR is inactivated, Akt activity drops to a very low level. In the case of $k_1 = 1$ the

time course profile of Akt is entirely different. Akt activity increases slowly during the time that BCR is fully active and then only drops slightly when BCR is inactivated before reaching a steady-state of activity which is still greater than 40%. These different types of behaviour can easily be explained. Akt activity responds quickly to BCR activity changes when $k_1 = 10$ and thus it increases quickly when BCR is active. Due to the second term in the first equation of (7.3), it also decreases quickly when active BCR equals zero and phospho-Akt and phospho-novel-PKC levels are high. In contrast, Akt activity increases much more slowly under active BCR with $k_1 = 1$ and thus its levels are not so high when BCR is inactivated. Therefore less phospho-Akt is inactivated by phospho-novel-PKC. Phospho-novel-PKC levels decrease as well due to the loss of BCR activity and high levels of active Akt remain. Though the behaviour is very easy to explain in terms of the model, it is still very interesting and might hold some real insight into biological behaviour.

7.5 Discussion

In this chapter a pathway model was developed for a crucial part of the 'inside-out' activation of integrins. The idea was that this model should be based on on existing data for leukocyte adhesion processes. To start off with as simple a model as possible, the adhesion data was not taken into account, but only the Western Blot data for Akt and PKC was modelled under different conditions. However, even the data of just these two species could not be explained by the structure of the pathway which was assumed to be true (see Figure 7.1). It was clear that PKC could not just have an activating role but must also have an inhibiting role upstream of Akt. The literature concerning this subject confirmed that classical as well as novel PKC isoforms exist and given the inhibitors used in the experiments it all pointed towards the hypothesis that novel PKC isoforms

inhibit Akt. Taking this into account, the simplest model that could explain the given data, was a BCR-Akt-novel-PKC feed-forward loop.

Since the system was said to have reached a steady-state after two minutes, the point of time when the Western Blots were started, the steady-state equations were solved for phospho-Akt and phospho-novel-PKC. Using the levels of activity shown by the Western Blots for these two species, conditions for the parameters k_2 and k_4 were found depending on k_1 and k_3, respectively. As the values of k_1 and k_3 themselves did not determine the behaviour of the results, but the ratio of k_1 to k_3, the equations were solved in MATLAB and plotted for a variety of k_1 to k_3 ratios, where k_3 equals one. Two different types of behaviour were found. For low ratios, the amount of phospho-Akt increased until it reached the steady-state value. However for high ratios, the amount of phospho-Akt increased much more quickly, growing beyond its steady-state value and then dropped again to steady-state.

The difference between these two cases became clearer when including more realistic oscillatory and pulse-like activity profiles of BCR rather than constant activation. Here we found that an oscillatory activity profile of BCR also leads to oscillations in phospho-Akt. The amplitude of these oscillations was much larger in the case of high k_1 to k_3 ratios rather than small ones. Even more interesting were the simulation results when a step function was used as BCR activity profile to resemble a pulse of BCR activity. For $k_1 = 10$, Akt activity increased very quickly during the pulse of BCR activity before dropping slightly due to the influence of active PKC. However, as soon as BCR was inactivated, Akt activity dropped to a very low level. In the case of $k_1 = 1$, Akt activity increased slowly during the time that BCR was fully activated and decreased only slightly after BCR was inactivated. Although the behaviour is very easy to explain in terms of the model and thus not counterintuitive, it is still very interesting and might

206

hold some real insight into biological behaviour.

It has been shown in a different context that the Akt pathway has low-pass filter characteristics [Fujita et al., 2010]. A low-pass filter is an electronic filter that passes low-frequency signals but reduces the amplitude of signals with frequencies higher than a cut-off frequency. Fujita et al. [2010] demonstrated with experiments, as well as with mathematical modelling, that the Akt pathway exhibits this behaviour. Their work shows that high frequency signals of the EGF receptor travel downstream and result in a similar time course profile in phospho-Akt. However, phospho-Akt's downstream target is poorly induced. On the other hand, weak, sustained signalling by EGFR strongly induces the Akt target. If this is also the case in the pathway considered here, the parameter k_1 could be highly relevant to characterise certain cell phenotypes. Cells with a low value for k_1, relative to k_3, translate a pulse of BCR activity to weaker but sustained Akt activity, which would lead to high activity of downstream targets, in this case Rap-1 and further downstream integrin. However, cells with high values for k_1, translate a pulse of BCR activity into a pulse of Akt activity which would lead to only weak induction of the downstream target Rap-1 and thus only a weak induction of integrin. Thus gathering time-course data of different cells under BCR stimulation and identifying those with a high or low ratio $\frac{k_1}{k_3}$ might help to understand the results of the adhesion assays more easily.

This model might also help to understand the different cell behaviours seen *in vivo* in terms of cell adhesions and the integrin pathway. Even if the pathway here does not have low-pass filter characteristics, the emerging behaviour is nonetheless interesting as the Akt signal generated in cells where the k_1 to k_3 ratio is small is much greater when seen over longer periods of time than when it is large. Thus, depending on how the signal is passed on downstream, the cell specific parameters might be decisive for the cell phenotype.

207

What the modelling in this chapter has shown is that in order to understand cell behaviour, steady-state data is not sufficient in a lot of cases. Time-course data also needs to be gathered. It would be very interesting to get time course data for the pathway modelled here and compare the modelling results to it. Hopefully this will be possible in the near future. Once this simple pathway is verified, it can be extended to include the treatment with PDBu as has been done in experiments and other species such as classical PKC, Rap-1 and integrin. The adhesion data available can then also be used to further parameterise the model which could then help to explain more about the experimental results.

In order to develop a complete model of integrin signalling in leukocytes, outside-in signalling under static and flow conditions will also have to be studied in detail and added to the current model. Including this in the single-cell model of Chapter 4 would give a multiscale model of leukocyte movement through tissue. The replacement of the matrix elements by endothelial cells lining a blood vessel could further give a modelling environment to study leukocyte rolling, adhesion, arrest and extravasation. A complete model of the integrin pathway would also be very useful to be included in the multi-cell model of the previous chapter. In addition, not only the signalling which is important in leukocytes, but also E-cadherin and integrin cross-talk would have to be taken into account in order to study the processes involved in cancer cell adhesion more accurately and in more detail.

Chapter 8

Conclusions and Future Work

The aim of this thesis was to develop a multiscale model of cell migration and cancer cell invasion in a two-dimensional domain by extending the work of Ramis-Conde et al. [2008b]. This model was supposed to couple explicit cell-matrix interactions with cell-cell interactions and intracellular dynamics in order to study the combined influence of these different factors on migration and invasion. The idea was to consider cells and matrix fibres as individual elements, which have a mutual interaction such that the matrix fibres guide the cell movement but are in turn reorientated by the cells moving across them. Additionally, intra- and intercellular dynamics of cell-cell adhesion, repulsion and cell division were to be taken into account. This has been done in stages in the different chapters of this thesis and with each model, data analysis techniques have been developed in order to be able to quantify the simulation results and make them comparable to experimental data. Furthermore, Western Blot data concerning the integrin signalling pathway has been studied and used to develop a model of part of this complex pathway.

In Chapter 4 we formulated a modelling framework for single cell migration on

two-dimensional matrices in which the cell as well as the matrix fibres are individual elements or agents. In this model the focus was on the most fundamental processes underlying cell migration at the level of cell–matrix interactions. Using this approach, we investigated the influence of matrix stiffness on cell migration and found that the reorientation of the matrix fibres due to cell traction forces might be an important part of this process. We also ran simulations to test whether the model could reproduce experiments that showed the preference of cells for stiffer matrices. The results agreed with these experiments and suggest that matrix reorientation, or the lack thereof on stiff matrices, may be an important factor in durotaxis. This is particularly interesting as the physical interactions between cells and matrix cannot be isolated in experimental settings. Therefore in those setups it is difficult to distinguish between the roles of the physical interactions and the intracellular signalling pathways coupled to these. The model makes it possible to ignore the intracellular signalling. It shows that purely the interactions between cells and the matrix, together with the physical structure of the matrix, can explain the preference that cells have for stiffer rather than softer matrices.

Furthermore the relationship between persistence time as well as actual cell speed, and maximum cell speed and matrix fibre length or matrix density were examined and showed a nonlinear relationship. Finally we also investigated the sensitivity of the results towards a change of certain, so far unconsidered, parameters and found that perturbations in those parameters had little impact on the results.

In Chapter 5 we presented a natural extension to the model in Chapter 4 by placing two cells into the domain. The cells were then tracked over three days of real time and their behaviour classified by calculating the distribution of the lengths of time that the cells followed each other. This was done under different

conditions. First the matrix stiffness and the initial distance between the cells were altered. Then different fibre lengths and densities were considered before finally both, the fibre lengths and densities, were varied. Summing up the results, the simulations showed that fibre length and matrix stiffness seem to have an impact on the behaviour of two cells placed on the extracellular matrix whereas fibre density mainly affects the variability of the lengths of time that cells follow each other rather than the median of these values.

In Chapter 6 we formulated a multiscale model of cell colonies and used it to study the transition from healthy cell behaviour to invasive, cancer-like behaviour. We developed the model on the basis of the cell migration model from previous chapters by including cell division as well as intra- and intercellular cell-cell interaction dynamics. First we studied the intracellular E-cadherin–β-catenin dynamics and developed an ordinary differential equation model to describe the E-cadherin–β-catenin complex formation and disruption as well as the transition of these complexes between the cell's interior and the different cell-cell contact sites. The model parameters were estimated by fitting the number of E-cadherin–β-catenin bonds between two cells to cell-cell separation force data. We obtained a good fit to the data after some adjustments to the model and it proved essential to assume that there is no limit on the number of bonds that can form between two cells and that the E-cadherin–β-catenin complexes are redistributed upon attachment of further cells. This led to the assumption that the forces that are measured between a pair of cells in isolation is higher than the forces between any two cells in an epithelial layer.

In addition to the intracellular dynamics and resulting adhesion, we included a simple cell cycle in each cell, which allows the cells to divide when they are in M-phase, and cell-cell repulsion governed by the Hertz model. Thus we could

consider growing cell colonies. We ran simulations of these growing colonies and varied the value of the endocytosis rate of E-cadherin–β-catenin complexes. We also introduced a threshold for the level of pressure, i.e. the magnitude of the repulsive force, under which a cell could still divide and studied its effects on colony growth. Thirdly, we worked around possible discrepancies between the modelling technique and the biology by assuming that the active force between two cells whose contact site had a diameter larger than one sixth of the cell's circumference was either zero or equal to the repulsive force alone. These three variations of the model were combined in simulations and the colonies were studied after three and after seven days. The results were compared to experimental findings regarding the distribution of the number of neighbours of cells in an epithelial layer. This led to the observation that there indeed has to be a threshold on the pressure under which cells divide. Furthermore, the results showed that the active force between two cells has to be just the repulsive force when cells get too close and zero when they are at a distance that we consider to be their 'natural state' in epithelia. This means that the forces exerted by the E-cadherin–β-catenin bonds are only active when the cells are further apart than they would be in naturally occurring epithelial sheets and thus the forces prevent separation rather than leading to closer adhesion. From a biological point of view this is a sensible conclusion as a constant tugging by the adhesion and repulsion forces to keep a balance would be wasting the cell's energy.

The simulations also gave some ideas regarding the development of invasive colonies. Some of the computational simulation results showed that invasion could be initiated without the down-regulation of E-cadherin and up-regulation of cell-matrix interactions. A relatively fast internalisation of E-cadherin–β-catenin complexes after cell-cell bond disruption and a highly deformable cytoskeleton seemed to be the main requirements. Due to the lack of cell-matrix interactions

in these simulations, this observed behaviour was, however most likely regulated by cell division and would have to be investigated further.

Finally we conducted simulations in which we combined the multiscale model of cell-cell interactions with the model of cell-matrix interactions. In these simulations we varied the E-cadherin expression level and profile, the integrin expression level as well as the pressure threshold for proliferation. For cell populations with homogenous expression levels of 75% and 50% as well as those with heterogeneous expression profiles, we ran 15 simulations for each combination of integrin expression levels of 10% and 50% with pressure thresholds on proliferation of 13000pN, 20124pN, 36966pN and 56916pN. These simulations revealed a large variety of behaviours from the development of large, contiguous cell clusters to the invasion of the environment by small fingering patterns, small cell clusters and single cells. It was clear from these results that especially the combination of E-cadherin and integrin expression levels control the integrity or invasiveness of the cell colonies and that the proliferative activity of the cells has little impact on the cell colony development. Only in the case of a high E-cadherin expression level, a low integrin expression level and a high pressure threshold for division, did the cells divide too much for the two dimensional space they were occupying and led to failure of the simulations. In a three dimensional setting this would have probably led to the formation of a hyperplastic cell colony.

In Chapter 7 a pathway model was built for a crucial part of the 'inside-out' activation of integrins. The idea was that this model should be based on existing data for leukocyte adhesion processes. The data clearly showed that PKC could not just have an activating role but must also have an inhibiting role upstream of Akt. Taking this into account, the simplest model that could explain

the given data, was a BCR-Akt-novel-PKC feed-forward loop. Solving the ordinary differential equations of this model numerically led to two different types of behaviour. Either the amount of phospho-Akt increased until it reached the steady-state value or the amount of phospho-Akt increased much more quickly, growing beyond its steady-state value and then dropping again to steady-state. The difference between these two cases became clearer when including more realistic oscillatory and pulse-like activity profiles of BCR rather than constant activation. Here we found that an oscillatory activity profile of BCR also leads to oscillations in phospho-Akt with their amplitude varying in size for different parameter values. Even more interesting were the results of the simulation when a step function was used as BCR activity profile to resemble a pulse of BCR activity. For some parameter values, Akt activity increased very quickly during the pulse of BCR activity before dropping slightly due to the influence of active PKC. However, as soon as BCR was inactivated, Akt activity dropped to a very low level. For other values Akt activity increased slowly during the time that BCR was fully activated and decreased only slightly after BCR was inactivated. As has been shown in previous work, it is possible that also here the Akt pathway has low-pass filter characteristics. If this is the case, the parameter choice could be highly relevant to characterise certain cell phenotypes. This model might help to understand the different cell behaviours seen *in vivo* in terms of cell adhesions and the integrin pathway. Even if the specific pathway considered here does not have low-pass filter characteristics, the emerging behaviour is nonetheless interesting and highlights that the cell specific kinetic parameters might be decisive for the cell phenotype. This model definitely showed the importance of gathering not only steady-state data but also time course data in order to understand signalling pathways.

A lot of extensions are possible to the work presented here to make it easily adjustable to different *in vitro* setting and also to move it towards *in vivo*. Firstly, there are many more possibilities to explore with the multiscale cell invasion model presented in Chapter 6. Due to the computational constraints, not all the possible combinations of the varied key parameter values could be simulated. This is definitely something that will be done in the future. It will be interesting to see how further combinations of different levels of E-cadherin expression, integrin expression and proliferative activity in homogenous cell populations and cell populations with different levels of heterogeneity will effect colony growth and spread in the surrounding matrix. In experiments this spread has been shown to not only occur in terms of individual cells or small clusters, but also as prolonged 'fingers' extending from the main mass of cells. We saw patterns resembling this at a very initial stage in some simulations but it would be interesting to see whether the simulations can also generate more mature and definite fingering patterns. For this it will be necessary to use or develop a further data analysis technique which can give a measure of the shape of a cell colony. Fractal dimensions of cell colony-matrix interfaces and scaling analysis have been used in previous work [Bru et al., 2003] and might be applicable here as well. It has been mentioned (private communications with biologists) that cells might have to polarise in groups in order for these strands or sheets of invading cells to appear. The way this seems to be controlled is by intracellular signalling that encourages cells to follow certain leader cells. This is something that can also be tested with the model. We have done some preliminary work on this and Figure 8.1 shows the results. Figure 8.1(a) shows the invasion of the matrix by a small group of cells which is led by three cells polarised in the same direction and other cells which simply follow the cells leading them. The group detaches due to down-regulation of E-cadherin at contact sites between those cells that are affected by this group
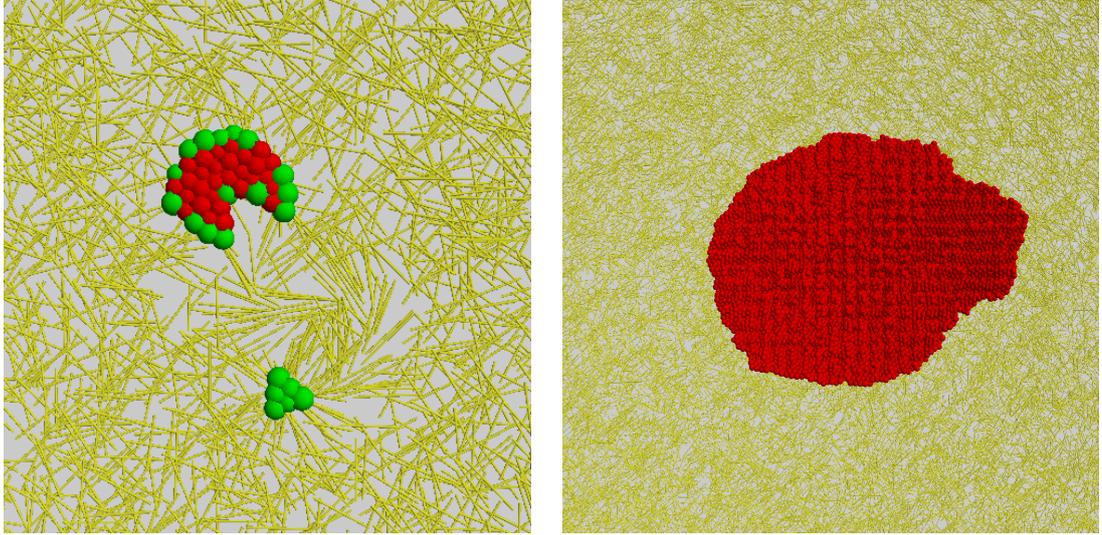
(a)   (b)

Figure 8.1: *Images of preliminary work that includes group polarisation in the cell invasion model. (a) shows a close-up image of a small group of cells that has polarised as a cluster and has down-regulated the E-cadherin at contact sites between cells belonging to the cluster and cells of the remaining colony. Cells coloured in green have acquired front-rear polarity whereas the red cells have apical-basal polarity. (b) shows an image of a large cell colony in which some cells at the right hand rim have polarised and signal to a large group of cells to follow their lead. In this case the E-cadherin bonds stay intact between cells of the polarised group and the remaining colony. Here no distinction is made in colour between the differently polarised cells in order to increase visibility of the colony shape.*

polarisation and those that are not. Figure 8.1(b) shows the extension of the mass of cells towards one side. Here again some cells polarise in the same direction and signal to other cells within a certain distance of them to follow them. In this case that happens without E-cadherin down-regulation. This group polarisation is a process that can be extended and included in the model in more detail to tie in with the current experimental interest in this area.

Clearly integrin expression is an important factor and the model in Chapter 7 showed how different the behaviour of the signalling pathway can be in different cells. Therefore extending this pathway model to a complete model of integrin signalling also including the integrin-E-cadherin crosstalk [Tsai and Kam, 2009, Danen, 2012, Goitre et al., 2012] and embedding it in the multiscale cell model will lead to the possibility of simulating cell populations with even more refined heterogeneity.

It would, furthermore, be interesting to vary the matrix characteristics that have been found to have an effect on cell migration in Chapters 4 and 5, in the multi-scale cell population model. For example a change in stiffness in different parts of the matrix would add another layer of complexity and could give very interesting results. It would also be very interesting to explore the effect of a more detailed model of the extracellular matrix on the simulation results. Such a model could, for example, take into account fibre-fibre interactions and the resulting matrix deformability and elasticity at a non-local level.

It will, however, be important to have more experimental data from different cell lines and tumours in order to run simulations that are biologically meaningful and can help understand the behaviour of cancerous cells. As mentioned in Chapters 4 and 5, more data on microscopic details on the extracellular matrix in different tissues and measurements of cell-matrix fibre interactions would also be useful to validate or adjust the model further to enhance the predictive potential. Such

a model could then also be used for the testing of different drugs and also for predicting combinations that inhibit cell movement and invasion while also decreasing the main cell mass.

When considering the testing of treatments and predicting patient outcome, the next natural step would be to consider the model in thee dimensions and we have examined this possibility. However, multiple problems arise when adding the third dimension. In the current model it is assumed that the domain in which the cell-matrix interactions take place is the circular base of the cell. In the multiscale model in which cells are modelled as spheres, the deformation of the migrating cells is not taken into account and the cell-matrix contact area is implicitly assumed to be a circular region of the same diameter as the cell. However, if this were to be moved to a three dimensional setting, the surface of the spheres representing the cells would be the contact areas and thus the mathematical problem of cell-matrix contacts would be a different one.

Not unrelated is also the problem that potential functions would have to be developed for the cell-matrix interactions. This could not be based on the Hertz model as it does not extend to the interactions of spheres with much thinner cylindrical elements. Thus, although moving this model to a three dimensional setting would be very interesting and has a lot of potential to help understand *in vivo* cancer growth and spread and could give the possibility of testing treatments and surgery, it would require quite extensive further modelling efforts. In addition, the simulations of the multiscale model in Chapter 6 have highlighted the drawback of this modelling technique by leading to simulations that take a very long time to run in the case of the development of large cell colonies. When considering the growth of actual tumours, whether in 3D or as a two-dimensional slice, the number of cells in the simulations would grow to millions. This would no longer be feasible without recourse to a supercomputer. Therefore, in future,

the model could be extended to a hybrid model. This has been done in previous work by Kim et al. [2007]. In their approach the central region of the tumour is modelled using a continuous description and only the outer rim is modelled at the single cell-level. This could lead to a model that still shows the details of the tumour-matrix interface but loses some of the details on the level of cell-cell interactions as a trade-off for computational speed and efficiency. Another approach could be to use all the information from the mechanistic model developed here, of how the results depend on the parameter values, and use this to develop an empirical model. This might be useful when investigating drug action on a regular basis or as a large scale study.

In conclusion, the models developed in this thesis have been able to give insight into cell migration, invasion and processes related to these. However, further investigations with, and validations of, these models are necessary in order to expand their capabilities from being descriptive to being predictive. Further extensions are possible which will allow the work developed here to be used as a tool for understanding a wide range of *in vitro* experiments and predicting their outcome. In future these models might even help to understand and predict *in vivo* cell migration and cancer spread.

Finally, data analysis techniques have been developed in this thesis in order to quantify single and two cell migration characteristics as well as analyse cell colony data. These techniques are universally usable for simulation as well as experimental data.

Bibliography

C. L. Adams, W. J. Nelson, and S. J. Smith. Quantitative analysis of the cadherin-catenin- actin reorganization during development of cell-cell adhesion. *Journal of Cell Biology*, 135:1899–1911, 1996.

C. L. Adams, Y.-Y. Chen, S. J. Smith, and W. J. Nelson. Mechanisms of epithelial cell-cell adhesion and cell compaction revealed by high-resolution tracking of e-cadherin–green fluorescent protein. *Journal of cell Biology*, 142:1105–1119, 1998.

B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular biology of the cell*. Garland Science, third edition, 2002.

R. Alon and K. Ley. Cells on the run: shear-regulated integrin activation in leukocyte rolling and arrest on endothelial cells. *Current Opinion in Cell Biology*, 20:525–532, 2008.

U. Alon. *An introduction to systems biology. Design principles of biological circuits*. Chapmann & Hall/CRC, 2007.

W. Alt. Correlation analysis of two-dimensional locomotion paths. In W. Alt and G. Hoffmann, editors, *Biological Motion*, pages 254–268. Springer-Verlag, Berlin, 1990.

A. Anderson and K. A. Rejniak, editors. *Single-cell-based models in biology and medicine.* Birkhaeuser, 2007.

A. R. A. Anderson, M. A. J. Chaplain, E. L. Newman, R. J. C. Steele, and Thompson. Mathematical modelling of tumour invasion and metastasis. *Journal of Theoretical Medicine*, 2:129–154, 2000.

B. Baum, J. Settleman, and M. P. Quinlan. Transition between epithelial and mesenchymal states in development and disease. *Seminars in Cell and Developmental Biology*, 19:294–308, 2008.

A. Bergman and K. Zygourakis. Migration of lymphocytes on fibronectin-coated surfaces: temporal evolution og migratory parameters. *Biomaterials*, 20:2235–2244, 1999.

S. Bernard and H. Herzel. Why do cells cycle with a 24 hour period? *Genome Informatics*, 17:72–79, 2006.

T. Brabletz, A. Jung, S. Reu, M. Porzner, F. Hlubek, L. A. Kunz-Schughart, R. Kneuchel, and T. Kirchner. Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment. *Proceedings of the National Academy of Science USA*, 98(18):10356–10361, 2001.

A. Bru, S. Albertos, J. L. Subiza, J. L. Garcia-Asenjo, and I. Bru. The universal dynamics of tumor growth. *Biophysical Journal*, 85:2948–2961, 2003.

D. M. Bryant and J. L. Stow. The ins and outs of e-cadherin trafficking. *Trends in Cell Biology*, 14:427–434, 2004.

H. Byrne and D. Drasdo. Individual-based and continuum models of growing cell populations: a comparison. *Journal of Mathematical Biology*, 58:657–687, 2009.

221

M. A. J. Chaplain and G. Lolas. Mathematical modelling of cancer cell invasion of tissue: The role of the urokinase plasminogen activation system. *Mathematical Models and Methods in Applied Sciences*, 15:1685–1734, 2005.

A. Chauviere, T. Hillen, and L. Preziosi. Modelling cell movement in anisotropic and heterogeneous networks tissues. *Networks and Heterogenous Media*, 2: 333–357, 2007.

Y.-T. Chen, D. B. Stewart, and W. J. Nelson. Coupling assembly of the e-cadherin/β-cateinin complex to efficient endoplasmatic reticulum exit and basal-lateral membrane targeting of e-cadherin in polarized mdck cells. *J. Cell Biol.*, 144:687–699, 1999.

Y.-S. Chu, W. A. Thomas, O. Eder, E. Pincet, J. P. Thiery, and S. Dufour. Force measurements in e-cadherin–mediated cell doublets reveal rapid adhesion strengthened by actin cytoskeleton remodeling through rac and cdc42. *J. Cell Biol.*, 167:1183–1194, 2004.

J. C. Dallon, J. A. Sherratt, and P. K. Maini. Mathematical modelling of extracellular matrix dynamics using discrete cells: Fibre orientation and tissue regeneration. *J. Theor. Biol.*, 199:449–471, 1999.

E. H. Danen. Abstract 5186: A new mode of integrin-e-cadherin cross talk that modulates invasion and metastasis. *Cancer Research*, 72:Supplement 1, 2012.

R. B. Dickinson and T. Tranquillo. Optimal estimation of cell movement indices from the statistical analysis of cell tracking data. *AIChE Journal*, 39:1995–2010, 1993.

P. A. DiMilla, K. Barbee, and D. A. Lauffenburger. Mathematical model for the effects of adhesion and mechanics on cell migration speed. *Biophys. J.*, 60: 15–37, 1991.

D. Drasdo and S. Hoehme. A single–cell–based model of tumor growth *in vitro*: monolayers and spheroids. *Physical Biology*, 2:133–147, 2005.

O. du Roure, A. Saez, A. Buguin, R. H. Austin, P. Chavrier, P. Siberzan, and B. Ladoux. Force mapping in epithlial cell migration. *Proc. Natl. Acad. Sci. USA*, 102:2390–2395, 2005.

S. Etienne-Mannevielle. Polarity proteins in migration and invasion. *Oncogene*, 27:6970–6980, 2008.

I. J. Fidler. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat. Rev. Cancer*, 3:453–458, 2003.

R. I. Freshney, editor. *Culture of epithelial cells*. Wiley-Liss, 1992.

P. Friedl and K. Wolf. Tumour-cell invasion and migration: Diversity and escape mechanisms. *Nature Reviews Cancer*, 3:362–374, 2003.

P. Friedl and K. Wolf. Proteolytic interstitial cell migration: a five-step process. *Cancer Metastasis Review*, 28:129–135, 2009.

P. Friedl and K. Wolf. Plasticity of cell migration: a multiscale tuning model. *J. Cell Biol.*, 188:11–19, 2010.

P. Friedl, K. Maaser, C. E. Klein, B. Niggemann, G. Krohne, and K. S. Zanker. Migration of highly aggressive mv3 melanoma cells in 3-dimensional collagen lattices reults in local matrix reorganisation and shedding of $\alpha 2$ and $\beta 1$ integrins and cd44. *Cancer Res.*, 57:2061–2070, 1997.

J. Friedrichs, A. Taubenberger, C. Franz, and D. Muller. Cellular remodelling of individual collagen fibrils visualized by time-lapse afm. *J. Mol. Biol.*, 372: 594–607, 2007.

Y. Fujita, K. A. ad Toyoshima, S. Uda, Y. Ozaki, H. Kubota, and S. Kuroda. Decoupling of receptor and downstream signals in the akt pathway by its low-pass filter characteristics. *Science Signaling*, 3(132):ra56, 2010.

M. H. Gail and C. W. Boone. The locomotion of mouse fibroblasts in tissue culture. *Biophys. J.*, 10:980–993, 1970.

J. Galle, M. Loeffler, and D. Drasdo. Modeling the effect of deregulated proliferation and apoptosis on the growth dynamics of epithelial cell populations in vitro. *Biophys. J.*, 88:62–75, 2005.

R. A. Gatenby and E. T. Gawlinski. A reaction-diffusion model of cancer invasion. *Cancer Research*, 56:5745–5753, 2006.

M. C. Gibson, A. B. Patel, R. Ngpal, and N. Perrimon. The emergence of geometric order in proliferating metazoan epithelia. *Nature*, 442:1038–1041, 2006.

A. Giese, R. Bjerkvig, M. E. berens, and M. Westphal. Cost of migration: Invasion of malignant gliomas and implications for treatment. *Journal of Clinical Oncology*, 21:1624–1636, 2003.

L. Goitre, B. Pergolizzi, E. Ferro, L. Trabalzini, and S. F. Retta. Molecular crosstalk between integrins and cadherins: do reactive oxygen species set the talk? *Journal of Signal Transduction*, 2012:Article ID 807682, 12 pages, 2012. doi:10.1155/2012/807682.

B. M. Gumbiner. Regulation of cadherin-mediated adhesion in morphogenesis. *Nature Reviews Molecular Cell Biology*, 6:622–634, 2005.

W. Guo and F. G. Giancotti. Integrin signalling during tumour progression. *Nature Reviews Molecular Cell Biology*, 5:816–826, 2004.

B. Harms, G. Bassi, A. Horwitz, and D. Lauffenburger. Directional persistence of egf-induced cell migration is associated with stabilisation of lamellipodal protrusions. *Biophys. J.*, 88:1479–1488, 2005.

H. Hatzikirou, G. Breier, and D. A. Cellular automaton models for tumor invasion. In *Encyclopaedia of Complexity and Systems Science*. Springer, 2009.

H. Hertz. Ueber die beruehrung fester elastischer koerper. *Journal fuer die reine und angewandte Mathematik*, 92:156–171, 1881.

T. Hillen. M^5 mesoscopic and macroscopic models for mesenchymal motion. *J. Math. Biol.*, 53:585–616, 2006.

L. Hinck, I. Nathke, J. Papkoff, and W. J. Nelson. Dynamics of cadherin/catenin complex formation: Novel protein interactions and pathways of complex assembly. *J. Cell Biol.*, 125:1327–1340, 1994.

J. Huelsken, W. Birchmeier, and J. Behrens. E-cadherin and apc compete for the interaction with β-catenin and the cytoskeleton. *Journal of Cell Biology*, 127: 2061–2069, 1994.

J. Hulsken and J. Behrens. The wnt signalling pathway. *Journal of Cell Science*, 115:3977–3978, 2000.

A. Huttenlocher. Cell polarization mechanisms during directed cell migration. *Nature Cell Biology*, 7:36–337, 2005.

D. E. Ingber. Cancer as a disease of epithelial-mesenchymal interactions and extracellular matrix regulation. *Differentiation*, 70:547–560, 2002.

D. E. Ingber. Can cancer be reversed by engineering the tumour microenvironment. *Seminars in cancer biology*, 18(5):356–364, 2008.

F. Jiang, K. Khairy, K. Poole, J. Howard, and D. Muller. Creating nanoscopic collagen matrices using atomic force microscopy. *Microsc. Res. Tech.*, 64:435–440, 2004.

R. L. Juliano. Signal transduction by cell adhesion receptors and the cytosleketon: functions of integrins, cadherins, selectins, and immumoglobulin-superfamily members. *Annual Reviews of Pharmacology and Toxicology*, 42:283–323, 2002.

R. Kalluri and R. A. Weinberg. The basics of epithelilal-mesenchymal transition. *The Journal of Clinical Investigation*, 119:1420–1428, 2009.

Y. Kang and J. Massague. Epithelial-mesenchymal transitions: Twist in development and metastasis. *Cell*, 118:277–278, 2004.

Y. Kim, M. A. Stolarska, and H. G. Othmer. A hybrid model for tumor spheroid growth in vitro i: theoretical development and early results. *Math. Mod. Meth. Appl. Sci.*, 17:1773–1798, 2007.

R. Kirmse, H. Otto, and T. Ludwig. Interdependency of cell adhesion, force generation and extracellular proteolysis in matrix remodeling. *J. Cell Sci.*, 124:1857–1866, 2011.

J. A. Lagarias, J. C.and Reeds, M. H. Wright, and P. E. Wright. Convergence properties of the nelder-mead simplex method in low dimensions. *SIAM J. Optim*, 9:112–147, 1998.

L. D. Landau and E. M. Lifschitz. *Theory of elasticity*. Pergamon Press, 1959.

D. A. Lauffenburger and A. F. Horwitz. Cell migration: A physically integrated molecular process. *Cell*, 84:359–369, 1996.

T. L. Le, A. S. Yap, and J. L. Stow. Recycling of e-cadherin: A potential mechanism for regulating cadherin dynamics. *J. Cell Biol.*, 146:219–232, 1999.

J. M. Lee, S. Dedhar, R. Kalluri, and E. W. Thompson. The epithelial-mesenchymal transition: new insights in signalling, development and disease. *J. Cell Biol.*, 172:973–981, 2006.

D. Leith. Drag on nonspherical objects. *Aerosol Science and Technology*, 6: 153–161, 1987.

C. A. Lemmon, C. S. Chen, and L. H. Romer. Cell traction forces direct fibronectin matrix assmeply. *Biophys. J.*, 96:729–738, 2009.

R. Levayer and T. Lecuit. Breaking down emt. *Nature Cell Biology*, 10:757–759, 2008.

C.-M. Lo, H.-B. Wang, M. Dembo, and Y.-L. Wang. Cell movement is guided by the rigidity of the substrate. *Biophys. J.*, 79:144–152, 2000.

H. Lodish, A. Berk, C. A. Kaiser, M. Krieger, A. Bretscher, H. Ploegh, A. Amon, and M. P. Scott. *Molecular cell biology.* W.H. Freeman and Company, seventh edition, 2012.

T. Ludwig, R. Kirmse, K. Poole, and U. Schwarz. Probing cellular microenvironments and tissue remodeling by atomic force microscopy. *Pflugers Arch. - Eur. J. Physiol.*, 456:29–49, 2008.

P. Macklin, M. E. Edgerton, A. Thompson, and V. Cristini. Patient–calibrated agent–based modelling of ductual carcinoma in situ (dcis): From microscopic measurements to macroscopic predictions of clinical progression. *Journal of Theoretical Biology*, 301:122–140, 2012.

M. Matsumoto and T. Nishimura. Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation*, 8:3–30, 1998.

227

P. K. Mattila and P. Lappalainen. Filopodia: molecular arcitecture and cellular functions. *Nature Reviews Molecular Cell Biology*, 9:446–454, 2008.

R. Mayor and C. Carmona-Fontaine. Keeping in touch with contact inhibition of locomotion. *Trends in Cell Biology*, 20:319–328, 2010.

S. McDougall, J. Dallon, J. Sherratt, and P. Maini. Fibroblast migration and collagen deposition during dermal wound healing : mathematical modelling and clinical implications. *Phil. Trans. R. Soc. A*, 364:1385–1405, 2006.

F. A. Meineke, C. S. Potten, and M. Loeffler. cell migration and organization in the intestitial crypt using a lattice-free model. *Cell Prolif.*, 34:253–266, 2001.

G. A. Mirams, A. G. Fletcher, P. K. Maini, and H. M. Byrne. A theoretical investigation of the effect of proliferation and adhesion on monoclonal conversion in the colonic crypt. *J. theor. biol.*, 312:143–156, 2012.

Y. Miyashita and M. Ozawa. Increased internalization of p120-uncoupled e-cadherin and a requirement for a dileucine motif in the cytoplasmic domain for endocytosis of the protein. *J. Biol. Chem.*, 282:11540–11548, 2007.

A. Mogilner. Mathematics of cell motility: have we got its number? *J. Math. Biol.*, 58:105–134, 2007.

A. Mogilner and L. Edelstein-Keshet. Regulation of actin dynamics in rapidly moving cells: a quantitative analyisis. *Biophys. J.*, 83:1237–58, 2002.

Y. Nakaya and G. Sheng. Epithelial to mesenchymal transition during gastrulation: an embryological view. *Development, Growth and Differentiation*, 50: 755–766, 2008.

T. J. Newman. Modeling multicellular systems using subcellular elements. *Mathematical Biosciences and Engineering*, 2:611–622, 2005.

D. Noble. *The music of life: biology beyond the genome.* Oxford University Press, 2006.

I. L. Novak, B. M. Slepchenko, and A. Mogilner. Quantitative analysis of g-actin transport in motile cells. *Biophys. J.*, 95:1627–1638, 2008.

H. G. Othmer, S. Dunbar, and W. Alt. Models of dispersal in biological systems. *J. Math. Biol.*, 26:263–298, 1988.

K. J. Painter. Modelling cell migration strategies in the extracellular matrix. *J. Math. Biol.*, 58:511–543, 2009.

S. P. Palecek, J. C. Loftus, M. H. Ginsberg, D. A. Lauffenburger, and A. F. Horwitz. Integrin-ligand binding properties govern cell migration speed through cell-substratum adhesiveness. *Nature*, 385:537–540, 1997.

L. E. Payne and W. H. Pell. The stokes flow problem for a class of axially symmetric bodies. *Fluid Mechanics*, 7:529–549, 1959.

N. Pecina-Slaus. Tumour-suppessor gene e-cadherin and its role in normal and malignant cells. *Cancer cell International*, 3:17, 2003.

C. S. Peskin, G. M. Odell, and G. F. Oster. Cellular motions and thermal fluctuations: the brownian rachet. *Biophysical Journal*, 65:316–324, 1993.

J. Pitt-Francis, P. Pathmanathan, M. O. Bernabeu, R. Bordas, J. Cooper, A. G. Fletcher, G. R. Mirams, P. Murray, J. M. Osborne, A. Walter, S. J. Chapman, A. Garny, I. M. van Leeuwen, P. K. Maini, B. Rodr?guez, S. L. Waters, J. P. Whiteley, H. M. Byrne, and D. J. Gavaghan. Chaste: A test-driven approach to software development for biological modelling. *Computer Physics Communications*, 180(12):2452 – 2471, 2009.

K. Poole, K. Khairy, J. Friedrichs, C. Franz, D. Cisneros, J. Howard, and D. Mueller. Molecular-scale topographic cues induce the orientation and directional movement of fibroblasts on two-dimensional collagen surfaces. *J. Mol. Biol.*, 349:380–386, 2005.

V. Quaranta. Motiliy cues in the tumor environment. *Differentiation*, 70:590–598, 2002.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. URL `http://www.R-project.org`. ISBN 3-900051-07-0.

I. Ramis-Conde, M. A. J. Chaplain, and A. R. A. Anderson. Mathematical modelling of cancer cell invasion of tissue. *Mathematical and Computer Modelling*, 47:533–545, 2008a.

I. Ramis-Conde, D. Drasdo, A. R. A. Anderson, and M. A. J. Chaplain. Modelling the influence of the e-cadherin-beta-catenin pathway in cancer cell invasion: A multiscale approach. *Biophys. J.*, 95:155–165, 2008b.

I. Ramis-Conde, M. A. J. Chaplain, A. R. A. Anderson, and D. Drasdo. Multiscale modelling of cancer cell intravasation: the role of cadherins in metastasis. *Physical Biology*, 6, 2009.

K. A. Rejniak. A single-cell approach in modeling the dynamics of tumor microregions. *Mathematical Biosciences and Engineering*, 2:643–655, 2005.

K. A. Rejniak. An immersed boundary framework for modelling the growth of individual cells: an application to the early tumour development. *Journal of Theoretical Biology*, 247:186–204, 2007.

K. A. Rejniak and L. J. McCawley. Current trends in mathematical modeling

of tumor–microenvironment interactions: a survey of tools and applications. *Experimental Biology and Medicine*, 235:411–423, 2010.

A. J. Ridley, M. A. Schwartz, K. Burridge, R. A. Firtel, M. H. Ginsberg, G. Borisy, J. T. Parsons, and A. R. Horwitz. Cell migration: integrating signals from front to back. *Science*, 302:1704–1709, 2003.

E. Sahai and C. Madsen. Cancer dissemination–lessons from leukocytes. *Developmental Cell*, 19:13–26, 2010.

S. A. Sanderius, M. Chuai, C. J. Weijer, and T. J. Newman. Correlaing cell behaviour with tissue topology in embryonic epithelia. *Plos One*, 6:e18081, 2011. doi: 10.1371/journal.pone.0018081.

D. K. Schlüter, I. Ramis-Conde, and M. A. J. Chaplain. Computational modeling of single cell migration: the leading roe of extracellular matrix fibers. *Biophys. J.*, 103:1141–1151, 2012.

O. Schmalhofer, S. Brabletz, and T. Brabletz. E-cadherin, β-catenin, and zeb1 in malignant progression of cancer. *Cancer Metastasis Review*, 28:151–166, 2009.

S. W. Schneider, P. Pagel, C. Rotsch, T. Danker, H. Oberleither, M. Radmacher, and A. Schwab. Volume dynamics in migration epithelial cells measured with atomic force microscopy. *Eur J Physiol*, 439:297–303, 2000.

M. Scianna and L. Preziosi. *Cellular Potts Models–Multiscale Extensions and Biological Applications*. Chapman & Hall/CRC, 2013.

C. Sonnenschein and A. Soto. The death of the cancer cell. *Cancer Research*, 71: 4334–4337, 2011.

C. L. Stokes and D. A. Lauffenburger. Migration of individual microvessel endothelial cells: stochastic model and parameter measurement. *J. Cell Sci.*, 99: 419–430, 1991.

B. Tanos and E. Rodriguez-Boulan. The epithelial polarity program: machineries involved and their hijacking by cancer. *Oncogene*, 27:6939–6957, 2008.

J. Tsai and L. Kam. Rigidity-dependent cross talk between integrin and cadherin signalling. *Biophys. J.*, 96:L39–L41, 2009.

S. Turner and J. A. Sherratt. Intercellular adhesion and cancer invasion: a discrete simulation using the extended potts model. *Journal of Theoretical Biology*, 216: 85–100, 2002.

F. van Roy and G. Berx. The cell-cell adhesion molecule e-cadherin. *Cellular and Molecular Life Sciences*, 65:3756–3788, 2008.

M. F. Ware, A. Wells, and D. A. Lauffenburger. Epidermal growth factor alters fibroblast migration speed and directional persistence reciprocally and in a matrix-dependent manner. *J. Cell Sci.*, 111:2423–2432, 1998.

R. A. Weinberg. *The biology of cancer*. Garland Science, 2007.

A. S. T. Wong and B. M. Gumbiner. Adhesion-independent mechanism for suppression of tumour cell invasion by e-cadherin. *Journal of Cell Biology*, 161: 1191–1203, 2003.

J. Yang and R. A. Weinberg. Epithelial-mesenchymal transition: At the crossroads of development and tumor metastasis. *Developmental Cell*, 14:818–829, 2008.

M. H. Zaman, R. D. Kamm, P. Matsudaira, and D. A. Lauffenburger. Computational model for cell migration in three-dimensional matrices. *Biophys. J.*, 89: 1389–1397, 2005.

E. Zamir and B. Geiger. Molecular complexity and dynamics of cell-matrix adhesions. *J. Cell Sci.*, 114:3583–3590, 2001.