Morphological analysis of epithelial cells
Morphologische Analyse von Epithelzellen

Der Naturwissenschaftlichen Fakultät
der Friedrich-Alexander-Universität Erlangen-Nürnberg

zur Erlangung des Doktorgrades Dr. rer. nat.
vorgelegt von Sara Kaliman aus Korčula, Kroatien
Contents

Summary 5

Zusammenfassung 11

Acknowledgements 15

1 Introduction 19
  1.1 General Introduction .................................................. 19
  1.1.1 Epithelial cells .................................................. 20
  1.2 In-vivo models for collective cell behavior ......................... 22
    1.2.1 Zebrafish, (Danio rerio) .................................... 22
    1.2.2 Fruit fly (Drosophila) ....................................... 24
  1.3 In-vitro models for collective cell behavior ....................... 25
    1.3.1 Madin-Darby Canine Kidney (MDCK) cell line .................. 25
    1.3.2 Collective cell motion ....................................... 26
    1.3.3 Modeling collective cell motion ................................ 32
    1.3.4 Response to mechanical stress ................................ 34
    1.3.5 Quantifying various properties of the growing mono-layer .................................................. 37
  1.4 Focus of this thesis .................................................. 37

2 Space tessellation by the epithelial cells 39
  2.1 Limits of applicability of the Voronoi tessellation determined
    by centers of cell nuclei to epithelium morphology ................ 39
    2.1.1 Introduction .................................................. 39
2.1.2 Experimental Materials and Methods .................................................. 42
2.1.3 Image analysis .................................................................................... 43
2.1.4 Properties of Centre of Mass Voronoi Tessellation .......................... 48
2.1.5 Results ................................................................................................. 51

2.2 Nuclei shape based tessellation method reconstructs epithelium morphology more accurately ............................................................ 62
2.2.1 Method .................................................................................................. 62
2.2.2 Results .................................................................................................. 65

2.3 Discussion and Conclusions .................................................................. 69

3 Morphological characterization and mechano-sensitivity of the MDCK epithelium system model ................................. 73
3.1 Introduction ............................................................................................... 73
3.2 Experiments ............................................................................................... 75
3.3 Growth of MDCK II monolayer on hard substrates ......................... 76
  3.3.1 Image analysis .................................................................................... 77
  3.3.2 Results ............................................................................................... 77
3.4 Novel growth regime of MDCK II cells on soft substrates .............. 79
  3.4.1 Results ............................................................................................... 80
3.5 Mechanosensitivity on hard substrates: Characterization of monolayer growth, cell organization, proliferation, steady state and irregularities ......................................................... 84
  3.5.1 Image analysis .................................................................................... 84
  3.5.2 Results ............................................................................................... 88
3.6 Discussion and Conclusions .................................................................. 103

4 Simulating the growth of epithelial cell clusters .............................. 107
4.1 Introduction ............................................................................................... 107
4.2 Cell tracking ............................................................................................. 109
  4.2.1 Tracking algorithm ............................................................................ 109
  4.2.2 Experiment No.1: Cluster with low cell density ......................... 110
  4.2.3 Experiment No.2: Cluster with formed bulk and edge .......... 115
4.3 Cluster growth in time ............................................................................ 118
4.4 Simulation of the circular cell cluster growth .................................... 121
  4.4.1 Formation of the bulk and edge regions in colonies grown from few cells ................................................................. 122
  4.4.2 Algorithm .......................................................................................... 122
  4.4.3 Results ............................................................................................... 127
4.5 Discussion ................................................................................................. 133
4.6 Conclusions .............................................................................................. 135
5 Organization of the epithelial cells and comparison to random systems

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Introduction</td>
<td>137</td>
</tr>
<tr>
<td>5.2 Method</td>
<td>138</td>
</tr>
<tr>
<td>5.2.1 Results</td>
<td>141</td>
</tr>
<tr>
<td>5.3 Comparison with packings of random ellipses</td>
<td>145</td>
</tr>
<tr>
<td>5.3.1 Packing of uniform ellipses</td>
<td>147</td>
</tr>
<tr>
<td>5.3.2 Packing of randomly distributed cell nuclei</td>
<td>149</td>
</tr>
<tr>
<td>5.3.3 Origin of the high correlation between NCR and cell density</td>
<td>159</td>
</tr>
<tr>
<td>5.3.4 Discussion and Conclusions</td>
<td>161</td>
</tr>
</tbody>
</table>

Appendix

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4 Appendix: Image segmentation errors and performance of CMVT method on other substrates</td>
<td>163</td>
</tr>
</tbody>
</table>

Bibliography | 175 |
Summary

The formation of an organism is a complex interplay of various cell types which proliferate and reorganize until they form a fully grown living being. This process is governed by widely studied biochemical and genetic factors. The physical manifestation of these factors constitutes cell morphology, which arises from forces produced either by the cell or by its neighbors. Quantifying typical cell shapes dynamically and over different tissue regions would improve our understanding of normal and abnormal tissue morphologies, which is essential for diagnosis but also it reveals many physical mechanisms underlying tissue growth. The process of tissue growth is a complex interplay of passive changes and active changes. The passive changes originate as a mechanical response of the cell to external forces, while the active changes arise from the energy-consuming acto-myosin complex. Unraveling the differences between active and passive responses extends beyond biology to the realm of biophysics where the field of active matter is extensively studied. Therefore, the focus of this thesis is the morphological characterization of tissues at different length scales in order to identify important physical limitations and rules which govern tissue growth. As a model system, we use unconstrained MDCK-II epithelial monolayers seeded at flat substrate with different rigidities.

The thesis is organized into five chapters, the first of which is a general introduction into in-vivo and in-vitro models for tissue growth, collective cell behavior, and cell mechanical properties. In the second chapter, we analyze how epithelial cells tessellate the space between each other. It has been almost 40 years since it was suggested that epithelial space tessel-
lation is similar to Voronoi tessellation. Relying on this result, numerous recent studies estimate cell morphology with Voronoi tessellation generated from cell nuclei centers of mass. However, the error of this approximation was never quantified. For the first time, we quantify the error of such approximation in respect to various morphological measures of a cell. Furthermore, we develop a novel and systematically more accurate method for reconstructing cell membranes from cell nuclei positions. To do so, we generate a tessellation from cell nuclei boundaries. The success of this tessellation must be found in a tight correlation between nuclei shapes and cell shapes at biologically relevant time scales. Since cell nuclei are more easily stained and segmented than cell membranes, this readily implemented method can be used widely and it builds a foundation for further morphological analysis of tissue. Since the method relies on precise nuclei identification, we develop and describe a novel routine for MDCK-II cell nuclei segmentation.

In the third Chapter we characterize large MDCK-II colonies on glass substrates and polyacrylamide (PA) gels with various rigidity. Mechnosensitivity was repeatedly confirmed for single cells, but once cells adhere to each other, they no longer seem to be sensitive to substrate rigidity. We show that, contrary to conventional wisdom, cells in large colonies are still mechano-sensitive. The evidence for this is found either on very soft substrates or – in the case of harder substrates – on long time scales. On soft underlying substrates, we observe a regime of growth not previously reported. The novel regime of growth starts with archipelago of 3D cell droplets which, upon reaching critical size, give rise to a cell monolayer characterized by very high and constant cell density. This monolayer is surrounded by 3D edge of unorganized cells which still divide. On the contrary, cells on harder substrates exhibit previously studied regime of MDCK-II growth. To further quantify this regime, we focus on morphological characterization of circular 2D cell colonies with initially uniform and low cell density. The colonies grow to large sizes, densify, and compartmentalize into bulk and edge regions. In the bulk region, cells undergo contact inhibition, whilst in the edge region, where cell density gradually decreases, cells move and proliferate. However, the distribution of cell density and cell colony area differs between hard gels and glass substrates. Moreover, steady state is always found at higher cell densities on hard gels than those on glass substrates. In the heart of these differences lies a different rate of cell proliferation. We found that large cells proliferate more on hard gels than on glass substrates. Also, both the nature and number of irregularities – which emerge in a monolayer structure with time – are very different on glass substrates. These in-vitro results indicate multiplex
response of fully differentiated cells to rigidity of their surroundings.

Characterizing growth of unconstrained circular epithelial colonies provides evidence of spontaneous large scale compartmentalization into bulk and edge regions. In the fourth Chapter, we report on our investigation of the necessary conditions for compartmentalization to emerge. Firstly, we measured cell velocity as a function of cell’s radial distance from the center of the cluster. The results suggest that cell velocity linearly depends on cell position in the edge region of the cluster, whilst in the bulk region the value of cell velocity is constant. With our simulation, we successfully reproduce cluster morphology over a period of eight days. This success confirms that measurements of cell velocity, cluster area growth in time and cell proliferation probability are consistent with spontaneous compartmentalization of cell colonies. Moreover, we utilize the simulation as a testing tool for the effect of different cell behaviors on global morphology of cell clusters. Importantly, our experimental measurements expand conventional wisdom on cell collective behavior and proliferation.

During colony growth cells organize with their nearest neighbors and change their shape. Movement and reorganization continue until steady state is reached. In the last Chapter we study density induced changes of cell shapes. We find that most of cell morphological measures change with cell density even though these measures do not depend on cell area. However, cells in steady state have the same morphological properties independent of cell density value. This result implies that steady state is defined by morphological organization and not by cell density as previously believed. Furthermore, we test similarities between organization of epithelial tissue and random packings of ellipses. Since cell nuclei are stiffer than cytoplasm, they occupy more and more of the total cell area as tissues densify. Thus, the ratio of cell nuclei and cell body area represents area packing fraction. This phenomenon enables a direct mapping between cell density and the area fraction of random packings. We therefore compare images of cell monolayers covering the whole range of attainable cell densities- and random packing of mono-disperse and poly-disperse ellipses. The poly-disperse ellipses have the same areas and elongations as the cell nuclei in a given image. Probability distributions reveal unforeseen similarities between random packings of cell nuclei and real tissue, which were the most evident for tissues in steady state. On the other hand, correlation functions affirm that cell reorganization, which accompanies a rise in cell density, is a necessary consequence of increasing area fraction (this arises from relative stiffness of cell nuclei). However, some measures remain unchanged in the tissue. This suggests an active cell mechanism which keeps these measures constant as tissue densifies. However, the
reason why cells want to keep, for example, cell elongation constant is yet to be understood.

We have identified many important principles regulating tissue organization on different length scales. The image analysis tools we develop in this work can be helpful for a broader scientific community – for instance in diagnostics. Our work indicates there are several new biological phenomena not previously considered. First of all, we identify a fast shape-coupling between cell nuclei and membranes. Furthermore, we show that cells in a tissue feel and react to rigidity of their surroundings in a complex manner. The large-scale cell organization seen in our experiments holds a plethora of not fully understood collective and individual cell behavior which we enrich with new information. Additionally, we discover striking similarities between tissue morphological properties and random ellipse packings which enables a new perception of cellular organization. All these phenomena open new perspective on the laws governing tissue formation. However, more studies are needed to deepen the presented results and reveal new laws beneath a fascinating process of tissue growth and organization.
Zusammenfassung


-organisation aufzudecken.
Acknowledgements

I would like to acknowledge all the people who contributed to this thesis.

• First of all, my warm gratitude goes to my supervisor Prof. Dr. Ana-Sunčana Smith for giving me this project, for support, for teaches and for believing in me. Without her, none of this would be possible.

• Furthermore, I would like to express my appreciation to our experimental collaborators, the cell and matrix mechanics group in Göttingen. Prof. Dr. Florian Rehfeldt supervised the project and Dr. Carina Wöllnik and Christina Jayachandran conducted most of the experiments analyzed in this thesis.

• This thesis would not be the same without Damir Vurnek and Jakov Lovrić. Their work was incorporated in the last two chapters of this thesis and I am very grateful to them. Also, many thanks to Philipp Linke for sharing his experimental data with me.

• Moreover, I am thankful to Prof. Dr. Diana Dudziak who shared her facilitates and provided intellectual contribution to this project.

• Prof. Dr. Gerd Schörder-Turk earned a special thanks for all conversations we had at the beginning of this project. It is him who provided us with the essential information on Voronoi tessellations.

• I would like to acknowledge the funding that I have received throughout my research, specially the Cluster of Excellence: Engineering of Advanced Materials (EAM) and the European Research Council
(ERC), but also the Research Training Group 1962/1 (RTG) in Erlangen.

- The proof reading and helpful grammatical corrections were done by Ivana Krce, Dr. Robert Blackwell, Dr. Jayant Pande and Dr. Karmen Ćondić-Jurkić, Dr. Michael Thomas and Damir Vurnek. Simone Gehrer, Dr. Caroline Hack, Dr. Ekaterini Georgiadou and Katharina Blessing translated thesis summary to German language. Thank you!

- I need to acknowledge all my colleges on this project, namely Damir Vurnek, Jakov Lovrić, Dr. Adriana Lepur, Dr. Aliee Maryam, Simone Gehrer and Wolfram Barfuß for contributing to this research with they work and more importantly with all the conversations which we had during and after work.

- Speaking of colleges, I could not forget also: Dr. Jayant Pande, Dr. Zlatko Brkljača, Dr. Zoran Miličević, Mislav Cvitković and Robert Stepić who became my friends.

- My very special appreciation goes to Karmen, Anita and Ekaterini for support and love I needed to finish this thesis.

- Big and bold thank you to my one and only parents.

- I could not finish this list, without expressing my gratitude to everyone who was not mention here, but they have been, nevertheless, the important and inherent part of this process.
1.1 General Introduction

Biophysics is an interdisciplinary field trying to merge physics and biology, two distinct disciplines that have been separately developing for centuries. However, it is less known that many biological discoveries were made by physicists, especially in genetics. For example Gregor Mendel, the founder of modern genetics, was a physicist. Erwin Schrödinger latter in his career got very interested in biology and wrote a book “What is life?” in which he proposed the existence of an aperiodic crystal containing the genetic information stored as a configuration of covalent bonds. Another physicist, Francis Crick, will latter credit this book as an inspiration for his research. It is therefore not a novelty that a physicists try to answer purely biological questions.

The term biophysics was coined by Karl Pearson, in 1892 in his famous book “The Grammar of Science” where he defined the field as: “this branch of science which endeavors to show that the facts of biology constitute particular cases of general physical laws”. He was one of the first people to emphasize the importance of quantification for biology, medicine, and social sciences. Despite his involvement in eugenics progress, Pearson was a brilliant statistician whose work covered almost all the main statistical methods used in biophysics today. Nowadays, biophysics covers all scales and closely relates to computer science, mathematics, and chemistry. In the following pages a short introduction into biophysical re-
search of animal cells is given, with a special focus on in-vitro models of epithelial tissues which are the focus of this thesis.

1.1.1 Epithelial cells

The word epithelium comes from Ancient Greek "epi" (on top) + "thele" (nipple) because originally this word described only the skin on the breasts. The epithelium is perhaps the first tissue that emerged during phylogeny while today, 600 million years later, over 150 types of epithelial cells participate in digestion, reproduction, sensing, respiration, and many other functions in mammals [1]. Epithelial tissue protects us from pathogens and secrete proteins that digest food or feed our young. Furthermore, it is epithelium that absorbs nutrients and detoxifies our bloodstream [2].

Epithelial cells can be classified by their shape (squamous, columnar, cuboidal) or by their layers arrangement. Epithelium arranged in a single layer of cells is called simple while the one consisting of multiple layers is called stratified. If epithelium just appears as multilayered it is classified as pseudostratified. One of the main functions of the epithelial tissue is to be a boundary and protection: as such it lines cavities in the body, as well as organs and blood vessels. In order to provide protection, it is necessary to epithelial cells to form very strong cohesion between each other.

They do so by forming tight junctions, adherens junctions, desmosomes, and gap junctions. Tight junctions (also called occluding junctions in vertebrates) are found only in epithelial cells. The name "occluding" comes from a transmembrane protein occludin which, together with claudin, is the major protein forming those strong junctions. The purpose of tight junctions is to seal the gaps between the cells in epithelia, creating a barrier for the diffusion of molecules across the cell sheet. Also, tight junctions prevent the diffusion of proteins in the plane of the membrane and in that way help to maintain a difference between the populations of proteins in the apical and basolateral membrane domains of the epithelial cell.

Desmosomes are localized spot-like adhesions which provide anchoring sites for intermediate filaments while gap junctions mediate communication between cells by allowing small molecules to pass from cytoplasm to cytoplasm of neighboring cells [3]. Adherens junctions in epithelium are typically formed by trans-dimerization of E-cadherins. This happens in a Ca\(^+\) dependent manner in a continuous belt beneath tight junctions. E-cadherin clusters are stabilized by actin filaments via catenins (β-catenin, α-catenin, vinculin etc.) [3, 4]. Mechanically, size of the adhesion zone
depends on the interplay between adhesion forces and cortical tension \[5\]. While adhesion forces tend to increase cell-cell contacts, actomyosin network tends to reduce them. Yet, adhesion and cortical tension are not independent because they are both supported by actin filaments. Cells have to actively remodel those cell junctions during growth, and this happens either by producing new cell junctions (cell divisions) or by removal of junctions (live cell extrusions and cell death) \[6\].

Another characteristic of the epithelium is the apico-basal polarity which is essential for the vectorial transport, sensory perception and barrier function. Epithelial polarity programme (EPP) controlled by a number of proteins and lipids is responsible for the establishment and maintenance of apical and basolateral domains \[1\]. The machinery of polarized trafficking used by epithelial cells to distribute proteins into apical and basolateral sides has been extensively studied \[7\], \[8\]. Beside different protein expressions, epithelial cells have membrane protrusions on the apical side such as microvilli or cilia which increase the surface area and help the absorption process \[9\]. The failure to generate or maintain cell-surface polarity leads to altered epithelial function and to pathologies such as polycystic kidney disease, hypercholesterolaemia and cancer \[8\].

Epithelial cells can quickly lose epithelial phenotype in a process known
as epithelial–mesenchymal transition (EMT), during which they lose junctions and consequently their apical–basal polarity, reorganize their cytoskeleton, undergo a change in the signalling programme, change cell shape and reprogram gene expression \[10\], \[11\]. All of this increases the motility of individual cells and enables the development of an invasive mesenchymal phenotype. A process very similar to EMT is recognized in dissemination of cancers, most of which (90%) in humans, arise from epithelia \[12\]. Beside cancer progression, epithelial tissue is very important during morphogenesis when cells of the cleavage-stage embryo undergo “epithelialization” and become adherent to one another \[13\].

While biochemical studies of the epithelium focus on the molecular mechanisms behind cell polarization, EMT, cell-cell or intracellular signalling etc., biophysical community sees the epithelium as a perfect system to study collective cell motion, cellular forces and organization. Directional collective motion is the key condition necessary for morphogenesis, wound healing and sometimes cancer progression. Epithelial studies include in-vivo, in-vitro, and in-silico studies of cell growth, trying to uncover both the chemical mechanisms as well as physical forces governing such processes.

1.2 In-vivo models for collective cell behavior

Many different model systems have been studied in this filed, from often complicated in-vivo models that offer more realistic but less controlled environments, to 3D and 2D in-vitro models, the latter being the system addressed in this thesis. Even though 2D in-vitro systems are relatively more convenient for detailed image analysis or well controlled changes of a single parameter in the system there are still many unanswered questions and contradictions in the published data.

Morphogenesis (morphê + genesis) is a process during which an organism develops its shape. In-vivo models try to analyze this process in simple living organisms that have rapid development and sequenced genome, like Drosophila, zebrafish and Caenorhabditis elegans \[15\].

1.2.1 Zebrafish, (Danio rerio)

The zebrafish, Danio rerio, has emerged as a vital model system for investigating the molecular mechanisms that drive the coordinated spreading of tissues in the early embryo. This is due to its external development,
large size, optical clarity, and genetic tractability [16]. The stages of morphogenesis that received the most attention so far are epiboly and gastrulation, which happen almost simultaneously.

During the early embryo development cells move [16], [17]. This phase is known as epiboly. Prior to this phase cells organize into three layers. The first layer is single-cell thick epithelium or the enveloping layer (EVL) which encloses the deep cells (DEL), which themselves eventually give rise to embryonic tissues. The third layer, the yolk syncytial layer (YSL), is an extra-embryonic tissue at the interface between the yolk and deep cells. All three embryonic layers undergo epiboly, thus zebrafish epiboly can serve as a model for studying the spreading of an epithelial sheet (EVL). Deep cells (DEL) will reorganize during gastrulation into ectoderm, mesoderm, and endoderm. These three layers will give rise to neuronal tissues and skin (ectoderm); the skeleton, heart, and connective tissue (mesoderm); and intestinal organs (endoderm) [18].

Beside plentiful biochemical studies of this process characterizing everything from genes to chemical signalling, a vast number of biophysical studies also contributed to the knowledge about morphogenesis. For example, one study has shown that a difference in tissue surface tension influences the spatial positioning of zebrafish germ layer tissues [19].

1.2.2 Fruit fly (Drosophila)

Drosophila (drósos, “dew”, and phílos, “loving” with the Latin feminine suffix -a) is one of the most important system models for the epithelial development. The fertilized egg develops into a larva, and the larva stage takes up to 6 days. Part of larva is imaginal disc which constitutes of proliferating epithelial cells which will eventually form most of the adult cuticular body [20]. Different imaginal discs will form wings, antennae, legs, and eyes [21]. For example, wings are formed from wing discs which grow from about 50 to 50,000 cells in four days [20]. Drosophila wings are two-sided epithelium, with one side being composed of elongated columnar cells, and the other of the flat cells (Figure 1.2 B). Biophysical studies analyze the Drosophila wing formation as a model for understanding, for example, the mechanism that globally orients planar polarity in the epithelium [22], conversion of the epithelial cell shape during morphogenesis [23], or the role of compartment boundaries [24].
1.3 **In-vitro models for collective cell behavior**

Beside in-vivo studies, cell culture has proven to be very useful in determining the main physical factors that affect the behavior of the cells because conditions of the growth can be controlled. As well as primary cells, immortalized cell lines can be used for this purpose.

1.3.1 Madin-Darby Canine Kidney (MDCK) cell line

One of the most established in-vitro system models for epithelium are Madin-Darby Canine Kidney (MDCK) cells (more than 6700 search results on PubMed). Those cells were derived by S.H. Madin and N.B. Darby from a normal kidney of an adult Cocker Spaniel in 1958 [25], and later it was proven that they originated from the kidney distal tube [26]. Even though they presumably originate from normal kidney cells it seems that, at least some strains, have become tumorigenic, which is an example of in-vitro neoplastic development.

As early as in 1969 it was reported that intravenously injected suspensions of MDCK cells produced brain metastases resembling adenocarcinoma in chick embryos [27]. Later study reported that MDCK cell line failed to form tumors in nude mice [28]. Collectively, data published on MDCK cell tumorigenicity suggests that there is variability within the cell line. To test if MDCK cell line has different tumorigenicity, one study compared the tumor-forming capacity and tumor latency of three different lots of MDCK cells obtained from the American Type Culture Collection (ATCC) in athymic nude mice and showed that there indeed exist lot-to-lot variations in the tumorigenicity of MDCK cells in a dose dependent manner [29]. Since MDCK cells are the most widely used cell line in influenza virus research [30] it is of great importance to develop MDCK subclones which are not tumorigenic. One such example is MDCK cell clone 9B9-1E4 which, when injected into adult and new born athymic nude mice, did not lead to progressive tumor formation [30], and was therefore later patented as a substrate for vaccine production.

Due to variability of the original MDCK cell line [31] [32], different strains were derived from it. Today, many different strains are available, including the parental line, MDCK I, MDCK II, MDCK.1, MDCK.2, superdome and supertube [33].

In all of our experiments we used exclusively MDCK II strain, which is the most commonly used strain. Those cells were developed form a higher passage of MDCK cells and they display much lower transepithelial electrical resistance (TER) values demonstrating "leaky" junctions. Further-
more, they do not form gap junctions and they are larger and taller compared to the smaller and flatter type I cells [33].

A great advantage of this epithelial cell line is that it demonstrates clear apico-basolateral polarity in 2D and 3D cell culture, well defined cell junctions, and a rapid growth rate [33]. In three-dimensional soft matrigel, they form a spherical enclosure of a lumen that is enfolded by one layer of polarized cells with an apical membrane exposed to the lumen side [34]. These structures can be altered by introducing the hepatocyte growth factor (HGF), which induces the formation of linear tubes [35]. Those 3D models are more often used in biochemical studies as a system model for studying formation of epithelial polarity [36].

From a biophysical point of view, the best studied regime of growth is the one on 2D surfaces where MDCK-II cells form sheets and exhibit contact inhibition [37]. This regime was analyzed in the context of collective cell motion [38],[39],[40] mechanical properties of the monolayer [41], morphological properties of the monolayer and quantification of various cell properties such as cell area and volume [42], cell proliferation [43], or colony growth in time [37].

### 1.3.2 Collective cell motion

Collective motion of cells is necessary whenever growth takes place in the life of multi-cellular organisms [44], whether in healthy tissue (morphogenesis or tissue repair) or in cancer tissue (invasive cancers). It has therefore been studied intensively in the last decade. Collective cell motion is studied in-vivo in 3D systems e.g. formations of zebrafish lateral line primordium [45], or Drosophila tracheal development [46]. However, in-vivo imaging is often complicated to both perform and analyze.

This led to 2D in-vitro models emerging as a very popular choice for biophysical studies of collective cell motion. The MDCK-II cell line is one of the most popular choices in biophysical studies of collective cell motion. Another popular choice are endothelial and fibroblast cells. Fibroblast cells and other non-cohesive cell lines are usually used in addition to epithelial cell lines in order to emphasize the role of cell cohesion in collective cell behavior.

**Biochemical point of view** The key question is which physical and biochemical factors are necessary for the observed collective cell phenomena. Is the tissue growth driven by the cell proliferation pressure alone, or is it driven by migration forces generated by the individual cells and if so,
is the cell-cell cooperation necessary for the overall tissue growth? One of the first modern studies that deal with this question [47] used live Hoechst stain on human umbilical vein endothelial cells (HUVEC) in the standard wound healing experiment and tracked the cells. This study used siRNA against 100 known migration regulators and monitored changes of the fibroblast growth factor (FGF) induced cell migration. Parameters observed in this study were sheet migration, cell density, single cell velocity and cell–cell coordination. The migration regulators mainly affected sheet migration rate directly with only some doing so indirectly by affecting cell proliferation, which in return affected cell migration. This means that besides affecting the cell proliferation, migration rate could also be increased or decreased by silencing different migration regulators in the cell sheet with the same cell density. Those migration regulators could be affecting individual cell velocities or cell-cell coordination. Furthermore, this study showed that growth factors influence the process of leader cells formation, while nearby cells follow through growth factor independent, coordinated migration. Therefore, endothelial sheet migration couples growth factor triggered formation of leader cells at the sheet boundary with growth factor independent follower behavior, the latter being a combination of cell motility per se (partially arising from cell proliferation) and cell–cell coordination.

It has been previously suggested that Merlin protein is a contact sensor important for formation of adherent junctions [48]. However, a recent study pointed out that this protein is also crucial for collective cell migration in MDCK monolayers [49]. The authors report the strongest effect of Merlin knock down on collective cell migration. In stationary monolayers Merlin was found to be localized at the tight junction (TJ) complex; however, it relocalizes to cytoplasm during migration which happens simultaneously with activation of Rac-1. Merlin-deficient cells show increased motility but they do not coordinate their movement with the neighboring cells. This study proposed a biological mechanism which links mechanical forces at cell-cell junctions to the resultant cell movements.

**Biophysical point of view**  Physicists mostly study cell velocities and their correlations. Data from the experiments is analyzed using many different theoretical approaches which try to understand physical laws governing this process. However, different experimental conditions induce and emphasize different phenomena in collective cell behavior. It is therefore important to mention these conditions when talking about the results of experiments concerning collective cell behavior.
**Stripe shaped PDMS stencils**  Modern wound healing experiments are one of the most popular setups for collective cell behavior experiments. In 2006, Nikolic et al. [38] have shown, using MDCK cells line, that ERK1/2 MAPK activation - previously identified as the essential trigger for collective motion of the cells [50] - could be initiated even without a cell injury. This work introduced utilization of the PDMS stencils which are removed once the cells have reached confluence.

Subsequent work of Silberzan’s group [39] has shown that finger-like protrusions of cells appear in those experiments and are always dragged by a leader cell [1], thereby proposing that leader cells are essential for collective motion of the cells. In a later study the same group measured the velocity of MDCK and fibroblast-like normal rat kidney (NRK) cell lines for 300 µm and 1 mm wide stencils [55]. This was also one of the first studies to utilize PIV method in the analysis of collective cell behavior. The average cell velocity decreased over time because the density of the cell monolayer trapped in the stencil increased. Once the stencil was released, cell velocity started to increase and density dropped. For epithelial MDCK cell line, spatial correlations of the velocity fields were much higher than for fibroblast-like NRK cells. Typical spatial correlation length (calculated for the fluctuations in the velocity fields) was estimated to be 200 µm which corresponds to about 15 cells.

Interestingly, another study [56] conducted in similar experimental conditions with 300 µm wide stencil did not report finger-like protrusions but a mechanical wave forming during MDCK expansion. This wave appeared as fluctuations in cell velocities, strain and stress rates, cell area and traction forces. Furthermore, a minimal model assuming cell reinforcement and fluidization was proposed to explain this phenomena.

**Micro-channels**  Cell collective migration was also quantified along the length of the stripe shaped micro-channels. Collective cell behavior dependency on the width of a stripe was analyzed with stripes varying from 20 to 400 µm [57]. In this experiment MDCK cells were allowed to

---

[1] The term “leader cell” was introduced in 2003 in the work of Omelchenko et al. [51]. Their study has shown that in IAR-2 cell line leader cells undergo strong reorganization of the cytoskeleton with microtubules penetrating the lamella, and that in those cells RhoA is inhibited. However, the existence of such cells had been reported in literature before, although not termed as “leader” cells. For example, back in 1966 very similar behavior was reported in chick embryo epithelial cells [52], and in 2000 on MDCK cells [53]. Recently, it was demonstrated on MDCK cells that killing the leader cell by micro-manipulator leads to suppression of collective motion until new leader cell is initiated [54].
migrate along the fibronectin stripe pattern surrounded by non-adhesive surface. It was found that velocity of the leading edge has a constant value over time and depends only on the width of the channel. One should point out that this conclusion was obtained from experiments done on relatively short time scales (up to 15 h). In narrow channels the leading edge moves relatively fast (about 39 µm/h), cells are more elongated and have high velocities even in the direction opposite of that of the edge movement. In wide channels the leading edge moves with constant velocity (about 22 µm/h) and cell velocity drops with distance from the leading edge. Furthermore, swirl-like cell movement is observed only in the wide channels. Spatial velocity correlation length was measured along both the length and the width of the channel. Correlation length along the length was estimated to be 120-200 µm. On the other hand, the perpendicular correlation length depended on the channel width for the channels thinner than 100 µm.

Another study used PDMS stamps to construct channels varying from 100 to 300 µm. Cell movements were analyzed alongside those channels with PIV and direct cell tracking (MDCK cell nuclei were stably transfected) [58]. This study confirmed decrease of the cell velocity with the distance from the leading front, accompanied with simultaneous increase in the cell density. By successive spatial coarse-graining, characteristic length scale was estimated to be 80 µm, while time averaging estimated the characteristic time scale to be 1.1 h. Those numbers were interpreted as correlation length and correlation time of the fluctuations in the cell migration. Furthermore, decoupling between diffusive and directed motion was investigated by Fick’s first law of diffusion with a constant flux term corresponding to the cell velocity of 9 µm/h. Above the critical cell density of 2,200 cells/mm², the density gradient correlates with the flux in a density dependent manner and linear fitting reveals collective diffusion coefficient of about 1500 µm²/h.

**Small circular confinements** Small circular confinements are another popular choice of geometry imposed on a cell monolayer. In one of the studies, [59] low density MDCK monolayers were seeded onto fibronectin coated small circular confinements with the diameter ranging from 50 to 1000 µm. Following the seeding, cells moved randomly until they reached density of 2000 cells/mm². Consequently, cells displayed collective rotation typical for this type of geometry. However, collective rotation was observed only for patterns with diameter of 100 µm and 200 µm. The reason lies in the maximal spatial correlation length which was measured.
to be 170 µm, and therefore only patterns with a radius smaller than this length could induce collective cell rotation. Similar value of the correlation length was also measured in an unconfined cell monolayer [55], [60] pointing to the fact that this value might be the “natural” correlation length for MDCK cells. Average velocity of the synchronized rotation was 25 µm/h for cell densities up to 4,000 cells/mm², whilst for higher densities it decreased significantly. Even in this confined system cell velocities increased with radial distance. To test the effect of the cell-cell cohesion in the collective motion, E-cadherins were suppressed with overexpressing Snail-1 in MDCK cells. This suppressed collective rotation and made it transient. Similar behavior was observed for cells from benign breast tumors, whilst malignant breast tumors exhibited completely random and chaotic movements.

A year later another study confirmed collective rotation in circular confinements smaller than 200 µm in diameter [61]. In this study radial and orthogonal velocity were averaged and observed over time. Fourier analysis revealed the time periodicity of typically 7 h for radial velocity $\langle U_r \rangle_0(t)$ but didn’t confirm the same for orthogonal velocity, $\langle U_θ \rangle_0(t)$ even though fluctuations were observed. Hepatocyte growth factor (HGF) and Blebbistatin affected cell oscillations without affecting the velocity of individual cells.

**Unconfined circular cell clusters** Unconstrained circularly seeded cell clusters are another popular experimental system for studying collective cell migration. The cluster usually contains small number of cells (typically 5,000) which form a monolayer in 24 hours. Once the monolayer is formed, cell cluster expands radially.

Using such model system it was proven that each cell, and not only leader cells, produces force [40]. This was evidenced by traction produced by cells many rows behind the leading edge. However, average normal traction did slowly decayed with distance from the edge. In the case of entirely self-propelled cells, radial stresses should have been precisely zero everywhere, however this was not the case. Radial stresses and cell density increase towards the center of the cell cluster and moreover the ratio between them was always constant. The authors propose that collective motion of a cell monolayer arises from global tug-of-war which integrates local force generation into a global state of tensile stress. Two years later it was shown that local cellular motion follows local orientation of the maximal principal stress even when local cell geometry displayed no preferred orientation [62]. Both stresses and cell velocities are cooperative and het-
erogeneous. Therefore, this kind of cell motion is termed “plithotaxis” (from the Greek “plithos” denoting crowd, plenty, abundance) since it is an emergent property of a multicellular collective system.

In unconstrained low density circular MDCK-II monolayer, swirls were reported [60]. Those swirls formed on both polyacrylamide (PA) gels and glass substrates, but their characteristic size had a different density dependence. This study was the first to show that deformability of the substrate influences the size of a cell group that moves together. Substrate deformation field was generated from the displacement of the beads used in traction force microscopy (TFM), whilst the velocity fields were obtained with PIV analysis from the bright field images. In both cases the average value of the field was subtracted and only field fluctuations were quantified by correlation function. Characteristic length of the spatial auto-correlation function—which was calculated from the fluctuations of substrate displacement vectors—increased with cell density. Furthermore, fluctuation of the deformation vectors correlated in time with velocity fluctuation fields.

Spatial velocity correlations proportional to the root-mean-square velocity imply that effective friction is dominated by the cell-cell interactions, while if dependence is $\xi_{vv} \propto v_{rms}^a$ with $a < -1$ the friction is dominated by cell-substrate interactions. MDCK cell line has shown increase in the correlation length with the rms velocity pointing that cell-cell interactions are the dominant friction term [63]. Fibroblast cells (3T3), on the other hand, have shown decrease of the correlation length with rms cell velocity, while immortalized human bronchial epithelial cells (HBEC) showed two phase behavior: for low rms velocities cells increase correlation length (cell-cell cohesions dominate the friction) while for high velocities they decrease correlation length of the spatial velocity-velocity correlation function (fluid-like behavior). The correlation length for MDCK cells reported in this study was much lower than in other reports [59, 55]. For low cell velocities the correlation length was about 30 $\mu$m while for higher velocities (about 15 $\mu$m/h) it climbed to 50 $\mu$m.

Some studies pointed to an analogy between active cell systems and passive colloidal systems approaching glass transition. Glass-like dynamics of collective cell motion was reported for MDCK monolayers growing from a drop with 5,000 cells seeded on relatively soft (shear modulus of 0.42 kPa) PA gels [64]. This study confirmed that cells slow down with rise in cell density and that the size of a group of cells that move together increases with cell density on deformable substrates. For cell densities below the critical value, cells flow like a fluid while above the critical density relaxations happen at time scales much larger than those of cell division and cell behavior is reminiscent of a glass. This was proven by calculat-
ing the structural factor from the intensity of the bright field images at various time and length scales. The results were described using damped harmonic oscillator (DHO) model. The first term in this model is describing self-diffusivity. However, another term is needed to describe data from cell monolayer and this term, called Brillouin peak, reflects elastic response to density fluctuations. From the width of the first peak one can extract diffusion coefficient, which was not sufficient to describe the cell velocities which were larger than the diffusion velocities for all measured cell densities. However, extrapolating to larger cell densities pointed to the existence of cross-over cell density above which individual cell motion is predominately diffusive. The critical cell density was estimated to be 2,800 cells/mm$^2$. This transition is reminiscent of glass transition controlled by density in colloidal fluids. In addition to structural factor, vibrational density of states (DOS) was also investigated. DOS has shown two set of peaks. The first one, commonly known as boson peak, happens at the time scales of 1 to 1.6 h and length scales comparable to cell size. Since cell migration was much smaller than length of a cell body per hour, this peak was associated with cell body shape fluctuations. Second peak found at time scales of 0.4 to 1 h and length scales smaller than cell size was associated with cell divisions.

A very interesting recent study stepped out of the conventional research conditions and studied collective migration of the a MDCK II monolayer guided by electric fields (corresponding electric current density of 0.3 mA/mm$^2$ was used) [65]. This study showed that cells follow the direction of the electric filed, which pointed to the fact that besides extensively studied chemotaxis, durotaxis and plithotaxis, yet another kind of taxis was present in cells. All the cells except the leader cells aligned their motion with the electric field, even when alternating currents were applied (minimal period of 2.5 min). These lines of studies are based on the Emil du Bois-Reymond’s 1843 discovery of the “wound field”, later investigated by Robert O. Becker mostly on salamanders [66].

1.3.3 Modeling collective cell motion

Various models of collective cell behavior are proposed. Most of the models treat cells like individual objects which are moving and dividing. While some models, e.g. vertex models, are interested in the final morphology of the monolayer, others, such as self-propelled particle models, are interested in cell movement and treat cells simply like individual particles. Another group of models are continuous models which treat tissue
as a fluid and propose that various partial differential equations describe properties of the tissue, e.g. cell density.

Vertex models that consider both cell morphology and motility have been proposed. One example is the two-dimensional cellular Potts model (CPM) which was originally proposed for cell sorting [67] but later studies added cell motility terms and studied collective cell motion within this framework. CPM works on the lattice and individual cells are subsets of that lattice. Favorable shape of the cell is determined by the Hamiltonian-like expression consisting of the cell-adhesion term and cell volume constraint. In a study of collective motion of endothelial monolayers [68], classical CPM was expanded with self-propelled motion of the cells and positive feedback between cell displacements and cell polarity. This model was able to repeat some of the observed experimental results on endothelial cells; for example, that 5-20 cells are moving together in narrow groups. Another study added the cell migration term and studied the effects of cohesion energy and motile force on the observed collective cell phenomena such as velocity correlations and mean-square displacement [69].

Another example is a study of tissue fluidization based on the self-propelled Voronoi model (SVP) [70] where cells have self-propelled motility defined by polarity vector and in addition they feel the interaction force $F_i = -\nabla E$ and the mechanical energy functional has quadratic terms in area and perimeter.

In modeling collective cell motion self-propelled particle models emerged as a common choice. One of the first proposed models [71] was inspired by ferromagnetic interactions. This model assumed that particles in the following time step align their velocity with the surrounding particles with an additional noise term. If particle density is sufficiently high and the value of the noise term is low, collective motion of particles in a random direction will emerge spontaneously. Most of the models that followed were trying to repeat some of the experimentally observed phenomena with flocking-based mechanism of collective cell motion.

In one of such models cells are represented by two particles which interact by repulsive force and once the distance between them exceeds a certain threshold value, the cell divides with a given rate. The equation of cell motion is driven by motility force of a cell, intracellular repulsive and friction forces plus the forces that result from interaction with other cells in a given radius. Those forces account for volume extrusion and adhesion, friction between cells and the noise term [72]. This model qualitatively reproduced the long-range alignment of cell motility forces and traction force maps in the circular epithelial colonies measured by Trapat et al [40]. Furthermore, the result of this model was an increase of the swirl size with
cell density and importance of the adhesion term in observing finger-like protrusions at the cell border. Similar model in which every cell is represented by two particles which interact with each other and with surrounding particles was proposed by Zimmermann et al. In this model cell movement is defined by slightly different functional consisting of the motility force of a cell and intracellular attractive force, whilst interactions between cells are described by repulsive/attractive term. Cells tend to move away from each other by adjusting magnitude and direction of their self-propulsion force due to contact inhibition of locomotion (CIL).

1.3.4 Response to mechanical stress

Whether it is intestinal epithelia, or lung alveoli or embryonic epithelial cells regularly experience tension. It is therefore of great importance to understand the effect of external mechanical stress on epithelial cells and their mechanical properties. Moreover, during Drosophila morphogenesis gene expression changes with external forces. Likewise, in tumors the gene expression is defined by stress distribution. Therefore, there is a general interest in the biophysical community in how external stress influences tissues. Physically speaking, cells react to external forces as viscoelastic materials. When force is applied on short time scales cells react to it as elastic materials, while on the time scales exceeding the relaxation times - typically tens of seconds up to several minutes - cells behave like liquids. More precisely, one could define the creep function \( J(t) \) for a material that is at a time \( t = 0 \) loaded with constant force \( F \) and where deformation \( d \) is measured as a function of time \( t \): \( J(t) = d(t)/F = j_0 \cdot (t/\tau_0)^\beta \). Here \( j_0 \) is softness of the material, \( \tau_0 \) the arbitrary time scale and \( \beta \) is the time scaling invariant exponent. If \( \beta \) is equal to zero the formula represents a response of the elastic material to the force, and if \( \beta=1 \) the equation reads Newton's law of viscous deformation \( d(t)/F = j_0 \cdot t \). Viscous materials are unable to elastically store energy, which is thermally dissipated, and upon removal of the force the deformation persists. Viscoelastic materials as well as cells are both elastically storing and dissipating the energy. Typical value of \( \beta \) in cells is between 0.1 and 0.5.

On the other hand, cells are living and they constantly reorganize their cytoskeleton network at a great expense of the metabolic energy which makes them fundamentally different from nonliving materials. There are many techniques available for measuring cells mechanics, namely: magnetic tweezers, laser tweezers, atomic force microscopy, cell poking,
Figure 1.3: Viscoelastic behavior of the cells. Blue curve represents the response of a viscous material (PDMS silicone oil); purple curve of an elastic material (polyacrylamide-bis-acrylamide (PAA) gel); and red of viscoelastic material (F9 embryonic carcinoma) measured with a magnetic bead. Picture taken from [76]

microplates, and cell stretchers. Each of these methods has its limitations and its advantages. Atomic force microscopy (AFM), as one of the standard methods for measuring the elasticity of cells, has low scanning speed but high space and time resolution as well as range of applied forces (up to 100 nN) [76]. AFM measurements on the MDCK cells showed spatial (both within one cell and between neighboring cells) and temporal changes in mechanical properties of living cells [78]. However, it was found that the tip plays an important role in AFM rigidity measurements. For example, MDCK cells probed with a spherical tip have elasticity of 380 ± 107 Pa while pyramidal tips showed a two-fold increase in elasticities [79]. Moreover, the topology of the underlying substrate influences the tissue elasticity. For example, MDCK-II cells on non-porous substrates are stiffer in comparison to cells on substrates with pores of various sizes [80]. Also, large size pores induced more of the fluid-like behavior in MDCK cells. Beside porosity, substrate rigidity can also alter the mechanical properties of the cells [81].

Unlike other cell types (for example fibroblasts), in the case of MDCK-II cells membrane tension—mainly originating from actin binding sites—is a significant contributor to the overall tension ($T_o = 0.1mNm^{-1}$ out of $T_o = 0.3mNm^{-1}$) [82]. Those measurements were done with a combination of
AFM-indentation and membrane–tether pulling approach.

Still this knowledge is not sufficient to predict the behavior of the cells when they are stretched. If one were to apply knowledge from soft matter physics to a cell system, one would expect to observe fluidization of the cells subjected to the shear [83]. The theory of the soft glassy materials (SGM) sees those materials to be composed of discrete elements having different energies. Therefore, each element of a SGM can be linked to a particle moving in a landscape of quadratic potential wells of a certain depth. If the cells belong to a class of soft glassy materials, mechanical stretching promotes more frequent hopping between local energy minimums, and therefore fluidization of the system [76]. However, results in this field are contradictory: early studies show softening of the cells under shear, whereas later discovery of polymeric cytoskeleton is suggesting stiffening [84]. This debate was addressed in a study that was able to isotropically and biaxially stretch and unstretch cells for 4 seconds and monitor, on the nanoscale, their mechanical properties and remodeling dynamics. The results showed that epithelial cells (including MDCK cell line) fluidize immediately after the stretch, and slowly recover with time [84]. Therefore, cell reaction to shear and normal stress seems to be fundamentally different.

Still, tissues don’t necessarily need to respond to stresses like the single cells so. On a monolayer level, cells adhere to and pull on each other through myosin contractility, which leads to higher tension. On longer time scales tissue is remodeled by cell divisions and apoptosis which introduce local stress into the system. This source stress has isotropic component from divisions and apoptosis but also an anisotropic component because cells have a preferred axis of the division. This source stress together with cell number balance implies viscoelastic behavior of the tissue with the relaxation time connected to the rates of cell divisions and apoptosis [85]. Therefore the tissue can be seen as an elastic material on short time scales while the liquid-like behavior emerges from dynamic reorganization of the tissue that is either internal (divisions and apoptosis) or external (stretching).

The first detailed study of the mechanical properties and the stretching response of MDCK monolayer was done with cells cultured on a temporary substrate created by polymerizing a drop of collagen between the two rods. Similarly to the results in single cells [84], in monolayer high cyclical strain led to partial fluidization of the cells but the biological mechanism remains unknown [41]. Monolayer stiffening in lateral direction was 20 ± 2 kPa and average strain at failure was as high as 69 ± 14 %. When actin cytoskeleton was depolymerized the stiffness decreased by 50%.
1.3.5 Quantifying various properties of the growing monolayer

Another important topic of biophysics is the quantification of various morphological and other cell properties. Such analysis is often nontrivial for single cells and on the level of the cell agglomerates it becomes even more complex. This could be one of the reasons why only a limited number of studies have focused on this problem.

Specifically for MDCK monolayers, one of the most important studies in this filed was done by Puliafito et al. [37] who quantified distributions of MDCK cell areas, colony area, relative size of the mother and daughter cells, duration of cell division, mean cell elongation and number of neighbors at various time points of the cluster development. Furthermore, in one of the pivotal studies in this filed, beside traction forces, distribution of the cell density within the MDCK colony [40] was reported. Zehnder et al. [42] measured the fluctuation of the cell volume in the monolayer while the cell proliferation was measured by Streichan et al. [43].

In general, properties of the MDCK monolayers are rarely quantified for high cell densities since the image analysis, especially from bright field images, becomes very difficult. Therefore, in order to improve the available data in the field, the appropriate image analysis routines need to be developed and meaningful statistical methods applied for each specific problem. Moreover, while the organization problems have been studied for particle packings, foams and emulsions, only a few studies applied the same principles to the cell packing [86], [87].

1.4 Focus of this thesis

Overall, epithelial morphology is insufficiently studied and many cell properties remain to be quantified. In this thesis we focus on characterization of cell shape and cell organization over different length scales (10^{-5} – 10^{-3} m). Moreover, we investigate various cell properties such as proliferation probability, coupling between nuclei and membrane shape, cell perimeter and elongation as a function of cell area etc. We use large circular MDCK-II colonies as an epithelial system model. The organization of the cells is studied for colonies seeded on polyacrylamide (PA) gels with various elasticity and glass substrates. All substrates are coated with collagen-I and chemically treated in the same way, meaning that the only thing that the cells "feel" is the rigidity of the substrate. We show that large scale organization of the cells is strongly influenced by substrate rigidity,
while the micro scale organization is influenced predominantly by mean cell density.

One of the most important novelties is proving the mechano-sensitivity of a cell monolayer and discovering the fast coupling between the shapes of the cell nuclei and the cell membrane. This property is utilized in the novel tessellation method for the determination of epithelial cell morphology from the shape of the cell nuclei. Moreover, we show that during the densification process, which eventually ends at the steady state density, the cell organization changes. We also show that the changes are passive mechanism arising from increased stresses on the cells which have relatively high stiffness of the nuclei. On the other hand, the lack of reorganization for some measures such as cell elongation must arise from an active mechanism. Furthermore, we quantify the cell migration in our system and simulate compartmentalization of the large epithelial clusters into contact-inhibited bulk region and low density edge region.
2.1 Limits of applicability of the Voronoi tessellation determined by centers of cell nuclei to epithelium morphology

The following section was already published as open access article in Frontiers in Physiology (Volume: 7, Article: 511) in 2016 with the title: “Limits of Applicability of the Voronoi Tessellation Determined by Centers of Cell Nuclei to Epithelium Morphology”.

2.1.1 Introduction

Over the last decade, a global effort to understand the underlying principles of morphogenesis, wound healing and cancer progression has generated a tremendous momentum in studies of epithelial tissues [88]. Consequently, significant work to characterize their architecture and growth is being performed on in vivo and in vitro model systems, a prototypical example of the latter being the MDCK cell monolayers [41], [40], [37], [49], [89], [42], [64], [61], [43]. Such progress is founded on the remarkable advance of molecular biology and imaging techniques, whose output data forms the basis for the quantitative analysis of the tissue development [90]. However, optimally harvesting this data depends on the development of image analysis tools. One commonly used technique for gaining informa-
tion about the internal tissue organization is based on the construction of appropriate space tessellations.

For epithelial cells, it was suggested already in 1978 that the polygonal Voronoi tessellation (VT) well approximates the tissue structure [91]. This prompted the development of in silico models, which adopt the polygonal nature of cells and are parametrized to reproduce the distributions of morphological features such as the area and the perimeters of the cells. These models typically use a free energy functional, which is minimized to yield optimal positions of points [92], [93], generating the tessellation. Alternatively, vertex models optimize the cell area and the boundary-length between cells. The parameters of the free energy function yield insights into the mechanical state of the tissue [23], [94], [95], even though the one-to-one correspondence with cells in acquired images cannot be established.

Besides modeling, VT are often applied in direct analysis of fluorescence microscopy data. Tessellations offer simple, fast and fully automated access to tissue morphology, which is otherwise difficult to obtain for a large number of cells. Tessellations are frequently generated from the centers of mass of cell nuclei (CMVT), which themselves are determined from segmented images (Figure 2.1 A-C). Today, CMVT make an integral part of automated image analysis packages used, for example, to delineate cancerous and healthy tissue in histopathological samples. One of the first attempts to use CMVT in a clinically relevant situation was to estimate cell areas and perimeters in primary lung carcinoma [96]. More accurate and complex procedures developed over time include one using CMVT to characterize a number of morphological measures of cell shapes in different cancers with poor and good prognosis [97]. Recently, CMVT became the foundation of an automatic analysis routine and is today used for the analysis of biopsies to distinguish cervical inter-epithelial neoplasia from normal tissues [98], [99].

For a long time CMVT was occasionally used in studies of reconstituted tissues as part of the effort to elucidate biochemical and physical principles of tissue growth [88]. It was applied in an in-vitro characterization of the effect of cell medium on the growth of two colorectal cancer cell lines [100]. With the advance of imaging, CMVT was employed more frequently, due to its simplicity and accessibility of nuclei staining procedures. In recent years it was used to estimate areas of MDCK cells within a monolayer, and to understand fluctuations of the cell volume [42]. Furthermore, it was applied in the quantification of the cell proliferation rate as a function of the cell area [43], and in evaluating the time dependence of the average cell elongation in MDCK colonies [37].

This wide spread usage of CMVT demonstrates that this technique is
becoming an accepted and fairly common method for reconstruction of the cells’ shapes. It is therefore surprising that a procedure for quantitative and systematic analysis of this approximation is still lacking in literature. While visual comparison of the tessellation with the images of the cell membrane suggests that CMVT can capture a number of cell shape characteristics at least in some tissues, direct correspondence of reconstructed shapes and the true cell morphology was not validated in a quantitative manner on a statistically sound sample of any system. Actually, the accuracy of CMVT may vary in different tissues and therefore the applicability of the tessellation should be tested whenever CMVT is used, particularly for diagnostic purposes.

In this paper we develop a protocol for the analysis of the accuracy of CMVT and apply this procedure to MDCK monolayers. We show, on a sample of 15,000 MDCK-II cells (cell areas 74-274 \(\mu m^2\)), generated by imaging three day old model-tissues grown on collagen I coated elastic polyacrylamide gels (E=11-34 kPa), that CMVT indeed reasonably captures the cells’ shapes. Similar results are obtained for different growth conditions (cells grown on substrates with a Young’s modulus of E=0.6 kPa and glass), but with slightly lower statistics. However, due to its intrinsic polygonal nature, CMVT cannot reproduce the curved cell boundaries or avoid cutting through the nucleus interior (Figure 2.1C). Consequently, while instructive, the correspondence of the CMVT and the original data cannot be significantly improved with better imaging.

To quantify the accuracy of the CMVT we analyze a number of classic shape measures, namely cell area, perimeter, and number of neighbors of each cell. Furthermore, we investigate cell anisotropy measures; cell elongation (ratio of major and minor principle axis of cells), variations in
boundary-lengths (deviation of the mean boundary-length that a cell has with each neighbor), and co-alignment between cell body and its nucleus (the angle between major axes of cell and nucleus). To assess the quality of CMVT, we also determine morphological measures directly from fluorescent images of the plasma membrane immuno-stained for β-catenin.

### 2.1.2 Experimental Materials and Methods

**Tissue Culture, Gel Preparation, and Fluorescent Staining**

MDCK-II cells were purchased from ECACC (# 00062107) and cultured in MEM Earle’s medium (Biochrom, #F0325) supplemented with 5% fetal bovine serum (FBS, # F0804, Sigma-Aldrich), 2 mM L-glutamine (Sigma-Aldrich, # G7513), 1% P/S (penicillin, streptomycin) (Gibco, LifeTechnologies, 15070-063) at 37 and 5% CO₂. Cells were trypsinized and passaged every 2 or 3 days before reaching 80% confluence.

Elastic poly-acrylamide (PA) gels (E = 0.6–34 kPa) were prepared as described earlier [89]. In brief: Mixtures of acrylamide (40% solution, BioRad) and bis-acrylamide (2% solution, BioRad) were polymerized for 60 min on plasma cleaned glass cover slips (No. 1, 25 mm ø, VWR) that were functionalized with 3-aminoproyltriethoxysiliane (APTES, Sigma-Aldrich, Germany) for 15 min and incubated with a 0.5% solution of glutaraldehyde in PBS (Sigma Aldrich, Germany) for 30 min. The hydrogels were subsequently coated with Collagen-I (BD Biosciences) at 0.02 mg/mL in a 50 mM HEPES buffer using the bi-functional cross-linker Sulfo-SANPAH (Pierce, Thermo Scientific). For quality control, the Young’s modulus E was measured macroscopically by a cone and plate rheometer (MCR 501, Anton Paar, Austria).

Cells were fixed using a 10% solution of formaldehyde for 5 min, then washed and blocked with a 2% BSA (Sigma Aldrich) solution in PBS, and permeabilized for using a 0.5% solution of Triton X 100 (Carl Roth, Germany). Filamentous actin was stained using phalloidin–tetramethylrhodamine B isothiocyanate (TRITC, # 77418-100UG, Sigma-Aldrich), β-catenin with a combination of primary (anti-beta-catenin AB produced in rabbit, Sigma-Aldrich) and secondary antibody (anti rabbit IgG-FITC, Sigma-Aldrich), and the DNA in the nucleus was labeled with Hoechst (# 33342, Molecular Probes, Life Technologies).

---

1This experiment and image acquisition were done by Christina Jayachandran under supervision of Prof. Dr. Florian Rehfeldt in the 3rd Institute of Physics - Biophysics, Georg-August-University, Göttingen.
Following this procedure, we obtain tissues of MDCK cells grown on stiff gels in the density range of about 4,800-8,600 cells/mm². On glass we explore densities of 4,500-5,700 cells/mm² while on very soft substrates (E=0.6 kPa) we systematically recover densities of 12,500-13,000 cells/mm².

Image Acquisition

Fluorescence microscopy was done with an inverted microscope (Axio Observer.Z1) using a 20x objective (N-Acroplan) and the HXP 120 illumination source (all from Zeiss, Göttingen). Images were recorded with a Zeiss AxioCam (MRm Rev. 3 FireWire) using the Zeiss AxioVision software package (Rel. 4.7). For the different fluorescence channels, the following acquisition times were used: Hoechst (nucleus) 400 ms, TRITC (actin) 200 ms, and FITC (β-catenin) 500 ms. All images were saved in an uncompressed TIFF format in a resolution of 1388 (H) x 1040 (V) = 1.4 Mega Pixel with 8 bit depth resulting in a pixel length 0.31 µm. Stitching of images is not performed to avoid introducing small shifts, which could affect the evaluation of the morphological properties of cells extracted from such images.

Confocal imaging was performed to provide a deeper understanding of the origins of errors associated with CMVT. Images were acquired with a Leica LSM 5 laser-scanning microscope equipped with a white light laser and a 63x oil immersion objective.

2.1.3 Image analysis

Segmentation of the cell nuclei images

Voronoi tessellations are typically obtained from immuno-fluorescent images of cell nuclei. These images need to be segmented with great accuracy to correctly determine all individual objects (nuclei) and their centers of mass. As summarized in a couple of recent reviews [101], [102], a number of methods were developed for the segmentation of nuclei images over the last fifty years. One common approach is used when a shape with several concave points appears in the image after thresholding. In this approach, this object is interpreted as two or more merged nuclei [103], [104]. However, most common are approaches based on the watershed algorithm [105],[106], which is either applied to the original data, or on a distance-transformed binary image [102]. Watershed-based methods are implemented for example in widely used software such as ImageJ and CellProfiler. In most of these methods, difficulties arise when nuclei cover
a large fraction of the surface. In most of these cases, the performance can be significantly improved by manual pre-processing of images.

We here develop a fully automated procedure optimized for cell monolayers where the cell nuclei do not overlap in principle. The routine (Figure 2.2) is particularly tuned to recognize boundaries between cell nuclei, hence, avoiding undercounting in the relevant image. Its foundation is a minimum intensity mask that is built around each nucleus, before the local threshold is applied. It is implemented in two stages. The first stage involves building a mask based on local intensity minima (space between cell nuclei) and in the second stage, the mask is superimposed on the original image before a local thresholding is performed. The binary image obtained is used to determine the boundaries of the cells' nuclei. Our fully automated method works with accuracies larger than 99% as determined by manually counting segmentation errors (see Supplementary Section 1).

In the first step we find low intensity regions in the image. Initially, the contrast of the image is increased by linear remapping of the original image intensity range onto the intensity interval [0, 255] (Figure 2.2 C). To determine the mask, the environment of every pixel in the image is tested, by creating a 36 arms of the stencil on the circle of fixed radius $r$, with the reference pixel being in the center (Figure 2.2 A). Radius $r$ is the first manually adjusted parameter and depends on magnification and camera resolution used to acquire the images. The arms of the stencil closing 180° are coupled resulting in 18 arm pairs. If extremal pixels of the arms are brighter than the reference pixel by predefined value $\Delta I_c$ (Figure 2.2 B), the reference pixel is stored in a 2D array, which is updated at every step. After all pixels in the image are tested the 2D array forms an image of low intensity regions (Figure 2.2 C), roughly representing the space between nuclei. For distinct experimental conditions (staining method, magnification, and camera), $r$ and $\Delta I_c$ are kept constant for all images ($r=9, \Delta I_c=3$).

In the following steps (two and three) the image of low intensity regions is post-processed and the mask is created. In the second step, pixels of low intensity which are not part of the space between nuclei are eliminated by performing a set of morphological operations to the image of low intensity regions (Figure 2.2 C). Specifically, a pixel is set to be white (element of the region) if more than five pixels in its 3x3 neighborhood are white. Disconnected small white objects are removed from the image. The resulting image is then subject to dilatation, skeletonization, and to the removal of spur pixels, yielding a one pixel thick network (Figure 2.2 E) that is a precursor of the future mask.

The image obtained still suffers from dangling branches in the network, which may be false boundaries (yellow arrow in Figure 2.2 E), or
Figure 2.2: Figure 2. Nuclei segmentation protocol. (A) 18 pairs of stencil arms. (B) Intensity values along the two arms of the stencil making a par (blue line in the (A) panel). In this particular example, there is at least one pair of arms for which the difference in brightness between the referent pixel and the arm ends $\Delta I$ is larger than $\Delta I_c$, as shown in the graph. Consequently, this pixel contributes to the image of low intensity regions. (C) Image of nuclei after contrast increasing. (D) Image of low intensity region acquired after the first step in the segmentation procedure. (E) Precursor of the mask emerging from the second step of the segmentation procedure. Examples of false and disconnected dangling branches are indicated with yellow and red arrows. (F) The completed mask. (G) Segmentation mask is superimposed onto the enhanced image before thresholding. (H) A region of interest from the original image. (I) Result of the first step of the threshold procedure. (J) Result of the local thresholding procedure. (K) Boundaries of the nuclei (red) extracted from image after thresholding. Scale bars are 10 $\mu$m.
parts of missing boundaries (red arrow). To reconstruct these, a search for a matching dangling branch is performed in the radius of 27 pixels (8.4 µm), around each disconnected end in the image. If a match is found and the extension of the branch found closes an angle that is less then $\pi/4$ with the original branch, the two dangling ends are connected with a straight line. If such a connection is not possible, and if it is shorter than 4.2 µm the dangling branch is removed.

Unconnected branches longer than 4.2 µm are extended in the straight line until a connection with the rest of the network is made, if they are found to cut through a middle of a cell that has an area 140% of the average cell in the image. Using this procedure, most of the network becomes enclosed, which completes the mask (Figure 2.2 F).

**Thresholding**

Staining the DNA of the nuclei with Hoechst can result in large intensity variations from one nucleus to another (DNA content, DNA compaction). As a result, a simple threshold applied to the image underestimates the size of darker nuclei and overestimates bright nuclei. To avoid this inconsistency, a local threshold procedure is applied in a three-step fashion. In the first step, a mask is superimposed to the original image, which provides a set of well separated nuclei. In a second step a threshold with a very high value is applied prior to an object search (Figure 2.2 I). Pixels belonging to each nucleus are memorized. In the third step, the mean intensity value of the original image at those pixel positions is calculated. The local threshold value is set to 60% of the original mean intensity value for each nucleus individually (Figure 2J). The objects obtained are used to find the nuclei boundaries (Figure 2.2 K) with an inbuilt MATLAB procedure.

Estimation of errors introduced by the segmentation routine. Segmentation issues most often occur at high densities, when two neighboring nuclei are of very different brightness, which results in merging two nuclei into one (undersegmentation), or due to the inhomogeneous intensity of a nucleus, leading to the recognition of two objects (oversegmentation). To evaluate the accuracy of our approach, these errors are corrected manually in 0.65% of cases on a sample of 10,000 cells. Further details, together with the evaluation of segmentation errors for tissues grown on glass and very soft substrates, are shown in the Section 1 of the Supplementary Materials. Importantly, even on very soft substrates (0.6 kPa), where the density is as high as 13,000 cells/mm² our approach recovers 99% of cells correctly, outperforming watershed based algorithms implemented in ImageJ and
CellProfiler software packages (Supplementary Section 1).

Calculating morphological properties of cell nuclei

The described protocol allows for the accurate determination of a number of morphological measures for each shape (nucleus) in the image including the area and the perimeter. Moreover, centers of mass and orientations of cell nuclei were calculated from binary images using a built-in MATLAB function “regionprops” contained in the Image Processing Toolbox. This function returns measurements of shape properties for each connected component (nuclei) in the binary image. The connected components are labeled using the flood-fill algorithm with the connectivity four implemented in the “bwconncomp” function.

Membrane segmentation

Immuno-fluorescent staining, imaging and segmenting the membrane is a key step for building a reference set of cell shapes used as the source of the “true”, or the so-called “directly measured”, data. Here we use β-catenin staining that reveals the position of the cell membranes and cell-cell contacts. This picture is subdivided into 95 parts and for each part image contrast is increased (intensity histogram in each segment is linearly stretched such that 1% of data is saturated at low and high intensities). At this stage, h-minima transform is applied - all intensity minima with an intensity depth that is smaller than the critical value are suppressed using the “imhmin” function in MATLAB. Subsequently, watersheding is performed with pixel connectivity eight using a built-in MATLAB function based on the Fernand-Meyer algorithm.

To check for the sensitivity of the segmentation protocol, the analysis is repeated with several critical values for the minima depth (Figure 2.3 A-C, and supplementary sections 2 and 3). Setting the critical depth of the minima to 35 induced more oversegmentation errors, while the value 45 was associated with significant undersegmentations. Setting the depth to 40 resulted in the correct reconstruction of 98.39% cells. This was determined on a sample of 17,850 cells grown on hard gels by comparison with images that were manually corrected by combining nuclei and membrane pictures. A similar extent of errors is obtained for tissues grown on very soft gels, while larger deviations are generated on glass due to the relatively low intensity of β-catenin on cell-cell contacts at the observed densities (Supplementary Section 2).
Figure 2.3: (Sensitivity of the membrane segmentation protocol to the parameters of the watershed algorithm. Results for several choices of minimum intensity depths are presented (A-C). Arrows point to oversegmented (red) or undersegmented (white, yellow) boundaries. Scale-bar represents 10 µm.

Morphological features of cells are obtained from segmented images by a self-developed MATLAB routine. First, vertex pixels of cells are found in the image and sorted in a clockwise direction for each cell. Then all pixels between vertices are detected as boundary pixels. To obtain the boundary length between two vertices, the distance between successive boundary pixels, adopting values of 1 or $\sqrt{2}$, is determined and counted. The perimeter of the cell is the total boundary length between all vertices of the cell. The area is the sum of the pixel areas associated with the object, half of the area associated with boundary pixels and a third of the total vertices area. The number of neighbors is the number of vertices belonging to each cell. Elongation and orientation are obtained with the MATLAB “regionprops” function (see above).

2.1.4 Properties of Centre of Mass Voronoi Tessellation

Construction of the Voronoi tessellation

With a set of distinct points in a continuous space, the Voronoi cell is defined as the region that contains all locations closer to the specific discrete point than to any other [107]. Even though similar regions were published by Descartes and later by Dirichlet (2D and 3D case), the term Voronoi region is nowadays most commonly used. It is termed after Voronoi who studied those domains in a general n-dimensional space. In other words, if n centers of mass of nuclei are given $c_1, c_2, \ldots, c_n$, the Voronoi region associated to the center of mass of the cell nucleus $i$ is given by:
VT_i = \{ x \in X \mid d(x, c_i) \leq d(x, c_j), \ i \neq j \}

where X is a metric space with a distance function d in 2D Euclidian space. An algorithm to compute such tessellations is available online \[108\] and is implemented in C++, Python and MATLAB. We use the software package Computational Geometry for MATLAB based on Qhull for computation of Voronoi tessellations. The set of input parameters are the coordinates of seeding points and output is a list of vertices defining the tessellation. As seeding points we use the centers of mass of all cell nuclei that are completely within the field of view. As a result, one generates a set of polygonal non-intersecting objects, which is intrinsic to CMVT.

CMVT Morphological Measures

For an arbitrarily shaped polygonal object, all morphological measures can be obtained from the positions of the vertices \[109\]. Specifically, for the polygonal cell given by n vertices (x_i, y_i) characteristic for CMVT, the area A is given by:

\[ A = \frac{1}{2} \sum_{i=1}^{n} |x_i y_{i+1} - x_{i+1} y_i|, \]

while the perimeter is simply

\[ L = \sum_{i=1}^{n} \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2}. \]

The sum runs over a closed path spanned by all vertices and the n+1 element in the sum corresponds to first vertex. The number of neighbors is in principle equal to the number of vertexes, since corrections for vertices shared by more than three cells are negligible in our sample. The elongation e of the cell is calculated from the principle (orthogonal) moments of inertia I_1 and I_2:

\[ e = \sqrt{\frac{I_1}{I_2}}. \]

The moments are obtained from the diagonalization of the inertial tensor with components I_{xx}, I_{yy} and I_{xy} calculated in an arbitrary rectangular coordinate system spanning the xy plane. As a result one finds:
\[ I_{1,2} = \frac{1}{2} (I_{xx} + I_{yy}) \pm \sqrt{(I_{xx} + I_{yy})^2 - 4(I_{xx}I_{yy} - I_{xy}^2)} \]

where \( I_{xx}, I_{yy}, \) and \( I_{xy} \) are given by the raw moments of the density distribution within the cell, which is assumed uniform. More specifically:

\[ I_{xx} = m_{02} - \frac{m_{01}^2}{m_{00}}, \quad I_{yy} = m_{20} - \frac{m_{10}^2}{m_{00}}, \quad \text{and} \quad I_{xy} = m_{11} - \frac{m_{10}m_{01}}{m_{00}}. \]

Zeroth, first and second moments of a regular 2D polygon are listed here:

\[ m_{00} = A, \]
\[ m_{10} = \frac{1}{6} \sum_{i=1}^{n} (x_i + x_{i+1})(x_iy_{i+1} - x_{i+1}y_i), \]
\[ m_{01} = \frac{1}{6} \sum_{i=1}^{n} (y_i + y_{i+1})(x_iy_{i+1} - x_{i+1}y_i), \]
\[ m_{11} = \frac{1}{24} \sum_{i=1}^{n} (x_i - x_{i+1}) \left[ x_i \left( 3y_i^2 + 2y_iy_{i+1} + y_{i+1}^2 \right) + x_{i+1} \left( y_i^2 + 2y_iy_{i+1} + 3y_{i+1}^2 \right) \right], \]
\[ m_{20} = -\frac{1}{12} \sum_{i=1}^{n} (x_i^3 + x_i^2x_{i+1} + x_ix_{i+1}^2 + x_{i+1}^3)(y_i - y_{i+1}), \]
\[ m_{02} = \frac{1}{12} \sum_{i=1}^{n} (y_i^3 + y_i^2y_{i+1} + y_iy_{i+1}^2 + y_{i+1}^3)(x_i - x_{i+1}). \]

The standard deviation of from the mean boundary length is given by:

\[ \langle \Delta L \rangle = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} \left( \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2 - L} \right)^2}. \]

The orientation is determined by finding the coordinate system where
the off diagonal terms \( I_{xy} \) vanish:

\[
I'_{x'y'} = \sin \theta \cos \theta (I_y - I_x) + \cos (2\theta) I_{xy} = \frac{1}{2} \sin (2\theta) (I_y - I_x) + \cos (2\theta) I_{xy} = 0.
\]

This yields the orientation angle \( \theta \)

\[
\theta = \frac{1}{2} \tan^{-1} \left( \frac{-2I_{xy}}{I_{yy} - I_{xx}} \right).
\]

### 2.1.5 Results

#### Generation of the sample

The key step in the comparison of CMVT with the true morphology of the cells is the construction of the sample of cells, which will be used for this analysis. Our main sample (hard PA gels) consists of 23 images of cell nuclei and membranes. Those images are segmented with the procedure described above providing centers of nuclei as well as outlines of the cell membrane. The first criterion that a cell has to satisfy to be a part of the set is that it has to have correctly segmented nucleus and membrane. Yet, segmentation disparities are obviously small and amount to about 2% on hard and very soft gels, and about 7% on glass. This comprises the direct error of the nuclei and membrane segmentation (for details on segmentation errors at different conditions see Appendix 5.4).5.4)

Besides correct segmentation, in order to be part of the set, the entire neighborhood of the cell has to be within the field of view. This immediately excludes from the statistics all cells that are at the outer edges of the images (15% on hard gels, 17% on glass and 11% on very soft gels), since for them, it is not possible to unambiguously reconstruct the tessellation. Moreover, due to the differences in positions of the cells’ nuclei and cell membranes relative to the boundary of the image, there is a 1% difference in the number of cells excluded in the nuclei and membrane channel.

To eliminate false recognitions we, furthermore, introduce the criteria that 95% of the segmented nucleus must be contained inside the segmented membrane, which is not occupied by another nucleus by more than 5% of the total nuclei area. Cells that do not satisfy this criterion do not contribute to the statistics (about 11% for hard PA gels, 28% for soft gels and 14% for glass). This criterion is introduced to account for the fact that imaging of the membrane and the nuclei require focusing in different planes above the substrate. Namely, adherent junctions associated with \( \beta \)-catenin are, in our samples, positioned slightly above the midline of the
Figure 2.4: (A) Confocal reconstruction (z-x) of β-catenin and the nucleus shows that β-catenin is slightly above the equator of the nucleus. (B) Errors intrinsic to the data due to deviations from the epithelial tubular structure, which in 2D projections appear as nuclei protruding into neighboring cells. Scale-bars represent 10 µm. (C) Table of K-S test results and (D) the distribution of cell areas building the data set, obtained for three different segmentation parameters.
cell nucleus (Figure 2.4 A). Therefore the brightest point of the β-catenin picture can be above the equator of the nucleus along the z-axis. If even small deviation from the tubular shape of the cell takes place, the nucleus will appear outside its membrane in the 2D x-y projection (Figure 2.4 B). Beside this problem, which is intrinsic to the acquired data, similar effects arise if the threshold value set during the segmentation of a nuclei was set too low. However, this type of error is significantly reduced by the variable threshold introduced in the image analysis.

In the selection procedure described, a total of 16% of cells grown on hard gels, 25% on glass and 33% on very soft gels are excluded from the statistics. While this is a significant fraction, the advantage of this stringent set of criteria is the insensitivity of the representative set generated on the free parameters in the sampling protocol (Supplementary Section 2.3). This is evident from the assessment of probability distributions calculated for all morphological measures for the three segmentations used in Figure 2.3. Here, each ensemble of segmented cells is independently subject to the elimination procedure described above, resulting in three representative sets (Fig. 2.4 D). These three “true” sets are compared with Kolmogorov-Smirnov (K-S) test providing a p-value (probability that two distributions are the same) as well as the maximal distance between two cumulative distribution functions as presented in the table (Figure 2.4 C).

$$D = \sup \left| CDF(M^{Memb})_1 - CDF(M^{Memb})_2 \right|$$

For example, for the distribution of cell areas (Fig 2.4 D) the K-S test does not reject the hypothesis with p-value equal to 100%, and the maximal distance between two distributions being 0.32% and 0.2% respectively, showing that all three sets are statistically nearly identical. The number of neighbors, even though it is accepted by the test, has the smallest p-value due to discretized nature of this measure. Moreover, this measure is most sensitive to the segmentation errors occurring in the immediate neighborhood of the cell of interest. Most importantly, this analysis shows that the uncertainties of the measured morphological features are very small, and hence can be taken as excellent representatives of the true cell shape characteristics.

Our final set consists of 15,014 cells grown on hard substrates, which allows us to study the morphology of cells ranging in area from 74 to 274 µm² with an appropriate statistical accuracy. These cells are classified in 20 subsets according to their size (Figure 2.5), where each subset has a width of 10 µm² and contains at least 68 cells. For each cell, we determine
selected morphological characteristics, first from the images of the membrane and then from CMVT. These findings are then analyzed in detail as described below. Additionally, two smaller sets are constructed for testing CMVT on tissues cultivated on glass and very soft gels (Supplementary Sections 4 and 5).

Comparison of morphological measures obtained from membrane images and CMVT estimates

We first compare the probability distributions of measures emerging from tessellations $M^{CMVT}$ with directly measured ones from images of the membrane $M^{Memb}$ (Figure 2.6 A-C) for basic measures such as area, perimeter and the number of neighbors, and (Figure 2.7 A-C) for anisotropy measures such as the elongation, mean deviation of the contact angle and the co-alignment of principle axis of the cell nuclei and the cell body. The correlation

\[
Corr_{CMVT}^{Memb} = \frac{\sum_{i=1}^{N_k} (M_i^{CMVT} - \langle M^{CMVT} \rangle)(M_i^{Memb} - \langle M^{Memb} \rangle)}{\sqrt{\sum_{i=1}^{N_k} (M_i^{CMVT} - \langle M^{CMVT} \rangle)^2} \sqrt{\sum_{i=1}^{N_k} (M_i^{Memb} - \langle M^{Memb} \rangle)^2}}
\]

between the two measured and the CMVT estimated distributions of a morphological measure is shown in the inset. Here the average in the bracket denotes the average of the respective distributions and the sum runs over all cells in the set.

Furthermore, sorting by the measured areas (Figure 2.5), we build 20 subclasses (indexed with $k$), each containing $N_k$ cells ($N_k > 68$ and every cell in the subclass is denoted by an index $i=1 \ldots N_k$). The mean value of a
particular morphological measure \( M_{CMVT}^{k} \) and \( M_{Memb}^{k} \) is presented as function of the mean cell area in each subclass and shown in the second row of Figure 2.6 and 2.7 (panels D-F). We compare data on cell-by-cell basis and calculate the mean relative error \( Err_{M_{CMVT}^{k}} \) of a measure \( M_{CMVT}^{k} \) comparative to \( M_{Memb}^{k} \) in each subclass:

\[
Err_{M_{CMVT}^{k}} = N_k^{-1} \sum_{i=1}^{N_k} \left| \frac{M_{CMVT}^{i} - M_{Memb}^{i}}{M_{Memb}^{i}} \right|
\]

This error measure (third rows - Figure 2.6 G,H and Figure 2.7 G,H) denotes the average deviation of the CMVT estimated from the measured magnitude of a morphological characteristic of interest as a function of the mean cell area in a particular subclass. Exceptionally, for the number of neighbors (Figure 2.6 I), and the co-alignment (Figure 2.7 I) we report the mean difference

\[
\Delta M_{CMVT}^{k} = N_k^{-1} \sum_{i=1}^{N_k} \left| M_{CMVT}^{i} - M_{Memb}^{i} \right|
\]

In the insets of the graphs, we show the mean tessellation error calculated for all of N cells in the set

\[
\langle Err_{M_{CMVT}} \rangle = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{M_{CMVT}^{i} - M_{Memb}^{i}}{M_{Memb}^{i}} \right|
\]

In the case of number of neighbors and co-alignment of the nuclei and cell body, we calculate the mean difference

\[
\langle Err_{M_{CMVT}} \rangle = \frac{1}{N} \sum_{i=1}^{N} \left| M_{CMVT}^{i} - M_{Memb}^{i} \right|
\]

which averages the deviation over all cells in the sample.

Finally, we show the distribution of relative errors (Figure 2.6 J,K, Figure 2.6 J,K) or the distribution of differences (Figure 2.6 L and Figure 2.6 L) in the bottom row for several subsets of cells. Here we focus on particularly small and large cells, as well as a set of cells of intermediate size (as indicated in Figure 2.5 ), to see what type of cells actually contribute to the error of the tessellation.
Basic measures

The most commonly discussed morphological characteristic of cells in a tissue is the average area or cell density. The analysis of CMVT prediction shows that the distribution of cell area is reasonably well reproduced. This agreement is confirmed by the comparison of the probability distributions of the areas measured and areas of cells obtained from the tessellation (Figure 2.6 A), and the relatively high degree of correlation between the two distributions. However, further analysis over the subclasses (Figure 2.6 D), shows that the areas of larger cells are systematically underestimated, and the areas of smaller cells overestimated by the tessellation. Accordingly, the distributions of errors presented for small, midsized and large cells (Figure 2.6 J) are not centered at zero. For small cells the offset is toward positive values, while for large cells is it toward negative values showing a systematic error of the tessellation that makes small cells larger and large cells smaller. Consequently, the mean subpopulation error increases toward the two extrema in cell sizes (Figure 2.6 G). Nevertheless, the areas of the cells are reasonably well reproduced by CMVT, and the mean error is about 10%. Notably, the mean cell size of the set is estimated with 0.25% error, which justifies the utilization of CMVT in estimations of the mean cell density, a result that should be seen in the light of the uncertainty of the mean “true” area $\langle M_{Meas} \rangle$ of 0.04% (Figure 2.4 D).

Significantly stronger deviations of CMVT from the true data can be seen in the distribution of perimeters (Figure 2.6 B). This measure is mainly affected by the curved nature of the cell wall, which the tessellation approximates with a straight line. Consequently, the length of the boundaries is systematically underestimated, shifting the whole distribution to smaller values. Naturally, best performance is obtained for relatively small cells (Figure 2.6 E). Nonetheless, the mean error of the tessellation is only 9.4%. Moreover, since the perimeter of the cell is strongly correlated with the cell area, the systematic size dependent errors prevail and are even somewhat enhanced (Figure 2.7 E,H,K). This is particularly acute for large cells where CMVT underestimates the cell perimeter more than 16% on average (Figure 2.7 H).

The distribution of the number of neighbors (Figure 2.6 C) is well accounted for by the tessellation, even though the number of cells having 6 neighbors is systematically overestimated. The distribution is slightly skewed for both measured and tessellated data sets, and the dependence on the cell size (Figure 2.6 F) is well reproduced. The more detailed analysis of the errors generated (Figure 2.6 L) shows that despite centering at zero, overestimates of the number of neighbors are more common for in-
Figure 2.6: Comparison of CMVT (yellow squares) and directly extracted morphological measures (blue circles). The graphs associated with cell areas, perimeters and number of neighbors are shown in the first, second and third column, respectively. Top graphs (A-C) show the probability distribution generated by direct measurement and estimated from the tessellation. (D-F) Second row is the average dependence of morphological characteristics on the cell area. The associated deviations of CMVT are shown in the third row (G-I). The distributions of errors for small, intermediate-sized and large cells are shown in the bottom row (J-L).
termediate cell sizes, unlike for very large and small cells that tend to underestimate the number of neighbors. In fact, on average, in 22.3% of cases CMVT overestimates, and in 20% underestimates the number of neighbors by one. Actually, CMVT correctly estimates the number of neighbors only for 46.5% of cells, which yields a low correlation index. This poor performance is also associated with the segmentation errors in the immediate environment of the cell of interest. Therefore, we conclude that CMVT provides the mean number of neighbors in a quantitative manner, but not on a single cell level.

Measures of cell anisotropy

Several measures such as the elongation, the standard deviation of boundary length and the co-alignment angle between the principle axis of the cell and its nucleus are all sensitive to the anisotropic properties of the cell shape. For the cell elongation and the co-alignment of the cell’s body and nuclei CMVT only qualitatively represents the probability distributions (Figure 2.7 A,C). As in the case of perimeters, these measures show systematic errors. While the mean error of the tessellation remains between 10% and 15%, the correlation between the measured and the estimated distributions remains only about 0.5. Interestingly, the elongation seems to be independent of the cell size and is around 1.3. However, it is systematically underestimated by CMVT - the cells turn out more spherical than they are (Figure 2.7 D). Consequently, the distribution of errors is negatively skewed (Figure 2.7 J). At the same time, the co-alignment between the cell and its nuclei is underestimated by the tessellation (Figure 2.7 F). Naturally, the distribution of errors (Figure 2.7 L) is positively skewed, even though the maximum of the distribution is still around zero.

Overall, these results suggest that in isotropic cells no particular nuclear polarity takes place, as expected, a result that is captured by CMVT but only on average and not on the single cell level. However, since most cells in this tissue have an elongated shape, associated with the co-alignment of the cell and the nuclei, significant errors are generated by the tessellation. The variation of boundary length, which should be larger for elongated cells than for isotropic cells, itself increases linearly with the cell area (Figure 2.6 E), which is interesting in the light of the insensitivity of the elongation to the cell area, and the fact that linear increase is expected with the cell perimeter. Importantly, CMVT well reproduces this trend, but still has mean error of 11.2% on the level of the single cell.

Unlike basic measures, the morphology measures associated with cell anisotropy do not show errors that are strongly size dependent, but seem
Figure 2.7: Comparison of CMVT (yellow squares) and directly extracted morphological measures (blue circles). The graphs associated with cell elongation, standard deviation of boundary length and co-alignment of the nuclei and the cell are shown in the first, second and third column, respectively. Top graphs (A-C) shows the probability distribution generated by direct measurement and estimated from the tessellation. (D-F) Second row is the average dependence of morphological characteristics on the cell area. The associated deviations of CMVT are shown in the third row (G-I). The distributions of errors for small, intermediate-sized and large cells are shown in the bottom row (J-L).
systematic. Consequently, the distributions of errors do not change shape for different subclasses of cell sizes (Figure 2.6 J-L).

Cross-correlations between measures

From the previous discussion, it becomes evident that there must be a degree of correlations between various morphological measures. To quantify this, we calculate the correlation coefficient $\text{Corr}_{Ma}^{Mb}$ which measures the extent of the linear relationship between two estimated or directly measured properties of the cells’ shape $M_a$ and $M_b$:

$$\text{Corr}_{Ma}^{Mb} = \frac{\sum_{i=1}^{n} (M_{a,i} - \langle M_a \rangle)(M_{b,i} - \langle M_b \rangle)}{\sqrt{\sum_{i=1}^{n} (M_{a,i} - \langle M_a \rangle)^2 (M_{b,i} - \langle M_b \rangle)^2}}$$

The mean in the brackets denotes the mean of the distributions shown in Figure 2.6 A-C and Figure 2.7 A-C, and the sum is performed over all cells in the set. The correlation coefficients as defined above are presented in the table 2.8 C while the scatter plots representing strong correlation (area and perimeter), no correlation (area and elongation) and weak anti-correlation (co-aliment and elongation) are presented for the ”true” data (2.8 A) and the CMVT estimate (2.8 B). We see that CMVT captures appropriately the level of cross-correlation between various measures.

Interestingly, the only strong cross-correlation in the data is between the cell areas and perimeters, evidenced from scatter plots shown in Figure 2.8 A,B. The cross-correlation is even overestimated by CMVT (Figure 2.8 C), due to the inability of CMVT to capture small fluctuations of the cell boundary arising in the true data.

Surprisingly, cell areas are not strongly cross-correlating with any other measures, neither in the true set, nor in the CMVT reconstruction. Weak cross-correlations exist with the number of neighbors, and the standard deviations in the boundary length. The latter is the consequence of cross-correlation between perimeters and the variations in the boundary lengths, which is most likely of purely geometric origin, as well as the weak cross-correlations between the variation in boundary length and the cell elongation. Interestingly, we find that the co-alignment of the cell nuclei and the cell body does not cross-correlate with the cell area, despite the expectation that in smaller cells, stress on the nuclei will be strongly coupled to the stress on the cell membrane. Likewise, no appreciable relation between the cell area and the cell elongation emerges from our data. This trend is well reproduced by CMVT.
Figure 2.8: Scatter plots and cross-correlations. (A,B) Exemplary scatter plots showing the cross-correlation (left panels) or the lack of it (middle and the right panels) between various morphological measures. (C) Correlation coefficients between two morphological measures, as estimated by CMVT (yellow numbers on the left) and directly measured (blue numbers on the right in each column). Cross-correlation coefficient can adopt values between -1 for anti-correlated measures to 1 for fully correlated ones.

<table>
<thead>
<tr>
<th></th>
<th>PERIMETER</th>
<th>NEIGHBORS</th>
<th>ELONGATION</th>
<th>STD(ΔL)</th>
<th>CO-ALIGNMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>AREA</td>
<td>0.98</td>
<td>0.96</td>
<td>0.29</td>
<td>0.24</td>
<td>-0.02</td>
</tr>
<tr>
<td>PERIMETER</td>
<td>0.21</td>
<td>0.26</td>
<td>0.06</td>
<td>0.19</td>
<td>0.48</td>
</tr>
<tr>
<td>NEIGHBORS</td>
<td>0.04</td>
<td>0.06</td>
<td>0.24</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>ELONGATION</td>
<td></td>
<td></td>
<td></td>
<td>0.35</td>
<td>-0.28</td>
</tr>
<tr>
<td>STD(ΔL)</td>
<td></td>
<td></td>
<td></td>
<td>-0.33</td>
<td></td>
</tr>
</tbody>
</table>

-0.05 -0.01
In the context of other anisotropy measures, it was already anticipated in the previous section that elongation of the cells anti-correlates with the co-alignment of the cell and its nuclei. In other words, in elongated cells, the orientation of the nuclei follows the orientation of the cell, whereas in more isotropic cells, this correlation is lost. This trend is well captured by CMVT, as well as the lack of correlation of these two measures with other characteristics of the cell shape.
2.2 Nuclei shape based tessellation method reconstructs epithelium morphology more accurately

2.2.1 Method

The commonly used approximation for cell membrane shapes is Voronoi tessellation generated from the nuclei centers of mass. To improve the accuracy of reconstructing the membranes from the nuclei positions, one can utilize more information than just the centers of nuclei mass. Here we show that using nuclei boundaries to reconstruct the cell membrane positions may lead to interesting results.

Specifically, use of the shape based Voronoi tessellation where the cell is defined as all the positions in the image closer to the set of points from the boundary of a given cell than to any other set of points from the boundaries of other cell nuclei. In other words, we have n subsets of the plane \( p_1, p_2, \ldots, p_n \) and we are dividing the plane in such a way that the i-th part of the plain is closer to the i-th subset \( p_i \) than to any other subset. The Voronoi region associated with the subset \( p_i \) is given by:

\[
VT_i = \{ x \in X \mid d(x, p_i) \leq d(x, p_j), \quad i \neq j \},
\]

where \( X \) is a metric space with a distance function \( d \) in 2D Euclidian space and the subsets \( p_1, p_2, \ldots, p_n \) represent boundaries of n cell nuclei. This space tessellation method consequently has curved edges, resembling the real edges of cells in a tissue.

If cell nuclei were circles, this tessellation would be an additively weighed Voronoi diagram also known as Apollonius diagram for which \( d(x, c_i) = \|x - c_i\| - w \). Here, \( c_i \) represents the center of mass of the i-th circle, and the edges of this tessellation are sections of hyperbola. There are a number of algorithms available to solve this problem \[110\], \[111\]. The case of asymmetric particles is much less studied. For example, an algorithm for calculating the Voronoi tessellation in the case of 3D packing of nonspherical convex particles was proposed \[112\]. This algorithm is based on the so called Delaunay empty sphere method which calculates the trajectory of the imaginary empty sphere of a variable size. Another study offered a more practical algorithm: they discretize the surface of an object to a sufficiently dense set of vertices and compute the classical Voronoi tessellation for every point on the surface \[113\]. This algorithm was later practically used to study ellipsoid packing in 3D \[114\]. The study on surface
Figure 2.9: Two Voronoi tessellation methods (A) Cell nuclei image (stained with Hoechst) is segmented with our segmentation method (red) (B) Construction principle of CMVT is demonstrated on two schematic puzzles (left) and several cells (right). Cell boundaries are mid-lines between cell nuclei centers of mass. CMVT laid over (B) cell nuclei and (D) the membrane. (E) Cell membrane (stained with β-catenin) is segmented with watershed algorithm (blue). (F) Construction principle of SVT is demonstrated on two schematic puzzles (left) and several cells (right). Cell boundaries are lines between the edges of cell nuclei. SVT laid over (G) cell nuclei and (H) the membrane. Scale bar is 10 µm.

nanobubbles also used this kind of tessellation, but the algorithm was not declared [115]. To our knowledge, until now, this tessellation hasn’t been used in biophysics. Although it was previously termed Set Voronoi diagram or Voronoi S tessellation, we refer to it as shape based Voronoi tessellation (SVT) since it is based on the shape of the cell nuclei.

We develop a fast and simple algorithm for calculating SVT on a binary image with well separated objects (segmented cell nuclei). The precision of our input data is limited by the pixel size. With our segmentation procedure (see Section 2.1.3), over and under segmentation of the cell nuclei objects happens in less than 1% of the cases (see Appendix Table 5.18). However, how precisely the nuclei boundary is estimated is not quantified due to the lack of an objective method to estimate the nuclei boundary. Fortunately, one can see that those errors are very small (see Appendix Figure 5.23 [5.24]). We construct the SVT from the resultant segmented binary images of the cell nuclei.

Our algorithm is the following: we perform the distance transforma-
tion which associates to every pixel a Euclidean distance to the nearest white pixel. Consequently, the cell boundaries are extracted by the watershed algorithm applied on the distance transformed image. This algorithm gives us a connected set of pixels (connectivity eight) that are at the maximal possible distance from all the surrounding cell nuclei. Those pixels form a network which tessellates the space very similarly to the real cells in the tissue. Accuracy is defined by the pixel size, which in our case is 0.31 μm.

The output of this algorithm is a binary image defining the cell borders which is the same as the output from the membrane segmentation process. Therefore, the morphological measures of shape based Voronoi cells are calculated in the same way as the properties of the cells (see Subsection 2.1.3). Efficiency of our MATLAB routine is approx. 700 cells/min on a common desktop. This includes segmentation of cell nuclei, followed by construction of the SVT and the computing of all the morphological properties of shape based Voronoi cells.

A shape based Voronoi cell includes by definition a cell nucleus and predicts curved cell boundaries (see Figure 2.9 G). This was not the case for the CMVT method. If one visually compares shape based Voronoi cells to the real cells (see Figure 2.9 H), the precision of the approximation as well as the improvement to the classical center of mass Voronoi method becomes obvious. To prove and quantify this visual impression we perform our error estimation analysis established in the previous Section 2.1 on CMVT method. We divide cell’s morphological measures into basic and anisotropy ones and systematically compare the performance of both tessellations $X^V$ in reconstructing morphological properties of the cell membrane $X^M$. Based on the cell areas we construct 20 classes which we index with $k$ (see 2.1.5). Each class contains $N_k$ cells ($N_k > 65$ and every cell in the subclass is denoted by an index $i=1\ldots N_k$). We plot the distributions of both basic (Figure 2.10 A-C) and anisotropy measures (Figure 2.10 J-L). Furthermore, we present the mean value of a measure $\langle X^V_k / M \rangle$ for each cell area class (middle graphs - Fig. 2.10 D-F and M-O). Finally, for individual cells we compare the values of all the morphological measures of a cell and a Voronoi cell. We calculate the relative error $ErrX^V_k = N_k^{-1} \sum_{i=1}^{N_k} |X^V_i - X^M_i| / X^M_i$ of a cell estimate $X^V$ compared to the “real” value estimated from the cell membrane $X^M$ in each subclass (Figure 2.10 G,H,P,Q). When more appropriate, we determine the mean difference $\langle \Delta X^V_k \rangle = 1/N_k \sum_{i=1}^{N_k} |X^V_i - X^M_i|$ (Figure 2.10 I,R). In the insets of the graphs we show the mean tessellation error $\langle ErrX^V \rangle = 1/N \sum_{i=1}^{N} |X^V_i - X^M_i| / X^M_i$. 
2.2.2 Results

We see that SVT is systematically better than CMVT in predicting the distributions of all of the morphological measures. The most obvious examples are cell perimeter and elongation, the two morphological measures that CMVT systematically underestimates.

Cell area is also better estimated by the SVT method (Figure 2.10 A,D,G). The improvement can be evidenced from the distribution of the cell areas (Figure 2.10 A). Investigating the mean value of the cell area in the area classes reveals that the SVT method tends to slightly overestimate the areas if the small cells and underestimate the areas of the big ones, but much less than CMVT method (Figure 2.10 D). This is obvious in the plot of the relative error (2.10 G) where systematic improvement to CMVT method is seen for all cell sizes, even in the case of the average sized cells where CMVT method gives a very good estimate for the average cell area. Average relative error in estimating cell area is 7.6% in comparison to 10.5% in the case of CMVT. The average value of the cell area is significantly better predicted by SVT. As explained in 2.1.3 these values were very well predicted by CMVT with the error being less than 1%, but in case of SVT this error is now reduced to less than 0.1%. However, those numbers should be considered carefully since they are influenced by the fraction of the cells at the edges of the image which varies with the magnification and camera.

SVT predicts cell perimeters significantly better than CMVT with a mean error of only 5.7% in comparison to 9.4% error made by CMVT. This is a significant improvement to a standard method, and in general very good result. This means that the cell perimeter can be estimated if we assume coupling between cell nuclei and membranes. This connection is somewhat interrupted in large cells, but in the case of average and small cells, the coupling very well predicts the perimeter values.

The estimation of the number of cell neighbors is systematically improved by SVT for all cell areas (Figure 2.10 I). However, the SVT method still overestimates the number of cells with six neighbors, but much less than the CMVT method (Figure 2.10 C). Overall, the average deviation in estimating the number of neighbors is 0.62 which is an improvement in comparison to CMVT. The deviation of the number of neighbors has the same cell area dependence as for the CMVT method with small and big cells contributing the most to this error (Figure 2.10 I).

While cell elongation of small and average cells is very well predicted by SVT method, the value for large cells is underestimated (Figure 2.10 M). Since the large cells are the least represented in the sample (Figure 2.5) the improvement made by the SVT method is sufficient to correctly predict
Figure 2.10: Comparison of SVT (red diamonds), CMVT (yellow squares) and directly extracted morphological measures (blue circles). Top graphs (A-C, J-L) shows the probability distribution of a morphological measure $X_i^M$ (blue histogram), and $X_i^V$ for SVT and CMVT. Middle graphs (D-F, M-O) show the mean value $\langle X_i^V / M \rangle$ of a measure as a function of the mean subclass area. Bottom graphs (G,H,P,Q) present the relative error within the subclass $ErrX_k^V$ as a function of the mean subclass area. Exceptionally, for the number of neighbors (I), and the co-alignment (R) we report the mean difference $\langle \Delta X_k^V \rangle$. The mean tessellation error $\langle ErrX^V \rangle$ is in the insets.
the distribution of the cell elongations (Figure 2.10 J). The average error of the SVT method in estimating cell elongation is 9% but the distribution of the errors is symmetric (except for large cells). The CMVT method, on the other hand, suffers from the systematic error. All and all, SVT method gives a good estimate of the cell elongations. The average cell elongation of the total cell sample (15,000 cells) estimated from β-catenin pictures was 1.36 while estimations made by the SVT approximation was 1.34. This is a significant improvement to the estimate made by CMVT method of 1.24.

Standard deviation also depends on the cell area. The average value of the standard deviation in contact length is much better estimated by SVT than by the CMVT method (Figure 2.10 N). However, the average relative error in estimating this measure is improved by only 1.1% (Figure 2.10 Q), but the improvement is better in the case of the small cells.

The co-alignment between the cell nuclei and the cell body is overestimated by SVT because this method by definition predicts that the cell shape adjusts to the nuclei shape. The CMVT method underestimated the co-alignment because the cells are approximated by polygons and therefore the main axis of cell’s moment of inertia is badly reproduced. This can be seen in distribution of the co-alignment angle (Figure 2.10 L). The mean value of this measure is increasing with cell size. This tendency is also seen in SVT estimate (Figure 2.10 O), but the increased rate is underestimated. This phenomenon can be attributed to the asymmetric membrane protrusions developed by the large cells which cause the lack of co-alignment between cell nuclei and cell body. The mean difference between co-alignment estimates by SVT and β-catenin pictures is 12.5 degrees which is 3.3 degrees better than the one made by CMVT method.

Even though SVT method improved notably the estimation of cell perimeter and elongation the relative error still progresses with cell area (Figure 2.10 H) and the mean values are underestimated in the case of the large cells (Figure 2.10 E). Likewise, the area of large cells is bigger than the SVT estimate and the cell nuclei and body are less co-aligned. This implies non symmetric distribution of the relative errors.

The origin of these systematic errors in the case of the large cells are anisotropic membrane protrusions typically formed by these cells (2.11 D). Due to this unusual cell shape, cells are bigger and have larger perimeter and elongation. Furthermore, they do not co-align their body with their nuclei.

For obvious reasons, perimeter increases with cell area. For this reason, we want to explore the behavior of the rescaled perimeter ($P_{rescaled} = \frac{P}{\sqrt{\text{Area}}}$) with increases in cell area. Plotting rescaled perimeters as a func-
Figure 2.11: Origins of the cell perimeter underestimation by the Voronoi tessellations. (A) Rescaled cell perimeter of cell membrane (blue) and two Voronoi tessellations (SVT is red and CMVT yellow) as a function of the cell area. (B) Rescaled perimeters of a polygon obtained from vertex positions in the tissue (blue), SVT (red) and CMVT (yellow). (C) Difference between rescaled cell perimeters (membrane is denoted in blue and SVT in red) and the rescaled perimeters of a polygon having the same vertexes. This measure reveals how much cell membrane deviates from a straight line. (D) Examples of anisotropic cell membrane protrusions that typically big cells develop. Cells with the area larger than 220 $\mu m^2$ are marked with white circles.
tion of cell area reveals that larger cells have bigger rescaled perimeters (Figure 2.11 A). This means that the real perimeter increases with the cell area not only due to the trivial reason, but also because of the reorganization of the membrane with densification of the tissue. There are two reasons for this phenomenon. Firstly, the positions of the membrane vertices reorganize with densification of the tissue. This is seen as an increase in the rescaled perimeter of a polygon defined by membrane vertices (Figure 2.11 B). Another contribution is the curviness of membrane, which we measure as a difference between the membrane length and the length of straight lines connecting the membrane vertices. Bigger cells have more curved membranes (Figure 2.11 C). SVT does not predict this phenomenon, and estimates that all the cells have equally curvy membranes. Likewise, reorganization of the vertexes in a way that increases perimeter of the rescaled polygon for the big cells is not predicted by both tessellations (Figure 2.11 B). All together, SVT predicts perimeter increases exclusively due to the cell area changes, while CMVT predicts slight decrease of the rescaled perimeter due to repositioning of the vertices as the tissue densifies (Figure 2.11 A).

2.3 Discussion and Conclusions

In the first Section we established a method for systematic comparison of the tessellation made by epithelial cells and the Voronoi tessellation generated from the centers of the nuclei mass (CMVT). The comparison between the two is an important problem in the physiology of epithelium. However, to our knowledge, this comparison was done only once 50 years ago in the original paper by Honda [91]. Honda defined the approximated seed of the Voronoi cell based on the positions of the vertices in the epithelium and calculated the total distance of this seed to the lines from the vertexes. This distance should be zero in the case of the Voronoi cell and is very small for Voronoi-like cells. However, this type of quantification does not reveal the error made by the CMVT method which is today commonly used. To our knowledge, our work is the first estimation of the quality of the CMVT method. We believe that this data will be very useful to the researchers analyzing the morphology of MDCK cells. In the case of different cell lines, we suggest using a similar method not only in biophysics but also in histopathology where the utilization of CMVT is rising.

Furthermore, we have developed a novel algorithm for cell nuclei segmentation. When applied on the MDCK-II monolayer, our MATLAB implemented routine is less prone to segmentation errors than ImageJ and
CellProfiler (one of the most popular free-ware softwares for nuclei segmentation), having the segmentation error of maximal 1% in the case of extremely high cell densities (see Appendix 5.4). Cell membranes are segmented with a MATLAB implemented watershed algorithm applied upon h-minima transform. The segmentation errors of this method applied on the β-catenin images can rise up to 7% (see Appendix 5.4). However, those errors can be easily eliminated with our criteria that membranes that do not contain one and only one nuclei are not included in the statistic. The most fundamental problem of our comparison method lies in the very essence of the cell shape. While the cell nucleus is positioned relatively low (close to the substrate), the adherent junctions and even more so the tight junctions are relatively above the nucleus. However, in visualizing the cell borders one can target either the proteins associated with adherent junction (β-catenin) or the the proteins associated with tight junctions (for example ZO-1), meaning that the fluorescent signal from the cell nuclei and the cell borders will always come from different confocal planes and inconsistency of the two signals can not be avoided. This effect contributes to the relative error and mean difference of an SVT and CMVT cell estimate compared to the values estimated from cell membrane.

The widely used, nuclei center of mass based Voronoi tessellation is polygonal in nature. This method is however satisfying if one is interested in the mean cell density. The source of this error is only from the cells at the border of the image, and therefore the error depends on the size of the image. In the case of our images (430x320 µm) the relative error of the mean cell area was 0.04%. All and all, this tessellation qualitatively correctly captured the trends of the morphological measures and the cross-correlations between them.

The worst performance of CMVT method is in the case of cell perimeter and elongation which are systematically underestimated. Distribution of cell areas is relatively well reproduced, but cell-by-cell analysis revealed relatively large deviations for the sub-populations of small and large cells. Numbers of cells with six neighbors are overestimated by the CMVT method. Furthermore, the CMVT method underestimates the co-aliment of cell nuclei and body due to the lack of information on the nuclei shape. Even though the relative errors in the case of the standard deviation of the contact length are relatively high (more than 11%), due to their symmetric distribution, this measure is only slightly overestimated in the case of the small and middle size cells. Likewise, distribution is relatively well predicted.

Importantly, the same trends are recovered in tissues with relatively low cell density (data from colonies grown on glass substrate) and with
very high cell density (data from colonies grown on soft substrates). This means that relative area of the cell in a monolayer plays a role in the method performance. The limits of the CMVT are easily visually evidenced by many cell nuclei laying outside the associated Voronoi cell. The significant improvement of the cell shape reconstruction can not be obtained by more precise imaging and segmentation of the cell nuclei, due to the polygonal nature of this tessellation. Therefore, a more sophisticated tessellation method is needed to overcome the systematic errors of the CMVT.

An example of such a method is the tessellation based on the shape of the cell nuclei. We provide a very simple and fast algorithm for calculating such space tessellation which we name the shape based Voronoi tessellation (SVT). This simple method significantly improves the estimate of all analyzed cell morphological measures. The downside of the CMVT method were systematic errors in estimating cell perimeter and elongation. The SVT method accurately predicts the values of those measures. However, some of the the systematic errors for relatively large cells — which often have cytoplasmic extrusions — remain because the protrusions are not predicted by SVT. The reason why the big cells form such protrusions is not clear. For all other measures SVT as well gives better estimate than the standard CMVT method. The average values are better predicted and the relative errors are systematically lower than in the case of CMVT. The worst performance of the SVT method relative to the CMVT is in the case of the big cells for morphological measures of number of neighbors, co-alignment angle and standard deviation in contact length. Again, the reason lies in cytoplasmic protrusions not being predicted by the SVT. While the CMVT method overestimated co-alignment angle between cell nuclei and body, SVT method underestimates it. This overestimation is a fundamental property of this tessellation method.

Overall, SVT consistently outperforms CMVT in capturing all studied morphological properties, while maintaining its computational efficiency. SVT provides a remarkably good reconstruction of MDCK tissues architecture and is particularly useful in recovering cell elongation, perimeters and cell areas. In this light, and due to its simple implementation, SVT should become the method of choice for determining morphological properties of tissue.

From the biophysical point of view, the reason why SVT is so successful in predicting the cell shape must lie in fast interplay between the stresses exhibited on the cell membrane and the stresses transmitted to the cell nuclei. Our result implies that cell membrane and nuclei are interacting on the time scales faster than tissue reorganization. In other words, the cell
membrane and nucleus are coupled and the outside stresses are quickly transmitted to the cell nucleus which induces the cell division arrest (for details see the following Chapter 3.5.2). This coupling must happen though the cell cyto-skeleton (microtubules, actin filaments and intermediate filaments), however the role of each of these polymers, which have very different thickness and properties, is to be investigated. Likewise, the future studies should explain if this mechanism is active or passive (do cells use ATP to transmit the stresses to their nuclei).
CHAPTER 3

Morphological characterization and mechano-sensitivity of the MDCK epithelium system model

3.1 Introduction

Since the seminal paper by Pelham and Wang in 1997 [116], the phenomenon of mechanosensing in cells has been explored in great detail. The focus was, at first, on the single cell response to the rigidity of the substrate and the extra cellular matrix (ECM) composition [117], [118], [119]. This response has been explained in physical terms as a consequence of the adhesion force between the cell and the matrix. Cells on rigid substrates have larger focal adhesion sites, and force per focal adhesion site depends linearly on the area of the adhesion site [120]. FA sites are not only larger on stiffer substrates, but they are also associated with greater adhesion forces. This suggests a large-scale mechanosensing mechanism [121]. Consequently, a cell’s spreading area increases with substrate rigidity and with the concentration of ECM proteins [122].

However, cells in a tissue have contact not only with the ECM but also with each other. Therefore, cells which cohere with other cells experience another source of force in addition to focal adhesion. In a system of two epithelial cells in contact, force at the cell-cell contact is estimated to be half of the force between the cell and the ECM for every value of the latter force [123]. However, such estimations are not possible for more than two cells. It is, therefore, unclear how important the focal adhesion actually is once the cells form strong cohesions with each other.
Morphological changes in the cell cluster due to interplay between the focal adhesion and cell-cell cohesion forces has been studied on various cadherin expressing fibroblast cells plated on substrates with different poly-ethylene glycol (PEG) concentrations [124]. Fibroblasts with high cohesivity tend to aggregate even on low PEG concentrations (higher focal adhesion strength), and high PEG concentrations induce aggregation for every cohesivity value. These results suggest that the relative strengths of the normal and parallel adhesion forces play an important role in the final morphology of the cell cluster. The similar results were observed for the fibroblast cells seeded on polyacrylamide gels of various rigidity [125]. On soft gels fibroblast cells merged to form tissue-like structures while on hard gels cells spread and migrate away from each other. It may, therefore, be expected that the cells in clusters also “feel” the substrate rigidity and react to it.

It is not such a well known fact that many studies of the mechano-sensitivity in the cell monolayer failed to repeat the results obtained for the single cells. One of the first studies regarding this showed that bovine aortic endothelial cells - even though they are very sensitive to substrate stiffness as single cells- in a confluent monolayer displayed indistinguishable morphology on substrate rigidities between 180 Pa and 28,6 kPa [126]. Similarly, vascular smooth muscle cells grown on 25 kPa and 135 kPa polyacrylamide gels differed in clusters containing up to three cells, but no effect was found for clusters containing four to seven cells [127]. On MDCK-II cells it was demonstrated that different rigidities of polyacrylamide (PA) gels (0.15-90 kPa) [40] do not influence monolayer morphology nor the attainable steady state density. The authors argue that this is the result of long-range forces between cells transmitted by the underlying elastic substrate.

The MDCK cooperativity, however, was different on glass substrates and PA-gels [60]. This result indicates that, while glass substrates cannot be deformed, cells are able to form long range deformation patterns on PA gels. This fundamental difference between hard PA-gels (deformable substrates) and glass substrates (rigid substrates) could play an important role in monolayer mechano-sensitivity, and this should be considered in addition to merely the strength of focal adhesion forces. However, this study does not address the issue of morphological differences between monolayers on different substrates.

Furthermore, one biochemical study [128] showed a very different reaction of MDCK-II cells to the TGF–β1 growth factor depending on the substrate rigidity. In this study it was shown that TGF–β1 on soft substrates (< 1 kPa) induced apoptosis in the small droplet like cell clus-
ters, while on hard substrates (> 5kPa) it was the trigger for epithelial-mesenchymal transition. This is the only study which, to our knowledge, has observed a morphological difference between small clusters of cells on substrates of different rigidity. Even though, morphological differences between cell clusters on very soft and hard gels could be seen in the images, those differences were not the focus of this study and therefore were not commented upon.

Here we re-address the issue of tissue mechano-sensitivity. Contrary to conventional wisdom, we find that substrate rigidity plays an important role in determining the morphology and the steady state of the cell monolayer. The difference is particularly striking for very soft substrates (0.6 kPa) where cluster growth is fundamentally different. But even on relatively rigid substrates, cells in the monolayer “feel” whether the substrate is the hard gel (3-34 kPa) or the glass (infinite rigidity).

### 3.2 Experiments

Results in this chapter relay on three independent sets of experiments. For each set the gel preparation, coating, seeding, cell medium, fixation and microscopy were those described in 2.1.2. However, the number of seeded cells, the substrate rigidity and the duration of the experiment varied in the three different sets of experiments. All the experiments for this chapter were performed under the supervision of Prof. Dr. Florian Rehfeltd in the 3rd Institute of Physics - Biophysics, Georg-August-University, Göttingen.

The first set of experiments is analysed in Section 3.3. For this set, 30,000 cells were seeded in 7 µL droplets on the glass substrate. Fixation was conducted on days 1, 2, 3, 4, 6, 8, 10 and 12 after seeding. Cells were stained for Hoechst and phalloidin–tetramethylrhodamine B isothiocyanate. The seeding and the staining for this experiment were conducted by Christina Jayachandran and Carina Wollnik. I conducted the image acquisition.

The second set of experiments was performed for the Section 3.4. On soft gels, 90,000 MDCK II cells were seeded in 7 µL on the moist 0.6 kPa polyacrylamide(PS) gels and the cell medium was added very carefully after ca. half an hour. The clusters were periodically fixed up to day twelve (fixation was done after: 2, 4, 6, 8, 10, and 12 days) and stained for nuclei and actin. For single cell experiments, 30,000 cells were dispersed over the entire dish, and periodically fixed after 1, 6, 12, 24 and 48 hours. Cells were stained for Hoechst and phalloidin–tetramethylrhodamine B isoth-
iocyanate. The cells were seeded and stained by Christina Jayachandran, and I acquired the images.

The third set of experiments is analyzed in Section 3.5. For this set of experiments 30,000 cells in $7 \, \mu\text{L}$ were seeded on the polyacrylamide (PA) gels with various rigidity (1, 2, 3, 5, 11 and 24 kPa) and on glass substrates. Cells were fixed periodically after 2, 3, 4 and 6 days. Cells were stained with Hoechst, phalloidin–tetramethylrhodamine B isothiocyanate, anti-paxillin and EdU cell proliferation assay. For this set of experiments Carina Wollnik conducted gel preparation, cell seeding and staining. I conducted the image acquisition and cell staining.

3.3 Growth of MDCK II monolayer on hard substrates

The classical regime of growth refers to unconstrained growth of MDCK-II cells on hard substrates which is characterized by the gradual separation of a cluster into bulk and edge regions. We refer to this regime as the classical one since it is the only 2D regime this far reported in the literature. This regime has been extensively studied in the context of many different problems such as cluster growth and morphology [37], cell dynamics [55], [58], [59], [40], [62], the glass-like behavior of the monolayer [64] or its reaction to stretching [41]. However, a number of studies have used different initial seeding conditions or geometrical constraints on the cell monolayer. Popular choices of geometry are: strip-shaped PDMS stencils, micro-channels, small circular confinements and unconfined circular cell clusters.

In our studies we use relatively large, unconfined, circular cell-clusters. For such cell conditions, cell density had already been measured as a function of time and the distance from the edge of the cluster [40]. Those results showed an increase in the cell density in the middle of the cluster and a gradual decrease towards the edge of the cluster. Experiments conducted in our group [1] revealed that this feature is characteristic for unconstrained MDCK-II cells on hard substrates independent of the initial seeding conditions. If cells had been seeded in a droplet or dispersed over the substrate, or only a few cells had been seeded, they always eventually formed a single cluster with a bulk region with high cell density and an edge region where cell density drops toward the edge of the cluster.

---

1 Experiments conducted by Damir Vurnek in the laboratory of Prof. Dr. Diana Dudziak in the Department of Dermatology, University Hospital at FAU Erlangen.
3.3.1 Image analysis

The experiments described in 3.2 were analyzed in MATLAB using a self-written routine explained in Subsection 2.1.3. Cells seeded on the glass substrates display domes and various other irregularities (see Subsection 3.5.2) over time and this makes the image analysis more difficult on this particular substrate. For this reason, all the images were corrected by hand and the regions with domes are eliminated from the analysis. The steady-state cell-density is, accordingly, the density of the cells between domes. From the centers of the nuclei mass the Voronoi tessellation is constructed and the Voronoi cell area ($A$) is used as an approximation for the cell area. The mean cell density ($\bar{\rho}$) is an inverse of the mean Voronoi cell area ($\bar{A}$). The standard deviation of the Voronoi cell areas ($\sigma_A$) is used to calculate the standard deviation of the cell density ($\sigma_\rho$) in a given field of view using the following expression: $\sigma_\rho = -\frac{1}{\bar{A}}\sigma_A$.

To define the border between the bulk and the edge regions in the cell cluster, we calculate the average cell density at the center of the cell cluster (circle with radius of 1mm). The mean density in the inner circle we define as the reference cell density (see dashed lines in Figures 3.1 C,D). For every image (434 \(\mu\text{m} \times 325 \mu\text{m} \text{wide region}) we calculate the average cell density ($\bar{\rho}$) and the standard deviation of the cell density ($\sigma_\rho$). If the reference density is in the $\bar{\rho} \pm \sigma_\rho$ range the region is associated with the bulk of the cluster.

3.3.2 Results

With this seeding procedure cells in a cluster need about one day to form a monolayer, and from that moment on the cluster gradually grows and densifies (Figure 3.1 A). Once the bulk and the edge regions are formed, clusters continue to grow preserving this morphology.

Therefore, we distinguish between two phases in classical growth: the maturation phase and the aging phase (Figure 3.1 A, B). The maturation phase is characterized by densification of radially shaped uniform cell clusters (Figure 3.1 B, C) that will eventually compartmentalize into the bulk and the edge regions (Figure 3.1 C). In the bulk region, cells reach steady state density and become immobile. In the edge region the cell density gradient vector points towards the bulk region, and cells move and proliferate. As the cell cluster grows, both regions increase in size (Figure 3.1 D).

Usually, four days after seeding, cell density in the bulk reaches steady state or, in some cases, it is very close to steady state. Once reached, steady
Figure 3.1: Characterization of the classical regime of growth on the glass substrates. (A) Schematic representation of the cluster growth over time. (B) Average cell density in the bulk region and the edge region during the maturation and the aging phase. (C,D) Density profile throughout the cell clusters during the maturation phase (2 and 4 days after seeding) (C) and the aging phase (4, 8, and 12 days after seeding) (D). Error bars represent the standard deviation of the cell density in that field of view.
state density is constant over time (Figure 3.1 D) while the cluster grows from $151 \pm 35 \text{ mm}^2$ on day four up to $456 \pm 25 \text{ mm}^2$ on the twelfth day. The bulk and edge regions stay well separated as long as the cluster has space to grow. The growth of the clusters with well separated bulk and edge regions, we term the “aging” phase. This phase is also associated with various irregularities in the tissue organization. These are explained in detail in Subsection 3.5.2. Although the details of those phases differ for glass substrate and stiff gels they nevertheless share general features.

While classical regime of growth occurred regularly on the substrates stiffer than 3 kPa polyacry-lamide (PA) gels, on the gels with a stiffness of $0.6 \pm 0.2$ kPa a different regime of growth was observed. In the following section we describe the difference between the classical and this different regime of MDCK-II growth.

### 3.4 Novel growth regime of MDCK II cells on soft substrates

The following section has already been published as an open access article in the Biophysical Journal (Volume: 106, Pages: L25-L28) in 2014 under the title: “Novel Growth Regime of MDCK II Model Tissues on Soft Substrates” [89].

#### Image analysis

The results of the experiments described in 3.2 were analyzed in MATLAB and ImageJ. The density and the monolayer area in small clusters were determined using ImageJ (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/, 1997-2017). Here, the number of cell nuclei was counted within the area of interest, this area being determined using a freehand selection tool (Figure 3.2). The deviation arising from repeatedly selecting the same area is less than 1%. The main contribution to the reported deviation comes from the variation in cell density when changing the area of interest. Examples of areas of interest are shown in different colors in Figure 3.2. Nevertheless, the relative uncertainty in density remains below 10% and naturally becomes smaller as the size of the monolayer increases. For example, by choosing different areas of interest the mean cell density of the colony shown in Figure 3.2 varies maximally by 6% with respect to the mean cell density in the area bordered by the blue line.
For larger clusters the number of cells in a given image was calculated using the MATLAB routine described in Section 2.1.3 (for error estimation of this method see Table 5.19). The area of the colony was calculated after segmenting the actin stained images using the Otsu’s method [129]. For very large colonies a collage of images was made prior to segmentation. In total, 130 colonies were analyzed with sizes differing over five orders of magnitude.

In the case of the single cells, the cell area was determined after segmentation of the actin filaments. The error in threshold procedure is negligible compared to the variation of the cell areas. Mean single cell area was determined from 50 cells in the case of glass substrates and from 90 cells in the case of soft substrates (0.6 kPa PA gels).

### 3.4.1 Results

On glass, colonies start as small clusters with very low cell density of $700 \pm 200 \text{ cells/mm}^2$ (Figure 3.3, A, B), typically surrounded by a strong actin cable (Figure 3.3, B, C). Spread area of the single cells (Figure 3.3 A) on glass was found to be $(2.0 \pm 0.9) \times 10^{-3} \text{ mm}^2$.

In colonies grown on 0.6 kPa gels, however, we encounter a very different growth scenario. The average spreading area of single cells is $(0.34 \pm 0.3) \times 10^{-3} \text{ mm}^2$, which is six times smaller than on glass substrates (Figure 3.3 C). Clusters of only a few cells show that cells have a prefer-
Figure 3.3: Early phase of cluster growth on hard (A-C) and soft (D-F) substrates. (A) Well spread single cells, and small clusters with a visible actin cable 6 hours after seeding. (B) Within one day, clusters densify and merge together making small colonies, the edge of which is shown in (C). (D) 12 hours after seeding, single cells on 0.6 kPa remain mostly round and small. They are found as individuals, or within small, 3D structures (top). The latter nucleate a monolayer in their center (bottom), if the contact area with the substrate exceeds about $5 \times 10^{-3}$ mm$^2$. (E) Irregularly shaped clusters appear due to the merging of smaller droplets on soft substrates. A stable monolayer surrounded by a 3D belt of densely packed cells is clearly visible, also in larger structures (F).
Figure 3.4: Monolayer densities in colonies grown on 0.6 kPa substrates, as a function of the cluster size and age. Each cluster is represented by a single data point signifying its mean monolayer density. The bulk and edge steady state densities of monolayers grown on glass substrates are shown as black and red dashed lines. Error bars are omitted for clarity.

ence for cell-cell contacts rather than for cell-substrate contacts. The same conclusion emerges from the fact that droplet-like agglomerates resting on the substrate form spontaneously (Figure 3.3 C) and that attempts to seed one single cluster of 90,000 cells fail, resulting in a number of three-dimensional colonies. When the contact area with the substrate exceeds $4.7 \times 10^{-3}$ mm$^2$, a monolayer appears in the center of such colonies (Figure 3.3 D). The colonies can merge and, if individual colonies are small, the collapse into a single domain is associated with the formation of transient irregular structures (Figure 3.3 D). Ultimately, large elliptical colonies (average ratio of major to minor axis of $e = 1.8 \pm 0.6$) with a smooth edge are formed (Figure 3.3 E), unlike on hard substrates where circular clusters ($e = 1.06 \pm 0.06$) with a ragged edge are the characteristic phenotype.

Irrespective of the size of the cluster, in the new regime of growth, the internal structure is made up of two parts (Figure 3.3 D). The first compartment is the edge ($0.019 \pm 0.05$ mm wide), a three-dimensional structure of densely packed cells. This belt is a signature of the new regime since on hard substrates the edge is strictly two-dimensional (Figure 3.3 C). The second compartment is the centrally placed monolayer with a spatially constant density that is very weakly dependent on cluster size and age.
Figure 3.5: Hoechst stained cell nuclei in monolayers grown on relatively soft substrates. (A) Novel regime of growth on 0.6 kPa PA gels. (B) Intermediate regime of growth on 1.2 kPa PA gels.

(Figure 3.4). The mean monolayer density is 13,000 ± 2,000 cells/mm², which is an average over 130 clusters that are up to 12 days old and have a size in the range of $10^{-3}$ to 10 mm², each shown by a data point in Figure 3.4 (for details see Section 3.4). This density is twice the steady state density of the classical regime of growth that is reached in the bulk of the clusters.

**Intermediate regime of growth** While on soft substrates (0.6 ± 0.2 kPa) the novel regime of growth (Figure 3.8 F) is present, on the substrates harder than 3 kPa we exclusively observe the classical regime of growth. However, the transition from novel to classical regime happens smoothly. On the substrates with intermediate rigidity (1.2 ± 0.2 kPa) morphology of cells is characteristic for both classical and for the novel growth regime (Figure 3.4.1 B). The two regimes coexist in the neighboring cluster, and sometimes both regimes are present even within the same cluster. The coexistence is seen even six days after seeding. On the soft substrates (0.6 kPa) some cell clusters started the growth in the classical regime. However, this was a very rare event and it was transient since no clusters with the morphology of the classical regime of growth could be found from the third day on while on 1.2 kPa, co-existence of the two regimes persisted over time.
3.5 Mechanosensitivity on hard substrates: Characterization of monolayer growth, cell organization, proliferation, steady state and irregularities

To explore the subtle differences between different substrates, we seeded circular cell clusters on PA gels with elasticity ranging from 1.2 to 34 kPa as well as glass substrates (for details see 3.2). On 1.2 kPa the intermediate regime of growth is observed with characteristics of both the novel and the classical regime coexisting over long time periods.

Hard gels (>3 kPa) and glass substrates induce the classical regime of growth characterized by growing and densifying monolayer. We examined the mechano-sensitivity of cells during the classical regime of growth during the time period of several days. The overall area of cell colonies on various substrates was measured over time. Within the colonies we measured the distribution of cell density and the fraction of cells in the S-phase of cell division at various time points of development as a function of cell area. Furthermore, we quantified and explained the irregularities in the regular cell monolayer structure that emerged over time and examined the activation of paxillin and actin on the level of cell cluster as well as at cell level. We found that, with regard to all the quantified features, cell clusters on hard gels as substrates differ from those on glass substrates.

3.5.1 Image analysis

Quantifying cell density distribution

The grids of 5x images were stitched using the ImageJ (Rasband, W.S., ImageJ, U.S. National Institutes of Health, Bethesda, Maryland, USA) plugin based on the publication by Preibisch at al. [30]. Cell density distribution was estimated from the stitched images of the Hoechst stained cell nuclei. In the circular rings 51 µm wide, and centered in the center of the cluster the mean intensity was measured (Figure 3.6 A). Subsequently, the linear mapping of the mean intensity and the mean cell density was done for each cluster individually (Figure 3.6 B). For each cluster at least 20 different segments with an area of at least 0.5 mm² were used. The use of linear mapping is justified by the relatively constant total intensity which is illuminated by each cell nuclei.
Figure 3.6: Evaluating cell density from average intensity. (A) Cell cluster is divided into equidistant rings. The central region is a circle with a radius of 350 µm. (B) Image segments taken from different regions in a cell cluster are analyzed and a linear mapping between average cell density and image intensity is found for each cluster.

Quantifying steady state and maximal cell density

Steady state densities were estimated using our MATLAB routine (see Subsection 2.1.3) from a sequence of images taken with 20x magnification (800.8 µm x 575.6 µm and pixels size of 0.31 µm). The sequence was taken through the middle of the cell cluster from one edge to the other. For this purpose we measured the cell density in clusters on 5, 11 and 20 kPa gels (one cluster per substrate). All clusters were six days old. On the glass substrates we analyzed two clusters. The steady state density is estimated as the mean value of the density in the bulk of the cluster. The bulk region is defined as all 267 µm x 575.6 µm regions which are in the $\bar{\rho}_{\text{central}} \pm 2 \cdot \sigma_{\rho}$ range. Here $\bar{\rho}_{\text{central}}$ denotes the mean cell density in the central region (2 mm radius around cluster center) and $\sigma_{\rho}$ is the standard deviation of cell density between different 267 µm x 575.6 µm fields of view in the central region. The high density ring is defined as all 267 µm x 575.6 µm regions which have cell density higher than $\bar{\rho}_{\text{central}} + 2 \cdot \sigma_{\rho}$. Standard deviation is the deviation of the mean density of different regions between different clusters. On soft substrates the standard deviation, likewise, denotes the deviation of the mean cell density within 19 cell clusters after six days.
Quantifying cell proliferation

We investigated cell proliferation on various substrates with 5-ethynyl-2’-deoxyuridine (EdU) which is incorporated in DNA during replication instead of thymidine. This, therefore, marks the S-phase of the cell proliferation cycle. Since the extent of the EdU incorporation will depend on whether the cell has just entered the S phase or is at the end of it, the brightness of cell nuclei is highly heterogeneous (Figure 3.7 B).

To identify proliferation nuclei firstly we segmented the Hoechst stained cell nuclei (Figure 3.7 A). To do so, we used the watershed transformation based segmentation method, which is relatively fast. We did not use our standard nuclei segmentation method explained in the Subsection 2.1.3 due to the large size and low resolution of the images (pixel size is 1.23 µm). Used method is a standard segmentation method based on thresholding the image, followed by the watershed algorithm being applied to the distance-transformed binary image. The key step in this procedure is choosing the correct threshold level.

Our images are 3.2 x 2.7 mm, which often makes a choice of a single threshold level impractical. To avoid this problem and to lower the segmentation errors, we cut the original image into 108 segments with an area of 0.0653 mm² (271 x 241 µm) after 5% of the image at the edges had been removed due to the overlap between the subsequent images. On each image segment a local threshold was applied with the threshold value calculated using the Otsu’s method. Cell density is calculated as the number of the segmented cell nuclei in the 271 x 241 µm rectangular. Typically, one image segment has between 130 and 480 cells depending on the cell density.

To calculate the number of the cells in the S-phase of the cell division, we calculated the average intensity of the pixels at the position of a segmented cell nuclei. If the average intensity of the EdU signal at the position of a given cell nuclei is brighter than the 95% of the average intensity of the EdU signal in that image segment, we consider that this cell is undergoing the S-phase of the cell division (Figure 3.7 C,D). This value was chosen arbitrarily upon visual inspection. Choosing a different threshold would result in a systematic shift of the proliferation curve toward lower or higher values. The total number of the proliferating cells divided by the total number of the segmented cell nuclei in that image segment gives us the fraction of the proliferating cells in that segment.

We imposed the set of criteria for accepting the image segment in the final statistics. Firstly, EdU intensity should not be too bright (due to occasional circular bright artifacts in the EdU channel) or too low (outside of
Figure 3.7: Examples of cell nuclei images (A-B) and the results of segmentation process (C-D) (A) Hoechst stained cell nuclei have uniform illumination while EdU stained cell nuclei (B) in the same field of view have very different brightness. (C) EdU stained cell nuclei are automatically recognized. Cells counted as proliferating are marked with red dots. (D) Image of the cluster edge. Image is separated in segments. The segments which satisfied our five exclusion criteria are marked with red dots.
the cell cluster). Even though, the same staining procedure was applied to
every cluster, each cluster had different absolute brightness due to differ-
ent bleaching conditions. Therefore, the limit for too bright and too low
was determined individually for each cluster after visual inspection. Fur-
thermore, the number of cell nuclei should not be too low (minimal cell
density of the cell monolayer is set at 1,500 cells/mm²), and the value of
the Otsu’s threshold level should not be too low (outside of the cluster).
Finally, image segment should not contain objects which are too large to
be a cell nuclei (>315 µm²). When all five criteria are satisfied the image
segment is accepted (see Figure 3.7 D).

Per cell cluster, between 500 and 2,000 image segments were accepted.
We analyzed seven clusters on 11kPa PA gels (1 cluster two days old, 1
three days old, 3 four days old and 2 six days after seeding), and on glass
substrates (1 cluster two days old, 2 three days old, 3 four days old and 1
six days after seeding). This gave us, in total, a several thousand accepted
image segments per substrate (4,100 on gels and 5,900 on glass) with each
image segment having several hundreds of cells.

One possible systematic error which could arise from this method is
overestimating the number of cells due to the edge effect (we counted the
number of objects in the box) and due to over-segmentation, to which this
segmentation method is prone (for data see Table 5.18). However, due to
the relatively large size of the image segments and only a small percentage
of over-segmented cell nuclei, this error is not significant which is evident
from correct estimation of the bulk density.

3.5.2 Results

Quantifying the morphology of the cell clusters

Visual comparison During the first four days, samples seeded on the
glass substrates looked very similar to the samples seeded on the hard
gels and the differences became obvious only six days after seeding.

Two days after seeding cells were organized in relatively uniform and
sparse colonies (Figure 3.8 A, B) on the hard gels and on the glass sub-
strates. Four days after seeding, clusters on both substrates compartmen-
talized into a contact inhibited very dense bulk region, and a low density
edge region (Figure 3.8 D), and those regions remained well separated.
However, during the later cluster development visual differences between
cell clusters grown on those two substrates became much more obvious
(Figure 3.8 A,B). Six days after seeding, both the bulk and the edge re-
gions differed visually on the glass substrates compared to those on the
hard gels. However, no significant difference could be seen nor quantified for the various hard gels (3-34 kPa) which all had the same steady state density, density distribution within the cluster and also the cluster size.

**Distributions of the cell density within the clusters and their areas**  As the visual comparison suggested, the cell density distributions throughout clusters grown for two days on glass and PA gel substrates are very similar, with relatively constant cell density (Figure 3.9 A). Four days after seeding clusters have formed the bulk and the edge regions but their density distribution profile reveals certain differences. Four days after seeding, the mean cell density in the bulk region often varies between clusters. The reason most probably lies in the small differences between the initial seeding conditions (how the droplet of cells is placed on the substrate and how wet the substrate was). For this reason, not all clusters reached the steady state after four days. We did, however plot two clusters that had reached the steady state four days after seeding.

Six days after seeding the small differences between clusters on different substrates observed on the fourth day become more pronounced (Figure 3.9 B). On the hard gels, once the bulk and the edge regions are formed, the belt of very high cell density starts to form which results in the M-shaped density profile typically observed on day six. Cell density in this belt is significantly larger than the steady state density. Cell density in the belt locally reaches as much as 11,000 cells/mm\(^2\). However, the mean value (averaged over the clusters on 5-20 kPa) of the maximal cell density is 8,500 ± 500 cells/mm\(^2\).

On the hard gels, the edge region has stripes of higher or lower cell density at the same radial distance from the center of the cluster (Figure 3.8 A,C). Those stripes are typically 3-7 cell thick lines which are unevenly distributed. On the glass substrates those stripes are not so well pronounced (Figure 3.8 D). On the other hand, on the glass substrates we observed a high density ring close to the very end of the cluster (typically 150-700 µm away) (Figure 3.8 B,D).

Whether the substrate is glass or hard PA gel will influence not only the distribution of the cell density but also the steady state density (figure 3.9 D). On the hard gels, steady state density on five different clusters was 7210 ± 140 cells/mm\(^2\) while on glass substrates this density on both analyzed clusters was 7040 ± 50 cells/mm\(^2\). This difference is accompanied by a significant difference in the monolayer organization due to the dome formations (see Subsection 3.5.2). Moreover, on the glass substrates cell proliferation in never completely arrested and cell extrusion events are
Figure 3.8: Visual differences between MDCK-II monolayers in the classical regime of growth. (A,B) Sections of the actin stained clusters at various time points. Clusters were seeded on hard gels (11 kPa) (A) and on a glass substrate (B). (C,D) Cell nuclei in the edge region of cell clusters 6 days after being seeded on the hard gel (C) and on the glass substrate (D).
Figure 3.9: Cell density and the cluster area during the development of the colonies on hard gels and on glass substrates. (A-B) Distribution of the cell density within the colonies two days, four days (A) and six (B) days after seeding on hard gels and on glass substrates. (C) Steady state and maximal cell density on soft gels, on hard gels and on glass substrates. On glass and soft gels steady state density is the maximal density. The dashed line denotes the steady state density on hard gels which is greater than the steady state density on glass substrates. (D) Mean area and the standard deviation of the cluster area 2, 4 and 6 days after seeding on hard gels and on glass substrates.
much more frequent.

The clusters area is also larger on hard gels in comparison to glass substrates but this difference can be observed only after the bulk and the edge regions have been formed (Figure 3.8 E). Six days after seeding on the hard gels, average cluster size is $328 \pm 15 \text{ mm}^2$ (five clusters in statistics) while on glass substrates it is $276 \pm 8 \text{ mm}^2$ (two clusters in statistics).

**Paxillin activation on the cell and the cell cluster level** To further understand cells’ mechano sensitivity, we stained for paxillin, a protein which accumulates at the focal adhesion sites. On the level of the whole cluster, we see that actin activation (Figure 3.8 C,D) is accompanied with paxillin activation (Figure 3.10 A). Confocal microscopy revealed a clear difference in the focal adhesion sites on the glass substrates and hard gels (Figure 3.10 C,D). It is only the glass substrates that induce well localized focal adhesion sites connected with strong actin cables (Figure 3.10 C,D). Focal adhesion sites on both soft and hard gels look similarly dispersed and are not associated with strong actin cables. This is a surprising result considering that cells grow very similarly on hard gels and on glass substrates, while on soft gels they display a novel regime of growth.

**Cell Proliferation**

If we simply look at the distributions of the cell divisions (Figure 3.12) we immediately see that in young cell clusters (second day) seeded on the hard gels and on the glass substrates cells are proliferating throughout the cluster. On the fourth day, cells proliferate predominantly in the edge region. The difference which could be seen between the cluster on the glass and on 11 kPa is due to the difference in the bulk density between those two clusters. The presented cluster on the glass has a mean cell density in the bulk region of $5,700 \text{ cells/mm}^2$. This example demonstrates the distribution of the cell proliferation just before the contact inhibition in the bulk is reached. On the other hand, on 11 kPa the cluster has reached the contact inhibition and has a density of $7,100 \text{ cells/mm}^2$. Finally, six days after seeding, all clusters have reached steady state cell density in the bulk, and only the edge region is proliferating and moving. While on the glass substrates sporadic cell proliferations can be observed in the bulk region, on the hard gels this region is completely contact inhibited. Therefore, we can by visual impression conclude that the number of the proliferating cells depends on the local cell density.

Quantifying the fraction of cells in the S-phase of cell division as a function of the average cell density in that image segment reveals that local cell
Figure 3.10: Paxillin activation in cells on the different substrates. (A) Paxillin stained cell clusters on hard gels and on glass substrates fixed at different time points. (B) Paxillin at the edges of the cell clusters on hard gels and on glass substrates. (C,D) Confocal images of paxillin (green) and actin (red) activation at the position of the focal adhesion sites.
density is the main parameter determining the percentage of the proliferating cells (Figure 3.12). The fraction of the cells in the S-phase averaged over different cell density classes on the glass (Figure 3.12 A) and on 11 kPa PA gel substrates (Figure 3.12 B) does not reveal a significant difference between different time points.

Plotting all data points from the clusters on the glass substrates and on the 11 kPa gels shows a significant difference between the glass substrates and the hard Pa gels (Figure 3.12 C). The average fractions of EdU positive cells as a function of the average cell size in that cell density class seems to be well fitted with a quadratic function on the glass substrate. Data for the cells on the PA gels could be fitted with a similar quadratic function only for relatively larger cells. For relatively small cells, proliferation is inhibited in the same manner on both substrates. This indicates that sporadic cell divisions in the bulk region on the glass substrates are a consequence of the local cell density, which does not lead to complete inhibition of the cell proliferation. However, this state is the steady state because the cell density does not increase in time due to more frequent cell extrusions on the glass substrates. For a cell density of less than approximately 5,700 cells/mm² the probability of cell division is significantly greater on the hard gels than on the glass substrates. Larger cells and particularly intermediate size cells, divide more on the hard gels than on the
Figure 3.12: Quantification of the cell proliferation. (A-B) The mean and the standard deviation of the fraction of cells in the S-phase of cell cycle in the different cell density categories. Data was plotted independently for each day for the clusters grown on the glass (A) and for those on the 11 kPa gels (B). The fraction of cells in the S-phase as a function of the mean cell density in a given image segment. Red dots represent the data from 11kPa gels and the blue dots are from the glass substrates. The intensity of the color is related to the age of the cluster from which the image segment is taken.
We know that on the soft substrates small droplets of cells grow and, once they reach the critical size, they start forming the dense monolayer. We know, then, that cell divisions must happen in the crowded environment of small droplets where cells are densely packed on top of each other. Surprisingly, even for large cell clusters which have formed a monolayer, cells still predominately proliferate inside the 3D edge. This indicates that the growth of the colonies in the novel regime arises not only from the cell divisions in the monolayer but also from the divisions in the 3D edge (Figure 3.13 A).

The intermediate regime of growth combines characteristics of the novel and the classical regime of growth within the same cluster. We see that on the left-handside of the cluster in the Figure 3.13 B cells are dividing in a similar way to the cells in the classical regime of growth, while on the right-hand side of the cluster proliferation is mostly arrested, with most of the cells dividing at the 3D edge.

**Irregularities**

The aging phase is characterized by degradation of tissue structure in many different ways. Various deformations emerge over time on a cellular level (nuclei shape deformation and degradation of actin filaments) as well as on a multicellular level (domes, cell extrusions and “lacunas”). The consequence of all those phenomena is a slight decrease in the cell density and the order in the tissue.
Figure 3.14: Domes. Fluorescence microscopy images of the Hoechst stained cell nuclei in the bulk regions on (A) glass substrate and (B) hard PA gels. Three-dimensional reconstruction of the confocal images shows that those structures are indeed blisters lifted away from the basal side of the monolayer. (D) Once they start forming (three days with our seeding conditions), the number of domes does change significantly over the time. (E) The size of the domes reaches maximum six days after seeding.

Domes  The domes are rarely reported in current studies but many older papers reported and studied them. They are associated with an MDCK cell line cultured on impermeable substrates such as glass [27], [131]. MDCK cells originate from the distal renal tube of a kidney [27], [26] whose biological function is to transport sodium ions from the apical side to the basal side. Major regulator of salt and water reabsorption in a number of epithelial tissues, and as well in MDCK cells, is the epithelial Na⁺ channel (ENaC) [132]. When those cells are grown on impermeable substrates, segregation of ions is followed by diffusion of water, which leads to blister-like structures termed domes. In some MDCK clones, tubule-like structures are formed under the same conditions [133].

Under our experimental conditions, we observed domes on the glass substrates starting from the third day (Figure 3.14). They increased in size up to sixth day (Figure 3.14). They are usually between 0.01 and 0.04 mm high. However, we noticed that the size and the number of domes depends greatly on the cell medium used during the experiment.

The regular, almost hexagonal organization of the cells in the tissue is destroyed by such structures and they make even simple morphological analysis, such as the mean cell density, much more difficult. While they
are still present on the PDMS gels, we did not observe them on PA gels of either low (0.6 kPa) or high rigidity (3-34 kPa). We therefore suggest using PA gels as the substrate of choice in the studies of the MDCK monolayer regarding the cell organization and morphology within the epithelium.

**Cell extrusions on soft substrates** On the soft substrates until the fourth day the monolayer is very homogeneous, showing a close to hexagonal arrangement of cells. From day four, however, defects start to appear in the form of the small holes (typical size $(0.3 \pm 0.1) \times 10^{-3}$ mm$^2$). These could be attributed to the extrusions of viable cells, either from the belt, or from areas of increased local density in the monolayer (inset in Figure 3.17). This suggests that extrusions serve to release stress built in the tissue, and, as a consequence, the overall density is decreased.

Previous reports suggest that isolated MDCK cells undergo anoikis eight hours after losing contact with their neighbors [134]. However, in our experiment, it appears that, instead of dying, the extruded cells create new colonies, seen as an archipelago surrounding the mother cluster (Figure 3.17). The viability of the cells cast off is most strongly evidenced by the appearance of single cells and very small second generation colonies surrounding the big clusters as long as twelve days after seeding. Inter-

Figure 3.15: Cell extrusions on soft substrates. (A) Cell nuclei within the mother colony and in the neighboring archipelago of second generation clusters grown on 0.6 kPa gels on day twelve. The inset shows a scar in the tissue that is a result of a cell extrusion event. From the image of the cell nuclei (left) it is clear that there are no cells within the scar, while the image of the actin (right) shows that the cytoplasm of the cells at the edge has closed the hole. (B) Confocal microscopy of a cluster showing how neighboring cells enclose the hole left after cell extrusion. Cell nuclei are shown in blue, and actin filaments in red.
Figure 3.16: Three examples of the cell extrusions found on the glass substrate. (A) Cell nuclei image. (B) F-actin fibers. (C) Confocal images of the cell nuclei and actin ordered from the lower to the upper part of the monolayer.

Interestingly, no morphological differences were found in the first and second generation colonies. The ability to extrude viable cells may point to a new migratory pathway regulated mechanically by the stresses in the tissue. Live cell extrusions in very dense MDCK II monolayers have been reported [135], as one of the mechanisms that prevent overcrowding in the tissues. These cell extrusion are very similar to the extrusions of the apoptotic cells, but the signaling pathway is activated by stretch activated channels.

“Lacunas” Another phenomenon that we observe on glass substrates, after the steady state density is reached, is the appearance of holes in the cell structure. While on soft substrates these formations are quickly enclosed by the neighboring cells, on the glass substrates the voids occur frequently (see Figure 3.16 B). We use the term “lacunas” for these formations. The “lacunas” don’t exceed 600 $\mu$m$^2$ in area and are found all across the bulk of clusters. If found, on hard gels they are most commonly present in the transition zone between the bulk and the edge of clusters. However, they are not always present in the clusters grown on hard gels and cannot be considered as typical phenomenon.
in most of the cases those formations are remnants from the cell extrusion events. In some cases they could also be small domes that are just forming. As seen in the Figure 3.16 example 2, it is hard to tell whether the big dark spot is a forming dome or the cell with the half-moon shaped nuclei is being extruded. On the other hand, the small dark spot is obviously a single apoptotic cell being extruded. However, non apoptotic cells that are extruded are often observed in pairs and have half-moon shaped nuclei, as evidenced in examples 1 and 3. It remains to be explained why do we observe much more cell extrusions on the glass substrates than on hard gels.

Degradation of the actin filaments and nuclei shape and density inhomogeneities During the aging phase, the morphology of cell nuclei becomes irregular, often taking on bean-like shapes. On soft substrates, triangular nuclear shapes can even be found. The change in nuclei shape is accompanied by the blurry appearance of actin fibers. We observe the changes on both the glass substrates and on the soft and the hard PA gels.

It is known that MDCK cells synthesize laminin [136]. However, which laminin is synthesized and how much, is substrate dependent [137]. Moreover, it was shown that laminin (LN) 5 is synthesized only by subconfluent MDCK cells, while confluent layers degrade the existing LN5 matrix [138]. Changes in the ECM rigidity induce changes in the concentration of the lamin-A which in return changes the stiffness of the cell nuclei [139]. Alongside the rigidity sensing mechanism, cells feel differences in the chemical composition of the ECM. They do this by recruiting different integrins at the focal adhesion sites [140]. It is, therefore, very likely that
Figure 3.18: Time induced nuclei-shape changes in the cells seeded on the collagen-I freshly deposited on the substrate and the collagen-I that has been in the cell medium for eight days.

the MDCK cells induce changes in the extracellular matrix (ECM) themselves, and those changes are responsible for bean-like nuclei shapes and the degradation of the actin fibers.

However, the changes in the ECM are not limited to the in-vitro conditions. In vivo changes in the cell environment can emerge either from tissue stiffening, or from changes in the ECM due to the biochemical identity of the proteins or due to mechanical changes induced by cross-linking of, for example, collagen fibers, which makes the ECM effectively stiffer [141]. This stiffening of the ECM induces aging related changes in the epithelium such as skin aging [142].

To test whether the morphological changes of the cells are induced by spontaneous cross linking of the collagen-I, we seeded cells on the collagen-I coated glass substrates that had been in the cell incubator for
eight days, and the cell medium was changed every second day just as in the case of the cell culture. Cells seeded on eight-day old collagen-I substrates manifested aging related morphological changes after six days, just like the cells seeded on the fresh collagen-I (Figure 3.18). This implies that possible spontaneous cross-linking of the collagen-I fibers in the cell culture medium is not the cause of the morphological changes in the cell nuclei and actin fibers.

We then checked whether these changes are produced by the environment or whether the cells demonstrate the deformed morphology even after being reseeded on the new substrates. Again, the cells revealed morphological changes after the steady state density was achieved in the bulk of the cluster. We conclude that these changes must be a result of either cell induced changes in the composition of the ECM or of the pressure that cells exert on each other. However, we cannot say which of these two possible scenarios is the case.

**Loss of the tissue structure-RDF** Time affects not only the shapes of the cell nuclei and the appearance of actin fibers, it also affects the organization of cells in the monolayer. During the aging phase, local density inhomogeneities become more and more pronounced. This phenomenon occurs on all substrates (Figure 3.19). Regions with very densely packed cells can be observed next to the regions with relatively large cells. This is particularly interesting since the organization in the bulk region is stationary once the steady state density is reached.

![Figure 3.19](image)

Figure 3.19: Density inhomogenities emerge in a monolayer over time. (A-C) Cell monolayer stained for cell nuclei on glass substrates (A) and hard (B) and soft (C) PA gels after 8 days.

Radial distribution function (RDF) reveals decreased cell organization in the tissue over time. We normalized the RDF function with the average cell density within a radius of interest. The results were averaged for several thousand cells. During the maturation phase, due to the densification, tissue increases static correlations which can be seen in the narrowing of
the RDF peak (Figure 3.20). This is observed in the bulk (Figure 3.20 A) as well as at the edge of the clusters (Figure 3.20 B). After the steady state density is reached in the bulk, an aging phase begins which is associated with the degradation of the tissue structure (Figure 3.1 C,D) in both the bulk and the edge regions of the cluster.

Figure 3.20: RDF function during maturation and aging phase of the monolayer in (A) the bulk region and (B) the edge region.

### 3.6 Discussion and Conclusions

We have proven, contrary to the previous reports [40], [126], [127] that the monolayer of the MDCK-II cells, which is a standard system model of an epithelium, is indeed mechano-sensitive. The effect of mechano-sensitivity is most pronounced on extremely compliant (Young modulus of 0.6 kPa) and extremely hard substrates (namely glass). Importantly, we show that the cell density at which cells find the steady state depends on the rigidity of the underlying substrate. This means that a tissue steady state depends on physical properties of its surrounding.

On the soft substrates ($E = 0.6 \text{ kPa}$) we demonstrated a novel phase of growth of MDCK II model tissue in which cells organize in elongated clusters consisting of the very dense monolayer (constant density of $13,000 \pm 2,000$ cells/mm$^2$) surrounded with the 3D edge of still proliferating cells. Despite previous similar efforts, this regime of growth had not previously been observed. This finding is especially interesting in the context of the elasticity of real kidneys, for which the Young modulus has been found to be in the range of the elasticities studies herein [143]. Furthermore, the viscoelastic properties of the confluent MDCK II monolayer have recently been measured [144] using AFM-based oscillatory microrheology (OMR) and force cycle experiments (FCE). These methods have shown the apparent Young modulus of $0.85 \pm 0.1$ (for indentation depth $h_0 < 1 \mu\text{m}$) and $0.575 \pm 0.05$ respectively. This suggests that “soft” for MDCK II cells
means softer than they are, while “hard” substrates are those stiffer than the cells themselves. However, substrate rigidity can also influence the rigidity of the cells [81] and therefore more systematic measurements are still required in order to confirm this hypothesis.

Interestingly, the study on MDCK cells in a circular spatial confinement [61], reported the emergence of a 3D edge on the glass substrates after the cell density exceeded 7,000 cells/mm². The authors explain the mechanism of the edge formation thus: if the cell division happens near the edge, new cells position themselves on top of the cells at the edge (asymmetric forces form the neighboring cells). If, however, division happens in the bulk cells are extruded from the cell monolayer (symmetric forces form the neighboring cells). Despite the visual similarities, the emergence of the 3D edge in the novel regime of growth must come from a different mechanism, since the 3D structure is the initial state of the cell clusters and the monolayer forms in the middle when the critical size of the droplet is reached. The formation of the 3D edge must be a consequence of cells inability to spread and invade new territories when they are grown on extremely compliant substrates.

In contrast, the classical regime of growth is observed on the hard substrates (>3 kPa) where cells form slowly-densifying monolayers which spread more readily and have large and thin cells at the edge. Despite the numerous studies on the classical regime of growth, the morphological characteristics of the colony growth on the large time scales is still lacking in the literature. Our study, then, contributes to this field, reporting the precise cell density distribution over long time periods, the curve for the probability of the cell division as a function of the cell size on large cell statistic, the number and the size of the domes on the glass substrates, the cell extrusions, the cell-nuclei deformations and the density inhomogeneities. The most characteristic feature of the classical regime of growth is spontaneous separation into a contact inhibited bulk region, where cells have reached the steady state density, and a dynamic, still proliferating edge region, where cell density gradually decreases towards the very edge of the cluster.

Cell proliferation does not change significantly over time. On the both substrates (glass and hard gels) the proliferation curve can be determined. This curve can be approximated using a single quadratic function on the glass substrates. On the hard gels, in contrast, the curve follows a quadratic function only for relatively larger cells. However, both curves are convex and monotonic. This result is contrary to that of the previous study, which reported a concave shape of this function for MDCK-II cells on PDMS gels [43]. The reason for this discrepancy could be the different methods. While
we used EdU, and so measured the number of cells in the S-phase of the cell division, the other study used Fucci cell cycle markers, which marks the cells undergoing S, G2 or M phases of the cell division in different colors. Another reason could be the lack of statistics, which in our case was avoided by including several hundreds of thousands of cells in the measurement.

Increased cell proliferation on the hard gels is not unexpected given the larger overall cluster area and the larger cell densities on those substrates. However, the difference becomes obvious only after the formation of the bulk and the edge regions (six days after seeding, in our case). Alongside larger cell density and larger cluster sizes, hard gels also have different density distributions within the cluster. While on glass substrates the bulk region smoothly converts into the edge region, on the hard gels they are separated by a high density ring region where cells are packed more densely than in the steady state.

We also analyze the morphological changes of the cells over the time. On all analyzed substrates, cell-nuclei shape changes, within six days, from elliptical to asymmetric shapes: there are most usually bean-like but sometimes triangular. Change in the cell nuclei shape on both soft and hard gels is accompanied by a blurry appearance of the actin fibers. On the glass substrates, some cells still have strong basal actin fibers. The changes described arise either from pressure that cells “feel” from their neighbors or from the changes in the composition of the ECM. However, further experiments are needed to distinguish between those two scenarios. Due to this problem, which is always associated with in-vitro experimental conditions, we do not recommend experiments with MDCK cells which have been cultivated for longer than six days.

In addition to the morphological changes in cells, the organization of the cell monolayer also changes as the density approaches the steady state. On all analyzed substrates, the inhomogeneity of the cell density becomes more pronounced over time. On the glass substrates, large and numerous domes start to appear. The same formations are observed on PDMS gels, while PA gels seem to be porous enough not to facilitate the formation of these blisters. The domes destroy the monolayer organization to a great extent, and so we recommend the utilization of the hard PA gels for all studies dealing with the morphology of the MDCK monolayer.

The study of substrate-dependent reactions of MDCK cells to TGF−β1 growth factor [128] showed that the cells organized in the small droplets on soft substrates undergo apoptosis, while the cells on hard substrates (>5 kPa) undergo epithelial-mesenchymal transition. This study proves that morphological changes on different substrates are accompanied by
complex changes in biochemical signaling and possibly gene expression. This suggests that cell mechanosensitivity is a complex process which occurs even in fully differentiated cells which still react to purely mechanical stimuli and adjust their proliferation and organization accordingly.
4.1 Introduction

The initial step in monitoring collective cell behavior is cell tracking. Phase-contrast imaging enables chemically nontoxic monitoring of the cell monolayer (photo-toxicity is still present). However, phase-contrast images are not easily segmented and segmentation is the crucial step for correct cell tracking. In the phase-contrast, cell borders can be seen by eye only for very low cell densities. In the case of MDCK-II cells, the borders can be identified for monolayer densities below 2000 cells/mm$^2$. To avoid this obstacle, most researchers use particle image velocimetry (PIV), which is a method usually used to perform velocity measurements in fluids. This method is based on dividing the region into small pieces, called interrogation windows, and finding the maximum correlation function between an initial interrogation window and potential new window in the successive image. Other methods for indirect cell tracking are available, like optical flow constraint, but PIV has become a standard method in biophysics to analyze internal motions of the cytoskeleton, as well as global movements of cells in the tissue or fluid-like motion of bacteria suspensions [145]. However, the problem with PIV is that it does not enable tracking of individual cells and often cannot decouple cell deformations from cell movements. To our knowledge, the group of Silberzan was the first that utilized PIV analysis for the phase-contrast images of the cell monolayer to obtain the cell velocities [55]. Today PIV is a standard method
of analyzing cell velocities from the cell monolayer with very few studies using direct cell tracking.

One study that used direct cell tracking investigated wound healing with human umbilical vein endothelial cells (HUVEC) stained with Hoechst [47]. Cells were tracked with a self developed code written in MATLAB. This program identifies cell nuclei (bright objects) in each step and their coordinates are linked together based on minimum Euclidean distance to form trajectories. Another example is the study of a flow of MDCK cells in narrow channels which were stably transfected to express mCherry-labeled H2B histones [58]. In this study PIV analysis was complemented with direct cell tracking of single cells which showed random motion upon subtracting drift velocity.

Even though many cell tracking studies were done on various systems, only a few studies focused on quantification of cell velocity in respect to time, position and surrounding cell density. In unconfined cell clusters a drop in cell velocity with increase in cell density was reported [64]. In circular confinements, the cell velocity increased with the radial distance [59], while in micro-channels the cell velocity decreased with the distance from the leading edge and increase in cell density [58]. The speed of the leading edge was reported to be constant [57]. Even fewer studies analyzed the growth of the MDCK cluster with time. In the study by Puliafito et al. [37], area growth of an initially small colony (a couple of cells) was measured for eight days. The results pointed to an exponential growth in the first six days followed by subexponential one. The authors argued that the crossover happens when cells at the border of the cluster reach the maximal velocity of the cell motion, which they measure to be 15 µm/h. Unlike those results, wound healing experiment [146], dating from 1980, reported constant acceleration of the MDCK monolayer. In this experiment, the MDCK monolayer with various cell densities was wounded with a stiff blade leaving a rectangular 8 mm wide wound. The wound enclosure was followed for five days. It is therefore unclear whether cell cluster grows with a constant velocity or constant acceleration, or in some other regime. Furthermore, it is not known what is the exact distribution of the cell velocities with the distance from the leading edge; does it change in time and do all the cells move outwards?

We addressed this problem with circularly seeded unconstrained large clusters of MDCK-II cells. We analyzed cell velocity vectors by direct tracking of cells in colonies seeded in two experiments with different initial conditions. In the first experiment, the cell density was very low, which enabled us to use direct tracking from phase-contrast images for the first 48 hours. In the second experiment, we seeded the cell cluster
with 10.4% permanently transfected MDCK cells\(^1\) and tracked the trans- 
fected cells until contact inhibition in the bulk region was reached. The 
data from both experiments pointed to the spatial dependence of the cell 
velocity. In the bulk region velocity is constant, while in the edge region it 
linearly increases toward the border of the cluster.

Based on measured cell velocities and cell proliferation as a function 
of cell area (for details see Subsection 3.5.2), a Voronoi tessellation based 
simulation of epithelial colony growth was done. This simulation success-
fully repeated formation of the bulk and the edge regions in correct time 
interval with cell density and cluster area as in experimental conditions. 
Moreover, scenarios of various cell behaviors were investigated.

4.2 Cell tracking

4.2.1 Tracking algorithm

Our tracking routine reads positions of the cell centers of mass and 
tracks them over time. In the next time step (20 of 15 min) the algorithm 
searches all centers of mass in a given expected radius around the refer-
ence cell - one and only one center of mass, if found, is associated with the 
tracked cell (Figure 4.1). The procedure is repeated for a given number of 
images and only the cells which were tracked in all images are memorized. 
The total number of tracked cells and their velocity depends on the radius 
of the circle in which tracked cell is expected to be found in the next time 
step. For example, a large expected radius will enable tracking of the fast 
and large cells (cells closer to the edge of the cluster). The expected radius 
was set to be 7.8 \(\mu\text{m}\), 13 \(\mu\text{m}\) and 18 \(\mu\text{m}\) and cells were tracked for 2 hours. 
The radial distances were measured from the center of the cluster. If more 
than one expected radius enabled tracking of the same cell, the cell was 
counted only once. The angle of the cell motion was defined as the angle 
between vector pointing from estimated center of the cluster to the center 
of mass of a reference cell and the vector which represents the total shift 
in position of the reference cell during 2 hours.

\(^1\)Permanent transfection of the cell nuclei enables more precise image segmentation, 
however this method introduces chemical toxicity \[147\]. Therefore, we have used only 
10.4% of transfected cells to ensure that collective migration is not affected by the che-
mi-cal perturbation.
Figure 4.1: Schematic illustration of the tracking algorithm. In the next time step the center of mass of the reference cell is expected to be found within the expected radius which is set to be $7.8\,\mu m$, $13\,\mu m$ and $18\,\mu m$. Cells were tracked if only one center of mass was found within the expected radius.

4.2.2 Experiment No.1: Cluster with low cell density

Experiment\(^2\) The seeding procedure was similar to the mechano-sensing experiments, where 15,000 MDCK-II cells were seeded in $7\,\mu m$ droplet on $30\,kPa$ polyacrylamide (PA) gel. Images were taken every 20 min for two days in phase contrast. The setup for live cell imaging was built upon a Zeiss Axio Observer.Z1 equipped with the Andor Zyla 5.5 sCMOS camera. In order to sustain an environment with $5\%\,CO_2$ at $37$ and $80\%$ humidity, the ibidi Heating System was used in combination with their gas mixing system named "The Brick". Images were taken with $5x$ magnification (Zeiss # 441020-0000) which enabled the cell cluster to fit in four fields of view. The pixel size was $1.296\,\mu m$.

**Image segmentation** Input data are two $2.7 \times 3.2\,mm$ fields of view filmed at different positions within the same cluster. The first step in the segmentation process is background subtraction, done by subtracting the estimated background illumination for both fields of view upon analysis of the image intensity profiles. Noise reduction is conducted by Wiener filter (using $5x5$ pixels neighborhoods), which is a type of linear filter that adapts itself to local variance. If the local variance is large, the filter per-

\(^2\)The experimental data used in this paragraph was provided by Philipp Linke who performed this experiment as a part of his master thesis finished under supervision of Prof. Dr. Florian Rehfeldt in Göttingen.
forms little smoothing, while more smoothing is applied for small variance. The segmentation is performed on the denoised phase-contrast image (Fig 4.2 B). Firstly, the one pixel thick network representing local regions of low intensity (cell membranes in this case) is created as explained in Subsection 2.1.3. This mask is added to the phase-contrast image before recognition of the objects in the image. Object recognition is done by the MATLAB function ‘bwconncomp’ which is using the connected-component labeling procedure (flood-fill algorithm). By using this procedure, all the regions in the image connected with predefined connectivity (four in our case) will be found. In the Figure 4.2 C we show the result of our segmentation procedure. Between 5,900 and 8,000 cells were tracked in both fields of view, depending on the initial time point. The down side of phase-contrast images is that cell borders become blurry for cell densities larger than about 2,000 cells/mm². Therefore, tracking cells from phase-contrast movies has an inherent limitation and only low density clusters far from steady state density can be analyzed.

Results  Cells do not start collective motion until they form a confluent monolayer which, in the case of this experiment, happened after 7 hours (Figure 4.3). After the monolayer reaches confluence, cell movement becomes coordinated first for cells very close to the cluster border. It took approximately 26 hours for all cells in the cluster to start moving in coordinated fashion (Figure 4.5).

The first regime is characterized by the outward movement of cells close (less than 1 mm) to border of the cluster while the bulk cells have velocity of about 5 µm/h and a randomized directions of movement (Figure 4.4 D). Cell velocity slowly increased with time until a critical point
about 26 hours after seeding when a sudden burst in cell motility occurred. In the next 4 hours (26-30 hours after seeding), cell velocities in the bulk region increased to 15 µm/h (Figure 4.5 D). Moreover, the distribution of cell velocity directions in the bulk region quickly changed from random movement to directed outward movement (Figure 4.5 E,F). From the angle distribution, it can be seen that at 30 hours after seeding most of the cell velocity directions deviate less than 30 degrees from outward movement (Figure 4.5 E). In this period, cell density in the bulk region was about 1,600 cells/mm² and no significant change in the cell density was observed (Figure 4.4 F), indicating that the burst is not triggered by cell density alone, but also with maturation of cell-cell contacts.

The velocity linearly increased towards the cluster border for cells less than 1 mm away from the cluster border. The velocity of the cells at the border increased with time and during the second regime those cells moved with velocities of ~ 25 µm/h on average, while individual velocities reached up to 40 µm/h (Figure 4.4 E) which is much higher than 15 µm/h which was previously reported maximal velocity of those cells [37]. Once the distribution of cell velocities obtained its characteristic shape, it remained the same until the end of the experiment (Figure 4.4 A).

During the second regime, we observed cells moving in a swirling motion, starting from ~ 30 hours and lasting until ~ 43 hours after seeding (Figure 4.4 G). This swirling behavior has been reported previously in the literature and here we confirm existence of such cell behavior on cell densities of about 1,600 cells/mm². This cell density is consistent with the study by Angelini et al. [60] and the swirls are located towards the bulk region as reported in the study with micro-channels [58].
Figure 4.4: Radial dependence of cell velocity and cell density over time

(A) Mean cell velocity as a function of position in the cluster for various time points. The cluster is divided into 15 equidistant regions from the estimated center of the cluster to the location of the last tracked cell in the cluster. The first four regions are averaged and presented as one scatter dot in the plot due to smaller statistics of the cells. Likewise, the average velocity of the cells from the last region is not plotted due to small statistics in that region. (B-C) Visual representation of the tracking results in both fields of view. Cells were tracked for 2 hours starting from 20 hours (A) and 36 hours (B) upon seeding. Equidistant regions are represented with yellow circles. (D-E) The upper plot represents the scatter plot of individual cell velocities depending on the cell distance from the center of the cluster. The lower plots show histograms of the cells velocities angles relative to the outward movement from the center of the cluster. Color of the individual bars represent the mean distance from the center of the cluster of the cells belonging to a given angle bar. Mean velocities and their directions are plotted for the cluster 20 hours (D) and 36 hours (E) after seeding. 

(F) Cell density profile with respect to time. Cell density was estimated by hand counting and the associated error bar is $\pm 300$ cells/mm$^2$. (G) Swirl-like movement of the cells was observed close to the center of the cluster between 30 and 43 hours after seeding. Analyzed movie is contributed by Philipp Linke.
Figure 4.5: Transition from slow and random movement (A-C) to fast and coordinated movement of the cells in the bulk region (D-F). (A,D) Cell velocities as a function of the radial distances from the center of the cluster. (B,E) Distribution of the velocity angles. (C,F) Movie snapshots. Red circles denote tracked cells and red lines cell movement occurring within the next 2 hours. Movie contributed by Philipp Linke.
Figure 4.6: Image of the cell cluster with 10.4% of stably transfected cells 30, 40 and 56 hours after seeding. Tracked cells moving in the outwards direction are marked with yellow circles, while those moving inwards are shown in green. Cell movements within the next hours are marked with the red lines. Movie contributed by Damir Vurnek.

4.2.3 Experiment No.2: Cluster with formed bulk and edge

Experiment A MDCK-II cell line permanently expressing GFP with a nuclear localization sequence was used. Initially, 5,300 cells were seeded in a 1 µm droplet on the glass substrate coated with collagen-I. Out of those 5,300 cells, 10.4% were permanently transfected and the rest were wild type. The cell cluster was imaged in bright field and fluorescence FITC channel every 15 minutes on a ZEISS LSM780 confocal microscope with a Plan-Apochromat 10x/0.45 objective. The movie starts 24 hours after seeding and lasts until the end of the third day (48 hours). Images were stitched with Fiji plug-in published in [148] and the whole cell cluster was analyzed.

Segmentation Since cells with stably transfected nuclei were sparsely seeded, cell nuclei were well separated. We thresholded the images with Otsu’s method. In the next step we removed spur pixels and objects smaller than 21.5 µm². Objects in the binary image were found by “bwconncomp” MATLAB function, which uses flood-fill algorithm, and centers of the mass were found with “regionprops” function. Those points were subsequently

---

3 Transfection was done by Damir Vurnek and Simone Gehrer as a part of the collaborative project with Prof. Dr. Diana Dudziak in the Department of Dermatology, University Hospital at FAU Erlangen.

4 This experiment was performed by Damir Vurnek in the laboratory of Prof. Dr. Diana Dudziak in the Department of Dermatology, University Hospital at FAU Erlangen.
tracked in time with our algorithm (see Subsection 4.2.1). Cell density was estimated by hand counting cells in bright field images in the regions of $\sim 0.05 \text{ mm}^2$. Unlike phase contrast, cell borders are easily seen in bright field at higher densities ($>5,000 \text{ cells/mm}^2$), but appear blurry at lower densities.

**Results and comparison** After 24 hours the cells reached confluence in a circular cluster with a radius of about 1.5 mm. We consider this to be the initial point of the experiment. The initial density in the bulk region was $\sim 2,200 \text{ cells/mm}^2$ which is greater than the bulk cell density at the end of the first experiment.

In this experiment, cells in the bulk have never reached large velocities and they did not move outwards (Figure 4.6). However, the velocity of the cells deep within the bulk region was constant, similar to the first experiment. In the first experiment, cells in the bulk had average velocity up to 15 $\mu\text{m}/\text{h}$ and the cell density in this region was very small (maximally $\sim 1,800 \text{ cells/mm}^2$). In the current experiment, after 36 hours - when coordinated movement started – the density in the bulk reached about 5,200 cells/mm$^2$ and the constant cell velocity in the bulk region remained below 10 $\mu\text{m}/\text{h}$. Upon 56 hours, when the bulk density is very close to the steady state density (approx. 6,200 cells/mm$^2$), this velocity is smaller than 5 $\mu\text{m}/\text{h}$. Therefore, cell velocity deep in the bulk region is a spatial constant for both experimental conditions, but the value of velocity drops with an increase in cell density.

Similar to the first experiment, coordinated collective movement emerged over time. During the first period (about 24-36 hours), the cells move in all directions with a small and position independent velocity (Figure 4.7 A,B). About 36 hours after seeding, coordinated outward movement of the cells at the border of the cluster began (Figure 4.7 C,D). Within the next 20 hours (36-56 hours after seeding), more and more cells started to move outward and cell velocities increased with time. After 56 hours, all the cells less than 1 mm away from the cluster border were involved in movement and velocities at the border of the cluster reached 30 $\mu\text{m}/\text{h}$. Cell velocity increased approximately linearly from the bulk towards the border of the cluster. However, in this experiment only cells at the edge displayed outward movement, while the bulk cells moved in all directions with small velocities.

Even though the clusters in both experiments had similar initial size (approx. 2.3 mm and 1.5 mm radius), after 48 hours they had different cell velocity and cell density distributions. This discrepancy proves that
Figure 4.7: Cell velocity distributions over time. (A,C) Cell velocity as a function of the distance from the center of the cluster at the beginning of cluster growth (A) and in the later, more coordinated stage of the growth (C). (B,D) Histograms of the cell velocity angles deviating from the outward movement from the center of the cluster. Color of the individual bars represent the mean distance from the center of the cell cluster belonging to a given angle bar. During the first period, cells had random directions of movement, but slowly became more outwards directed (B). Cell velocity did not depend on the position within the cluster (A). During the second period, most of the cells moved outward and those cells were located in the edge region of the cell cluster (approx. 1 mm away from the cluster border). (D). Velocity of the cells at the edge of the cluster was monotonically increasing.
initial conditions play an important role in cluster development. It seems that not only is the morphology of the cell cluster important (cell density distribution, size and shape of the cluster), but the time the cells have spent in that particular configuration as well. For example, the cell density in the cluster at the beginning of the second experiment was similar to that from the first experiment two days after seeding, but the velocity distributions were dramatically different. Interestingly, for both clusters the region of linearly increasing velocity was approximately one millimeter large.

4.3 Cluster growth in time

Experiment

Trypsinized cells were counted and diluted to a concentration of 5 x 10^6 cells/ml. Monolayers were seeded as droplets in cell medium pre-filled chambers (2 well glass bottom, ibidi, Germany or 6-wells with glass chamber slides). Coating was done with 10 μg/cm^2 of collagen-I (Corning, Rat Tail).

For seeding individual cells, cell suspension was diluted and multiple wells were seeded at cell concentration of 1 cell/cm^2. After 24 hours attachment to the wells was checked and two wells with a cluster containing four and seven cells, respectively, were chosen for the growth experiment.

Intermediate size clusters were seeded as a 0.25-0.5 μl cell suspension droplet. Due to small droplet size this seeding is not highly controllable and the cell number was around 500 cells. The experiments were initiated when cells merged into a confluent monolayer (24-48h after seeding).

Large clusters were seeded as a small droplet (1 μl) of cell suspension placed in the middle of the chamber. The droplet contained about 10,000 cells. Upon attachment the cells formed a circular cell sheet.

Results

We assume that cell clusters with the same area and similar cell density distribution will behave the same upon activation of cell motility burst. Therefore, we can shift the initial time point for macroscopically seeded clusters. The results reveals a unique exponential fit with time-dependent exponent for the growth of all clusters independently of their initial conditions. Measured velocity of the cells at the border reaches up to 50 μm/h.

This section is based on the experiments conducted and analyzed by Damir Vurnek under supervision of Prof. Dr. Ana-Suncana Smith.
Time-dependent area growth can be approximated with the exponential function of the following form:

\[ A(t) = A_02^{D(t)t} \]  

(4.1)

\[ D(t) = \frac{1}{(\alpha + \beta t)}, \alpha = 0.586, \beta = 0.0422 \]  

(4.2)

Consistency with previously published data  The study by Puliafito et al. [37] reported an exponential growth in the first five to six days, followed by sub-exponential growth. Even though this seems to be contradictory to our growth curve, it is not. When shifted to the same initial area and the same time period, both curves coincide for the first five days. In this time period, the exponential curve with cell doubling time of one day fits the experimental data after which the sub-exponential growth period starts.

The investigation by Rosen et al. [146] reported constant acceleration of the cluster, while our growth curve implies small changes in acceleration over time. A linear fit of cell velocity as a function of time \((v = v_0 + \frac{1}{2}at)\) revealed that cells at the border of the wounded tissue, with the steady state density (6,600 cells/mm²) and initial velocity of 2.8 \(\mu\)m/h, had an acceleration of 0.23 \(\mu\)m/h². When the tissue at lower cell densities was
wounded (1,700 and 1,300 cells/mm²), cells at the edge had a postponed reaction, but once they started moving their acceleration was 0.29 µm/h² and 0.21 µm/h², respectively. When we plotted our data on the same time scale (Figure 4.9 D), we found that cells have the initial velocity of 2.5 µm/h and the acceleration of 0.21 µm/h², which is very close to the results by Rosen et al [146].

Therefore, this growth curve is consistent with previously published data. However, none of the previous experiments followed the growth on long enough time scales to observe the universal growth curve, which is indeed exponential, but with a decaying factor that accounts for decrease of the cell proliferation due to the densification of the tissue.
4.4 Simulation of the circular cell cluster growth

To our knowledge, the model by Zimmermann et al [73] is the only one that reproduced, at least qualitatively, the distribution of the cell density in the circular cell colony. In this model, cells tend to move away from each other by adjusting the magnitude and direction of their self-propulsion force due to contact inhibition of locomotion (CIL). Due to CIL, cell velocities were very low in the bulk region. Toward the edge, cell velocity increased and cell density decreased. However, colony growth was constant and cells divide with the same probability if the distance between two particles is larger than a threshold value. Likewise, final spatial distribution of the cells’ density in the cluster is U-shaped, which does not correspond to our experimental results.

All models, whether they are continuous, self-propelled particle models or models which consider motility and morphology of cells, are assuming some reasonable physical mechanisms which lead to experimental results that are available in the literature. However, we know that experimental results available for collective cell motion are extremely sensitive to initial conditions, time period of the experiment, as well as experimental details that vary from laboratory to laboratory. For example, difference in substrate preparation, cell medium, exact cell line or the passage of the cells used in the experiment can influence the result significantly. Furthermore, models are using only qualitative comparisons with experimental data and usually select a set of experimental phenomena that they want to reproduce.

On the other hand, our experimental results have been harvested from different experiments and stages of growth. Our aim was to devise a model that is able to reproduce observed cluster morphology under experimentally established rules of cell behavior. The main hypothesis is that the cell cluster growth (Equation 4.1), cell proliferation probability and distribution of cell velocities throughout the cluster must lead to formation of the bulk and the edge regions. We investigated the necessary conditions for this spontaneous compartmentalization, as well as various scenarios with our phenomenological Voronoi cell based cluster growth simulation.
4.4.1 Formation of the bulk and edge regions in colonies grown from few cells

The most consistent observation in all experimental conditions is formation of the bulk and the edge regions in cell clusters. However, the time needed to form these regions and their final size depends on the initial conditions. The size of the edge region varies from cluster to cluster, and it varies in time for the same cluster. We have chosen to simulate the growth of very small cell colonies until the bulk and the edge regions are formed.

The experiments on which we based our simulation were part of the experiments for measuring cluster growth with time (see Section 4.3). Two cell clusters are grown from very small colonies (less than 100 cells) and imaged 11 times in the phase-contrast for the first 7-8 days (see Figure 4.10). Within those 7-8 days, the clusters have reached steady state cell density in the bulk, which became evident through large dome-formations in the bulk region. As already explained in section 4.1, phase-contrast images do not reveal cell borders at higher cell densities. Therefore, it is impossible to quantify the exact cell density for these two clusters. However, the evidence of the domes in the bulk region can be used in this case as a reliable indicator that the cell density is either approaching, or already is, in the steady state.

The visual inspection of the clusters revealed that even though the initial conditions were very similar, the final morphology of the clusters differ (see Figure 4.10). Cluster B has reached the steady state in the bulk region (demonstrated by large dome formations) after 7 days, while cluster A still didn’t have domes in the bulk in the same period. Furthermore, the cluster B is much more circular than cluster A. Therefore, variability is inherent to the behavior of large cell colonies; however, common characteristics can always be found. We investigated the necessary conditions for formation of the bulk and the edge regions 7-8 days upon seeding a small cell colony.

4.4.2 Algorithm

We simulated cell cluster growth with a self-propelled Voronoi tessellation (VT) model following the algorithm presented in the Figure 4.11. Initial seeding conditions were chosen to match the initial conditions of the experiments: the initial area in our simulation is 0.055 mm$^2$ (this value is the $A_0$ from the growth curve), while the cell density is 1300 cells/mm$^2$ (72 initial cells).
Figure 4.10: Phase-contrast images of two different cell clusters grown from very small colonies until the bulk and the edge regions are fully formed. Images contributed by Damir Vurnek.
Relaxation of the Voronoi tessellation  The relaxation of cell shape in the simulation is based on Lloyd’s relaxation algorithm. This relaxation method places the seeds of the Voronoi tessellation in the geometrical centers of Voronoi cells generated from those seeds and eventually leads to centroidal Voronoi tessellation. Two iterations of the Lloyd’s algorithm lead the Voronoi tessellation generated from randomly distributed seeds to the Voronoi tessellation which has approximately same features as the real tissue (Figure 4.12 B). A better approximation for the local organization of the cells will be presented in fifth Chapter, but it is computationally much more expensive. Considering that our main interest is the global morphology of the cell cluster, we decided to use a simple approximation to describe the cell shapes. We found that Voronoi tessellation generated from a single seed and relaxed with the Lloyd’s algorithm acts as a simple descriptor for the cell shape, while being able to capture the main features of the tissue relaxation upon cell division and migration.

Cluster growth in time  We imposed boundary conditions to our system so that the growth of the cluster follows the experimental curve (for detail see Equation 4.1, 4.2).

Distribution of the cell velocities within the cluster  Once collective cell motion is triggered (see Section 4.2.1), velocities increase approximately linearly towards the border of the cluster except for the cells which are too deep in the bulk region. In this region cell velocity seem to be significantly smaller and constant. This constant velocity could still be large if the cell density in the bulk is very low, as is the case with the first experiment. If the density is high, as observed in the second experiment, cell velocity in this region is very low and has no preferred direction.. However, we do not have the data on cell velocity distribution for small cell clusters which
Figure 4.12: Comparison of the Voronoi tessellation generated from randomly distributed seeds relaxed with two iterations of Lloyd’s algorithm with the Voronoi tessellation generated from the centers of cell nuclei mass. (A) Visual comparison between the tissue at low cell density and the Voronoi tessellation generated from the centers of cell nuclei mass. (B) Probability distribution of the Voronoi cell areas for tessellations generated from centers of cell nuclei mass (blue), randomly distributed seeds (red) and randomly distributed seeds relaxed with two iterations of Lloyd’s algorithm (black). (C) Visual representation of tessellations generated from randomly distributed seeds and randomly distributed seeds relaxed with Lloyd’s algorithm. The insets represents the value of the average distance between the cell centroids and the seeds that generated the tessellation.

Figure 4.13: Input data for the cell velocity. (A) Velocity of the cells at the edge over time is defined by the growth curve with a damped exponential term. (B) At the eight day, the value of the velocity is low and constant in the bulk and increases linearly in the edge region to the value defined by the growth curve. (C) A noise term is added to the directional movement of the cells. The value of the noise is a random value between $-2\pi$ and $2\pi$ in the very center of the cluster and between $-\pi/6$ and $\pi/6$ in the edge region. In the bulk region the upper limit for the noise term linearly decreases to the $\pi/6$ value.
grow to large sizes. Therefore, we assumed the same cell behavior for the clusters seeded from a few cells and from few thousand cells, and tested if such assumption would lead to the formation of the bulk and the edge region upon eight days.

We assumed that during the first day all the cells have the same velocity in the random direction. After one day the cells have “sense” of direction. Initially, the cluster is so small that all the cells can be correlated with the cells from the border region and cell velocity is linearly decreasing towards the center of the cluster. However, once the cluster becomes sufficiently large, the cells in the bulk loose sense of direction and have low velocities in random directions.

Based on the experimental data (see Figures 4.4, 4.7) we assumed that cells more than 1 mm from the border of the cluster lose correlation with the cells at the edge. The direction of cell movement is random in the center of the cluster and in the edge region it is in an outward movement ± maximally 30 degrees. The upper limit for the deviation from the outward cell movement in the bulk region changes linearly from $2\pi$ to $\pi/6$ (Figure 4.14 C). The velocity of the cells in the edge of the cluster is defined by the growth curve (Figure 4.14 A) and the velocity in the bulk is set to be 2 $\mu$m/h based on the experimental data 4.7. The velocity increases linearly from the bulk region to the border of the cluster (Figure 4.14 B).

However, motion of the cells is also restricted by their close neighborhood. If the cell movement would place the cell closer than 7 $\mu$m to a neighboring cell, the movement is suppressed. The distance between the cells is defined as the distance between the Voronoi seeds.

**Cell divisions** In every step of the simulation every cell has a probability to divide dependent on its area by the curve presented in the Figure 3.12. Once divided, both the mother and daughter cells cannot divide in the next eight hours. This restriction makes cell proliferation less frequent. However, the proliferation arrest time is in agreement with the initial doubling times of the cell clusters. By following $N = N_0 * 2^T$ and counting the number of cells at various time points within the first two days of the experiment, we can easily obtain values for the initial doubling time of the cells. This value varies from 11 to 15 hours, which is consistent with our limit that the cell division cannot be faster than 12 hours. When cell division occurs, a new seed of the VT is created 5 $\mu$m away from the original one in the direction of the cell elongation. Cell proliferation follows the experimentaly obtained quadratic curve shown in Figure 4.14.
Figure 4.14: The quadratic curve approximating the fraction of the cells in the S-phase of the cell division on the glass substrate as a function of the cell area.

\[ \text{Div}(A) = aA^2 + bA + c \]

where

\[ a = -3.7 \times 10^{-6}, \quad b = 0.00341, \quad c = -0.305 \]

with two cutoff conditions:

\[ f(A) < 0 \rightarrow f(A) = 0 \]
\[ A > A_{\text{crit}} \rightarrow f(A) = f(A_{\text{crit}}), \quad A_{\text{crit}} = 539 \, \mu m^2 \]

**Defining the tessellation at the boundary**  To be able to define Voronoi cells at the edge of the cluster, three sets of additional seeds are placed around the cluster. The closest set of seeds is 25 \( \mu m \) away from the edge cells which corresponds to the typical size of the cells at the border. Additional seeds are not considered as cells. Cluster radius is defined as \( \sqrt{\frac{\sum \text{Cell Areas}}{\pi}} \).

### 4.4.3 Results

Implementing the cluster area growth, proliferation rate and the distribution of the cell velocity within the cell cluster resulted in a correct morphology of the cluster for the first eight days of development. During this period, the spontaneous compartmentalization in the bulk and the edge regions occurs (Figure 4.15 A).
Even though the initial conditions were the same, the cell clusters A and B differ after 7 days. However, in the early phase both cluster grow with low and uniform density. But after 4 days monolayer starts to sporadically densify (Figures 4.10). This morphology was reproduced by the simulation (Figure 4.15 A). Within one more day the densification becomes much more obvious. After six days, the bulk region is clearly denser than the edge region, which can be seen as blurry cell borders in the bulk region of cluster A. This trend becomes even more pronounced after 7 days. The 8th day brings to light the domes in the very center of cluster A indicating that this region has reached the steady state density. The simulation results correspond to the cluster morphology of the cluster A. The cell density distribution in the simulation shows the formation of the bulk region and the edge region where density continuously decreases after eight days.

Even though we have used results of the measurements made on macroscopically seeded clusters, we have successfully reproduced time-dependent behaviour of the global morphology of the cell colony grown from just a few cells. This supports the hypothesis that the same principles guide the cluster growth independently of the initial size of the colony. Likewise, the final morphology of these clusters is the same, with well separated the bulk and the edge regions. We have confirmed that cell velocity should linearly decrease from the border of the cluster towards the bulk region, where it saturates at the low and constant value. We have also shown that probability of the cell division is a quadratic function of the cell area, and that the cluster area growth is a damped exponential function of time. However, it is not clear whether this is the only set of conditions under which the bulk and the edge regions would form. Therefore, we investigated different growth scenarios.

The effect of varying conditions

Different bulk velocity For the cell velocity, we have assigned random directions in very small colonies (first day). The random regime is followed by the regime where velocity linearly depends on the cell position in the cluster. All the cells move outwards with the noise term which is redirecting their outward movement by a random value between $-\pi/6$ and $\pi/6$ (see Figures 4.4 E, 4.7 D). When the cluster grows bigger than 1 mm in radius, cells in the bulk region have a small constant velocity (2 µm/h), randomly directed. These assumptions were based on the experimental data.

We have seen that the cell velocity in the bulk region is strongly dependent on the cell density. It is therefore possible that this cell velocity is
Figure 4.15: Simulation of the cell cluster growth. (A) Growing cluster at various time points. (B) Sections from phase-contrast images of the clusters A and B (images contributed by Damir Vurnek) and the result of the simulation upon 8 days. Sections are representative of the region spreading from the very center of the clusters to the border of the cluster. (C) The average cell density distribution within the simulated cluster after eight days. Error bars correspond to the standard deviation of the cell density within the given ring.
larger than 2 µm. The simulation results show that larger cell velocities in the bulk region (6 µm) lead to lower cell density in the center of the cluster and higher density in the edge region. If the bulk velocity increases even more (12 µm), the bulk and edge regions would not form at all, and the density in the center of the cluster would actually be lower than in the edge region. On the other hand, if cell velocities are directed outward and have a simple linear dependence on the radial distance, the resulting clusters have relatively uniform cell density and the bulk and the edge regions do not form.

Arresting motion in the bulk region is necessary for the formation of the bulk and the edge regions. When cluster reaches a certain size, motion of the cells that are too far away from the cluster border is no longer coordinated with the edge cells. A distance at which the border cells lose information about the edge position is used as an input parameter in this simulation.

Cell velocity is not a linear function of the cluster radius. The cell tracking data is quite noisy (Figures 4.5, 4.7) and one could argue that the data is not linear, but linear-like or even quadratic. For example, let’s assume that the cell velocity is not linear function of the cell position \( v = \frac{v(Radius) \cdot r}{Radius^1} \) but instead it depends on the cell position like \( \frac{v(Radius) \cdot r^{1.5}}{Radius^{1.5}} \) or \( \frac{v(Radius) \cdot r^2}{Radius^2} \) (see Figure 4.17 A).

In this scenario, we need to look closely to the movement of the cells at the border of the cluster. The average radius of the cluster was defined as the radius of the circle that would have the same area as the total area of the colony. The cells that are further away than the average radius of the colony now have large velocities (Figure 4.17 A) and the cluster ends up with unreasonably large cells at the edge. However, we assumed that the value of the cell velocity saturates for the cells near the edge as indicated.
Figure 4.17: Effect of cell velocity distributions on cluster morphology. (A) Three different functions of cell velocity as a function of cell’s radial distance for days 5 and 8. For radial distances larger than the cluster radius cell velocities were set to the constant value. (B) Morphology of the cluster upon 8 days for various functions of the cell velocity radial dependence.

by the data from the first experiment (see Figure 4.4 A). Therefore, the velocity of the cells which are further away than the average cluster radius saturates (see Figures 4.4 A, 4.5).

It was observed that if cell velocity depends on cell’s radial distance as $v = \frac{v(\text{Radius}) \cdot r^{1.5}}{\text{Radius}^{1.5}}$ or $v = \frac{v(\text{Radius}) \cdot r^{2}}{\text{Radius}^{2}}$, the bulk and edge regions still form as in the case of $v = \frac{(v(\text{Radius}) - v_{\text{bulk}}) \cdot (r - (\text{Radius} - \text{Edge Size}))}{\text{Edge Size}} + v_{\text{bulk}}$ with a cutoff in the bulk region. However, the bulk size is too small, especially in the first case. The quadratic dependence resulted in the low cell density in the edge region with less than 2,000 cells/mm$^2$, which was not observed in the experiments.

Another interesting observed phenomenon is formation of a high density ring close to the edge of the cluster as a consequence of the velocity arrest for the cells which are further away than the average radius of the cluster. Very similar formations was seen in cell colonies grown on glass
Figure 4.18: Cluster morphology upon 8 days if the proliferation probability was constant (20% and 35%) or a linear function of the cell size.

Figure 4.19: Effect of shifting the proliferation curve on the cell cluster morphology after 8 days.

substrates, where the ring with larger cell density forms close to the very edge of the cluster after bulk and edge regions. We hypothesize here that this ring is a consequence of the velocity saturation for the cells in close proximity to the edge. However, why this occurs on glass substrates and not on hard gels has not been yet understood.

**Probability of cell division has a different cell size dependence** Furthermore, we tested the effect of cell division on the cell cluster morphology (Figure 4.18). We found that small and constant values of cell proliferation would not lead to formation of the bulk, but if the value is large enough, the bulk and edge would form. However, the proliferation, which does not depend on the cell area, would induce much more density inhomogeneities in the cluster morphology. Proliferation probability, which linearly depends on the cell area and has the same cut-off values, leads to a smaller bulk region and a lower density in the edge region.

**Cell division is faster or slower than assumed** The percentage of the cells positive for the EdU marker was measured as a function of the cell area. As explained (see Subsection 3.5.2) EdU marks the S-phase of cell
division. The EdU was applied for 4 hours before cell fixation and therefore it has marked the cells that have been in the S-phase and the ones that have entered it within the 4 hour period. However, we do not know how long the S-phase lasts nor is the time cell size dependent. Also some cells could be, for example, stuck in later stages of the cell division (for example, G2-phase).

To examine the effect of the proliferation probability on the cluster morphology, we conducted two simulations; one with a lower probability (50% lower), and one with a higher probability (50% higher). The proliferation probability is defined as a fraction number of cells which incorporated EdU marker into their DNA during the four hour period. For the lower probability, the bulk and the edge of the cluster did not form after 8 days. Higher proliferation probability induced higher cell density of the bulk and the edge regions after 8 days, which morphology was reminiscent of the cluster B.

4.5 Discussion

Cell velocities are surprisingly coordinated. All the cells in the edge region move outwards. The edge region is annular in shape and is in the order of a millimeter thick. However, the size of the edge region changes with time. For example, two clusters used for cell tracking had a radius of about 2-2.5 mm and the edge region was about 1 mm thick. On the other hand, clusters six days after seeding had a radius of 9 mm and an edge region of about 3 mm (see Figure 3.9). With time, cells further away from the boundary of the cluster are involved in synchronized movements. One possible explanation for this observation is that the movement of the cells at the edge of the cluster, known to be governed mostly by the leader cells, is transmitted and subsequently adopted by the follower cells. However, in this case hundreds of cells were found behind each leader cell that move outwards. We know that the typical correlation length of the velocity fluctuations for the MDCK cells is about 15 cell, which was demonstrated in many experimental conditions (see Introduction 1.3.2). The open question is how the movement of the leader cells is transmitted over hundreds of cells with a growing number of cells that get involved over time.

Another possible explanation is that the cells feel the local gradient of the cell density and move in the direction of lower cell density. However, the local difference of the cell density on the scale of couple of tens of micrometers is very low, and the mechanism of sensing such small differences would be intriguing.
Cluster area growth is a damped exponential function of the time. This implies that the velocity of the cells at the cluster border is a function of cluster radius. We can assume that local cell forces need to be integrated, as suggested by the work of Trepat et al. [40], [62]. We also know that only cells in the edge region move and proliferate, and the cell density is linearly decreasing. The number of cells behind each cell at the boundary of a cluster is a function of the edge size. To increase the number of cells behind each cell at the boundary with cluster radius, the edge size needs to be a function of the cluster radius. Another explanation for the velocity of the cells at the cluster border to be a function of cluster radius, is that all cells, even the ones in the bulk region, contribute to the collective movement. However, we know that this is not the case because the bulk region is completely contact inhibited upon reaching the steady state density. An alternative explanation would be that motility of the cells at the very edge depends on the size of the cluster. This could only be possible through biochemical signaling, but again, the mechanism of such cell behavior would involve very sophisticated chemical sensors.

We still do not understand the origin of the observed cell velocity distribution and time dependence. Further studies on this topic are required and a number of open questions remain, including is the cell velocity governed by the local density gradient, or is the density gradient a consequence of the velocity distribution? Does the number of dividing cells increase toward the edge of the cluster because the cell velocity there is larger and consequently, the cells are bigger? Or is the cell velocity larger due to the pressure created by the proliferating cells? Or finally, is this behavior a consequence of a very complex set of feedback mechanisms?

Future models of cell cluster growth should incorporate the origin of the cell velocity and include linearly decreasing cell velocity dependent on the cluster radius. None of the models so far demonstrate such cell behavior. The model by Zimmerman et al. [73] had the linearly decreasing cell velocity in the circular cell clusters, but the velocity was constant and not radially dependent. If the "passive" mechanism is in the origin of the cell behavior, the cell velocity should be driven by the cell density gradient and the pressure created by the cell divisions, which is cell density dependent. It would, therefore, be interesting to see if such model can reproduce the observed cell behavior.
4.6 Conclusions

In this Chapter, two important results are reported. First, tracking of individual cells over time as a function of cell position within the cluster. Second, the successful simulation of the cell colony growth from a few cells until the bulk and the edge regions fully form.

The distribution of cell velocities was quantified for two cell clusters with different initial seeding conditions. In both cases, a period of time after seeding was needed for collective outward cell migration to emerge. Again, in both cases, the edge cells move outward with velocity, which linearly decreases with cell’s radial distance from the boundary of a cluster. For cells in the bulk region a constant velocity was consistently observed. When bulk density was very low (below 1,800 cells/mm$^2$) as in the first experiment, cells in the bulk moved outward with relatively large velocities. On the contrary, when the bulk density was high (more than 5,000 cells/mm$^2$) as in the second experiment, cells in the bulk moved in random directions with relatively low velocities. We found that the time cells spend in a certain configuration also plays an important role in velocity distribution. Our results demonstrate that not only cell density is important in the collective cell behavior but also the maturation of the cell-cell contacts.

Furthermore, we do a Voronoi tessellation based simulation of cell cluster growth. The measured cell velocity distribution was used as an input parameter. Other parameters used in the simulations included cell proliferation probability as a function of the cell size (see Subsection 3.5.2) and the growth of the cluster area with time. The simulation reproduced the growth of the cell colonies seeded from just few cells for the first eight days. During this period, the colonies reached a radius of about 2.5 mm and the edge and the bulk regions were completely formed. The comparison between the simulation results and the images of the two clusters seeded with just a few cells produced over the eight days revealed very good agreement at all time points.

In addition, the simulation enabled us to investigate the influence of different cell behaviors on cluster morphology. Firstly, we found that suppression of cell movement in the bulk region is essential for the formation of the bulk and the edge regions. If the cell velocity was a linear function of the position, the regions would never form. Moreover, if the velocity distribution was not linear but quadratic, for example, the bulk and edge would form, but the cell density of the edge region was too low and did not correspond to the measurements. Cell proliferation, if too low, did not lead to the formation of the bulk and edge regions. We found that not only
the value, but also the shape of the proliferation curve influenced the final morphology of the cluster. If the probability of the cell to divide does not depend on the cell size, isolated small regions of very high cell density emerge. On the other hand, a linear proliferation function with the same cut off is not sufficient to form the bulk and edge regions. The morphology of a cell cluster on the glass substrate is typically circular with a high density bump close to the border of the cluster. This morphology is obtained if the velocity of cells located beyond the average cluster radius does not increase.
CHAPTER 5

Organization of the epithelial cells and comparison to random systems

5.1 Introduction

During morphogenesis and wound healing cells constantly move and divide until they reach a steady state when they inhibit both proliferation and locomotion. This process has been termed “contact inhibition” where we distinguish the contact inhibition of locomotion (CIL) originally found in fibroblast cells [149] and the contact inhibition of proliferation (CIP). Both contact inhibitions are disrupted in cancer cells which continue to move [150] and proliferate [151]. Many studies have tried to explain the role of biochemical signaling in contact inhibition [152] showing the role of, for example, Pac1 and PIX [153] as well as cadherins [154] in this process.

However, changes of a cell shape and organization in a tissue approaching contact inhibition are rarely analyzed. A study by Puliafito et al. [37] reported the narrowing of the cell area distribution followed by the appearance of the peak in the radial distribution function (RDF) and changes in the histogram of the number of neighbors as the monolayer densifies. Likewise, they measured that the number of very elongated cells (aspect ratio >3) decays over time. However, this morphological measure is quantified from the Voronoi tessellation constructed from the nuclei center of mass (CMVT) which is prone to a systematic error in estimating the cell elongation (see Section 2.1).
As explained in the third chapter, the substrate rigidity regulates the cell organization within the MDCK epithelial cell clusters on the large ($10^{-4} - 10^{-2}$m) scales. Here we study reorganization on the cellular level while the MDCK monolayer reaches contact inhibition. We are interested in the density dependent changes of a set of morphological measures.

5.2 Method

We decide on an intuitive set of morphological measures which describe the cell shape, the relation between the nucleus and the cell shape, and the connectivity of a cell with its neighbors. The following set of morphological measures is chosen: the cell and nucleus areas, the cell perimeter, the number of neighbors, the elongation of the cell and its nucleus, the area of the cell nucleus, the distance between the center of mass of the cell nucleus and the cell body (CMD), the ratio between the area of the cell nucleus and the cell body (NCR) and the standard deviation of the contact length with the neighbors. For the morphological measure of cell area we investigate only the change of its standard deviation (the mean value is an inverse of the average cell density) while for the other measures we investigate the changes of average values and the standard deviations. These measures, 17 in total, are plotted as a function of the mean cell density in a given image. Some of the morphological measures like the cell area, the perimeter, the distance between the centers of mass (CMD) and the standard deviation of the contact lengths with neighbors have to increase with the mean cell density. Therefore, for those measures we use the rescaled non-dimensionalized values like in references \[95\], \[155\]. We do this by fixing the area of each cell to a unit value. Consequently, the rescaled perimeter is $P_{\text{rescaled}} = \frac{P}{\sqrt{A}}$, where $A$ is the area of an observed cell. For example, a circle has the value of the rescaled perimeter as 3.55, a regular hexagon has this value as 3.72 while for a square the rescaled perimeter is 4. Therefore, this measure tells us how much the perimeter of a cell has changed due to changes exclusively in the cell shape. The rescaled cell area is obtained by setting the average cell area to a unit value.

Since segmented $\beta$-catenin images have a high segmentation error that increases with a decrease in the cell density (see Table 5.21) and the morphology of the cells in the MDCK monolayer at low cell density (less than $\sim 4,000$ cells/mm$^2$) is practically inaccessible with $\beta$-catenin images we cannot use the cell membrane images for this kind of analysis. Instead we use the shape based Voronoi tessellation (SVT) method explained in the Section 2.2.
The SVT method is a significant improvement on the Voronoi tessellation generated from the nuclei centers of mass (CMVT) which was commonly used to obtain the morphological features of MDCK cells in the monolayer [88], [42], [37]. However, this tessellation method also suffers from some systematic errors in predicting the mean values, standard deviations and probability distributions for some morphological measures which we have chosen to look at. In the case of the morphological measure of co-alignment between the cell nucleus and the cell body, the SVT method underestimates the alignment angle. Due to the SVT tessellation algorithm the co-alignment is a necessary result. Therefore we do not include this measure in our set of morphological measures. The other morphological measures are correctly reproduced by the SVT method which is evidenced in Figure 2.10.

We still need to check if the errors of the SVT method in predicting the rescaled perimeter and CMD of a cell increase with the average cell density. To do so, we compare the estimates of the average values and the standard deviations of those measures for three different average cell densities. Intermediate cell density (average cell size of $196 \pm 52 \mu m^2$) was represented with 2,500 cells from the monolayers grown on the glass substrate. High cell density (average cell size of $147 \pm 43 \mu m^2$) was represented with 15,000 cells from the monolayers grown on hard PA gels, while very high cell density (average cell size of $80 \pm 18 \mu m^2$) was represented with 3,000 cells from the monolayers grown on soft PA gels.

For the rescaled cell perimeter we have shown (Figure 2.11 A) that SVT slightly underestimates the value of this measure in comparison to the values from the segmented $\beta$-catenin images ("real" cell membrane). This effect is the most pronounced in the case of large cells, which arises from two reasons: firstly, SVT positions the cell vertices in a different way than the segmented cell membrane does (Figure 2.11 B) and secondly, big cells have more curved shapes when estimated from the segmented cell membranes (Figure 2.11 C).

The probability distributions of the rescaled perimeter and its SVT approximation at high cell density are shown in Figure 5.1. We see that, same as in the case of the cell perimeters, the SVT estimates of the rescaled cell perimeters have a systematic error (Figure 5.1 A). Moreover, the error also depends on the average cell densities. The difference between the mean values of the rescaled perimeter estimated from the segmented pictures of $\beta$-catenin and the SVT method for intermediate cell density was $0.32 \pm 0.18$, while for high cell density it was $0.12 \pm 0.08$ and for very high cell density $0.05 \pm 0.4$. We see that the values of the standard deviations also decrease with an increase in the cell density. Therefore, the average value
Figure 5.1: Probability distribution of the rescaled perimeters (A) and the rescaled distance between the center of mass of the cell nucleus and the cell body (CMD)(B) estimated from SVT and segmented images of β-catenin for 15,000 cells seeded on hard PA gels with an average area of 147 ± 43 µm² (high cell density).

and standard deviation of the rescaled perimeter for large cells are even larger than the SVT estimate.

Likewise, for the rescaled distance between the center of mass of the cell nucleus and the cell body (CMD) SVT underestimates the value of this measure in comparison to the value obtained from the segmented images of β-catenin (Figure 5.1 B). The difference between the average value of the rescaled CMD estimated from the segmented β-catenin images and the SVT method on intermediate, high and very high cell densities was 0.03. Therefore, SVT systematically underestimates the rescaled CMD and the error is not cell density-dependent.

We have chosen 65 images of the MDCK monolayer covering the whole range of available cell densities. The size of the images was 430x320 µm for 47 images (these images are acquired the same way as the images used in Chapter 2) and varying from 330x320 µm to 630x520 µm for 18 cropped images (these images are acquired the same way as the images used in Chapter 3). The pixel size was 0.31 µm for all the images. The cell nuclei are imaged at various time points and different positions within the cell clusters seeded on glass, hard PA gels and soft PA gels. Nuclei segmentation is done by the self developed MATLAB routine explained in the Section 2.1.3. However, this segmentation routine is associated with 0.1-1% of over- or under-segmentation error. Therefore, all segmentation errors in the images are corrected by hand (using ImageJ software) upon visual inspection. The processed images are consequently segmented once again and the second segmentation output is now as good as the manual segmentation. Upon nuclei segmentation, shape-based Voronoi tessella-
tion (SVT) is constructed to approximate the shape of the cells from the positions of the cell nuclei (see Section 2.2).

5.2.1 Results

For our set of 65 images we calculate the mean values of the 17 morphological measures and plot them as a function of the mean cell density in that image. We notice that the morphological measures have very similar values independent of whether the cell cluster is seeded on the glass substrates (blue and black) or the PA gels (red) (Figures 5.3 and 5.2). Likewise, the morphological measures of the cells from the cluster edge also do not depend on time. The cell morphology eight days (black) or two and four days (blue) upon seeding on the glass substrates is similar (Figure 5.3). Likewise, the monolayer in the transient classical regime of growth observed on the soft substrates (magenta) (Figure 5.3) has the same inner organization as monolayers seeded on the other substrates. Since glass substrates induce the formation of various irregularities (domes and "lacunars") that ruin the morphology of the monolayer at higher cell densities, we do not have images of the monolayer with the steady state on the glass substrates.

Steady state is therefore represented by the novel regime of growth and the classical regime on the hard gels. The values for all the cell shape measures saturate at the steady state cell density (Figure 5.3) independent of whether this density is reached during the classical regime of growth ($\approx 7,600$ cells/mm$^2$) or the novel regime of growth ($\approx 13,000$ cells/mm$^2$). This means that the cell organization is the same once the cells undergo contact inhibition (CI) independent of the cell density value when CI is reached. This implies that the steady state is not the cell density that the cells want to achieve but a certain organization which, once reached, leads to contact inhibition.

From Figure 5.3 we see that, even though the substrate rigidity is the parameter determining the organization of the cell clusters on large scales, on the micro scale the cell density is the key parameter which regulates the organization and morphology of the cells (Figures 5.2 and 5.3). However, every steady state is the same independent of the cell density.

Cross-correlation coefficient between the average cell density and the average values of 17 morphological measures in 65 images reveals that six measures have cross-correlation values smaller than 0.5, three measures have the value around 0.5 while for eight measures this value is larger than 0.5 (Figure 5.2 A). An example of the morphological measures which
weakly correlate with the mean cell density are the elongation of the cell body and the elongation of the cell nuclei. These measures seem to be almost constant for all the analyzed cell monolayers with an exception of the cell monolayers with a low mean cell density. For example, monolayers with a cell density smaller than 2,300 cells/mm\(^2\) (see Figure 5.2 B) have an average cell elongation of 1.43 ± 0.06 while in more dense monolayers the cells have an average elongation of 1.37 ± 0.03. Furthermore, the average number of neighbors has a weak correlation with the mean cell density. As in the case of cell elongation, the value of this measure changes at very low cell densities. The cells from a monolayer with the mean cell density smaller than 2,300 cells/mm\(^2\) have on average 5.93 ± 0.04 neighbors while at higher densities the cells have 5.96 ± 0.01 neighbors. The rescaled cell perimeter and the ratio between the nuclei and the cell area (NCR) clearly depend on the cell density, but their standard deviations do not show the same behavior. Interestingly, both the standard deviations of the cell area and the nuclei area per se depend on the cell density, but the standard deviation of their ratio does not.

The morphological measures for which the cross-correlation coefficient with the average cell density is much larger than 0.5 all decrease with the cell density except the ratio between the nuclei and cell area (NCR) (Figure 5.3). In other words, as the monolayer densifies the cells vary less and less in their area and number of neighbors and they create a relatively smaller membrane (or, at least, a relatively smaller 2D projection of the membrane). Cell nuclei area and its standard deviation strongly decrease as the mean cell area decreases. Other measures which strongly depend on the cell density are the ratio between the nuclei area and the cell area (NCR), the rescaled distance between the center of mass of the cell nucleus and the cell body (CMD) and the rescaled standard deviation of the contact length with neighbors (Figures 5.2 A and 5.3 A).

**3D considerations**  So far, we have considered only 2D projections of the cells and nuclei shapes. However, we need to check if the cells with larger areas are indeed larger. For MDCK cells the average cell monolayer height was measured as a function of the distance from the leading edge [62] on hard PA gels. The study showed that the MDCK-II monolayer height changes monotonically from zero at the border to approx. 6\(\mu\)m in the first 40\(\mu\)m. From 40 \(\mu\)m till 140 \(\mu\)m from the cluster border the height stayed approximately constant. The study on MDCK volume fluctuations on glass substrates [42] measured the average cell height to be 7.1 ± 0.7 \(\mu\)m but the average cell density in that study was relatively constant in the
Figure 5.2: Morphological measures that change with the mean cell density of the MDCK monolayer. (A) Cross-correlation coefficients between 17 chosen measures and the mean cell density. Some of the measures are strongly correlated with the cell density (red color), some measures are weakly correlated (blue color), and some are very weakly correlated (black color). (B) The cell and nuclei elongation are examples of the measures that are very weakly correlated with the cell density. However, the values on very low cell densities differ from the average values of these measures.
Figure 5.3: Morphological measures that change with the average cell density. (A) The ratio between the areas of nuclei and cells (NCR), the rescaled standard deviation of the contact length with the cell neighbors, and the rescaled distance between the center of mass of the cell body and its nuclei (CMD) are strongly correlated with the average cell density (correlation coefficient larger than 0.8). (B) The standard deviation of the rescaled cell area, the average rescaled cell perimeter and the standard deviation of the number of neighbors are correlated with the average cell density but the correlation coefficient is $\approx 0.65$. 
middle range of cell densities.

We measure the cell height for each cell (not averaged over the field of view) to evaluate the influence of the cell area on the cell height and the fluctuations between the cells. The height is estimated by hand from F-actin and Hoechst stained images of cells imaged with confocal microscopy (100x and 63x objectives, 0.2518 µm vortex depth).

We plot the cell height as a function of the cell area (Figure 5.4 A). We see that there are large fluctuations from cell to cell. The height fluctuations of the cells and their nuclei can also be evidenced by looking at the 3D reconstruction of the cell monolayer (Figure 5.4 E,F). However, it seems that smaller cells are on average higher. The cell height is estimated as the highest confocal plane with the signal from the cell nuclei or actin (see the inset in Figure 5.4 A). For the cells represented with the green dots the nuclei area and height were measured (Figure 5.4 B). We see that the nuclei height shows no clear dependence on the nuclei area. The cell volume is estimated as the height times the area even though this approximation leads to a systematic overestimate due to the nature of the cell shape. The data indicates that larger cells have larger volumes (Figure 5.4 C). In the cells where the nuclei occupy larger area fractions, they occupy larger volume fractions as well (Figure 5.4 C) but the values differ. For the analyzed cells, the cell nuclei do not occupy more than 35% of the total cell volume.

Therefore, we show that the results in 2D hold in 3D since the heights of both the cell and its nucleus vary insignificantly compared to the variations in the area. The cell height varied between 2.5 µm and 8 µm, while the cell area varied an order of magnitude (100 - 1400 µm²). Therefore, the cells with a bigger area have a larger volume, and larger area fractions of the cell nuclei indicate larger volume fractions. It seems that cells adjust their heights with their neighbors. However, due to the large fluctuations, more precise and numerous measurements are needed to reveal the exact dependence of the cell volume on the cell area.

5.3 Comparison with packings of random ellipses

We have seen that the cell monolayer reorganizes as the cell density increases, but the main mechanism remains unclear. In this section we try to understand the origin of this reorganization. As shown in Figure 5.3 with tissue densification the nucleus occupies a larger and larger fraction of the total cell body. When the cell density is low the nucleus occupies about 20% of the cell body but once the steady state density is reached the nucleus occupies almost 70% of the cell area. The correlation coef-
Figure 5.4: Changes of cells and nuclei heights. (A) The cell height as a function of the cell area. The green colored dots correspond to the cells for which the nuclei volume was also measured. The schema illustrates typical cell and nuclei shapes. (B) The nuclei height as a function of the nuclei area. (C) The cell volume (height times area) increases with the cell area. (C) The area fraction of the cell nucleus in a cell as a function of the volume fraction. (D) The image shows 3D reconstruction of a confocal image. The cell nuclei heights in a dense monolayer vary from cell to cell. (E) In the same field of view, the actin signal from the upper part of the monolayer shows variations in the cell height.
icient between the average cell density and the average NCR was 0.93. This happens because the cell nucleus is relatively stiffer than cell cytoplasm. Therefore, we look at the cell monolayer as a 2D packing of its nuclei where the mean nuclei-cell ratio (NCR) plays the role of the area fraction of a packing. The first question to answer is: which morphological measures change as a direct consequence of the reordering due to the increase of the packing fraction.

To simplify the problem, we approximate each cell nucleus with the best fitting ellipse and compare the organization of the MDCK monolayer with the organization of a random packing of ellipses representing the cell nuclei. Furthermore, we investigate a packing of mono-disperse ellipses at various area fractions to see the influence of exclusively the area fraction of a packing (average NCR) on the cell reorganization.

5.3.1 Packing of uniform ellipses

To test how the increase of the area fraction of ellipses influences the reorganization of SVT cells we pack 200 uniform ellipses. The packing of non-overlapping ellipses is done by house developed Phyton routine based on the algorithm presented in Figure 5.5. The ellipses have an elongation of 1 (circles), 1.25 and 2. The box size and the number of ellipses are fixed while the size of the ellipses is changed to achieve the area fractions of 0.2, 0.35, 0.5, 0.65 and 0.8. The random packing is repeated 25 times for each cell elongation and area fraction. From the ellipse positions we construct the SVT and analyze the morphological features of the obtained Voronoi cells (see Figure 5.6).

The results point to strong correlations between all the observed morphological measures and the area fraction of a packing. The value of the correlation coefficient is very weakly dependent on the elongation of the packed ellipses (Figure 5.7 A). Generally speaking, the packing reorganizes in such a way that all the Voronoi cell measures decrease with an increase in the area fraction except the number of neighbors which always increases. The only measure which is strongly dependent on the shape of the packed particles is the Voronoi cell elongation. At an area fraction of 0.2 the average Voronoi cell has an elongation of about 1.5 but as the packing fraction increases the elongation either increases or decreases depending on the shape of the packed particles. At an area fraction of 0.8 the cell elongation adopts the elongation of the packed particles (1.08 ± 0.06 for circles, 1.24 ± 0.07 for ellipses with elongation 1.2, and 1.9 ± 0.1 for

1The routine is developed by Jakov Lovrić from the PULS group
Figure 5.5: Algorithm scheme for the packing of random non-overlapping ellipses (reproduced from [156]). The algorithm is very similar to one published previously [157]. The difference is that in this routine we move and rotate a single randomly selected ellipse (instead of all of them) until a zero overlap criterion is met.
ellipses with elongation 2).

Even though the value of the correlation coefficient does not depend on the elongation of the packed ellipses, the morphological measures dependent on the area fraction in a different way. For example, the rescaled perimeter of the Voronoi cell has a strong anti-correlation with the NCR for all the packed ellipses. However, the plot of the average rescaled perimeter as a function of the area fraction of the packing (Figure 5.7B) reveals a different function behavior depending on the ellipse elongation.

The very strong correlations between all the morphological measures and the average NCR (area fraction) in mono-disperse systems show that the tissue reorganization is a consequence of an area fraction increase. However, the correlations are not so strong in the tissue.

5.3.2 Packing of randomly distributed cell nuclei

Cell nuclei have variations in shapes and areas. In other words they come from a certain non-degenerate probability distribution that changes with the cell density (Figure 5.8). This distribution alone can influence the value of the correlation coefficients. Therefore, we test the effect of this distribution and compare the organization of randomly distributed cell nuclei with the cell organization in the tissue.

As already explained, cell nuclei have elliptical shapes and we represent them with the best fitting ellipse. The best fitting ellipses are the ellipses with long and short axes having the same norm and orientation.
Figure 5.7: Changes of the Voronoi cell morphology with packing area fraction of mono-disperse ellipses.  (A) The cross-correlation coefficient between the area fraction of a packing and various morphological measures for ellipses with elongations of 1 (circles), 1.25 and 2.  (B) The average rescaled perimeter and elongation of cells as a function of the area packing of ellipses with elongations of 1.25 and 2.
Figure 5.8: Distributions of nuclei areas (A) and elongations (B). The low density data is collected from three images with a cell density of 2,500 ± 120 cells/mm² (675 cells in total) while the steady state density is represented with three images with a density of 7,400 ± 80 cells/mm² (1500 cells in total).

as the two orthogonal moments of inertia (second moments)\[109\]. Those values were calculated from the segmented nuclei images using the “regionprops” MATLAB function. The long and short ellipse axes and the total area covered by the nuclei are the input parameters for the packing algorithm presented in the Figure 5.5. We randomly pack non-overlapping cell nuclei found in 47 different 434 x 338 µm regions. For each tissue image we produce nine corresponding independent random packings.

The correlation coefficients between NCR and the other morphological measures are much different in the case of the non-uniform ellipse packing (Figure 5.9) than in the case of the uniform ellipses (Figure 5.7) with most of the measures having significantly lower correlations with the mean NCR in a packing of non-uniform ellipses.

Comparison of the correlation coefficients between the mean NCR and other morphological measures obtained from MDCK tissue and a random packing of ellipses representing the cell nuclei reveals many similarities between tissue and the random packing. However, there are also some differences. The most evident difference is in the case of standard deviations of NCR and CMD that have the opposite trend in the tissue and the random packing of the tissue nuclei. The differences are also pronounced for the standard deviation of the rescaled perimeter, the mean cell elongation and the standard deviation in cell elongation (Table 5.9 A). Plotting the mean values of these measures as a function of the mean NCR reveals small differences that are more pronounced for low cell densities when this set of measures has slightly larger values in a random packing of cell nuclei. However, these slightly larger values are sufficient to influence sig-
Figure 5.9: Comparison between the reorganizations of morphological measures with the area fraction of a packing in the case of tissue and random packing tissue nuclei. (A) Pearson’s correlation coefficient between the average area fraction of a packing (NCR) and the average value of cell morphological measures. (B) The graphs show the average values of the mean cell elongations and standard deviations of the cell elongations, the rescaled perimeters and the ratios between the nuclei and cell areas (NCR) as a function of the mean NCR. These measures depend on the average NCR in a different manner in the MDCK tissue and in a random packing of MDCK nuclei.
nificantly the correlation coefficients. The other 7 morphological measures have very similar correlation coefficients with the mean NCR for the cells from the tissue and a random packing of the cell nuclei.

**Distributions**

So far we commented propagation of the mean values of various morphological measures with the change in cell density or NCR but we haven’t considered the distributions of the morphological measures. To further compare the tissue and the random packing of the tissue nuclei we compare distributions of cell areas, cell perimeters, rescaled cell perimeters, and cell elongations in MDCK-II tissue and a random packing of cell nuclei. We plot distributions of those measures for 5 classes of cell density and represent them with at least three images with very similar cell density. We calculate the mean cell density and standard deviation of the cell density (average values in each image) within each class. The lowest cell density class (2500 ± 120 cells/mm$^2$) is represented with 695 cells. The medium cell densities are represented by the second class (3600 ± 90 cells/mm$^2$) with 1428 cells and the third class (4800 ± 50 cells/mm$^2$) with 1500 cells. The fourth class represents the steady state density (7400 ± 80 cells/mm$^2$) with 2478 cells. The last, fifth class is still dynamic region with cell density higher than steady state density in the bulk region (8000 ± 100 cells/mm$^2$) and it is represented with 2692 cells. Random packing distributions are obtained from nine independent random packings of the same cell nuclei.

We see that the area and perimeter distributions change with cell density (Figure 5.10) due to the cell size dependency of those measures. Contrary to the area and perimeter distributions, cell elongation and rescaled perimeter distributions change very little at different cell densities (Figures 5.11 A and 5.12 A).

All four morphological measures of the cells in the tissue are well reproduced by the Voronoi cells constructed from random packing of the tissue nuclei. However, the worst agreement is found at low cell densities (2500 ± 120 cells/mm$^2$) (Figures 5.10 5.11 B and 5.12 B). For the middle range cell densities, distributions from the random packing of cell nuclei also somewhat differ from the real tissue distributions, which is more evident in the case of cell elongation (Figure 5.11 C,D) and rescaled perimeter (Figure 5.12 C,D), while the cell area and perimeter (Figure 5.10) are estimated better.

At the steady state cell densities, the tissue is the most similar to the random packing of the cell nuclei in respect to all four morphological mea-
sures (Figures 5.10, 5.11 E and 5.12 E). Tissue organization at the densities higher than the steady state (8000 ± 100 cells/mm²) is very similar to the random packing in respect to cell elongation (Figure 5.11 F) and rescaled perimeter (Figure 5.12 F), but area and perimeter distributions differ (Figure 5.10). As explained in the third Chapter (see Figure 3.9), those extremely high densities are found in still dynamic ring between the bulk and the edge region.

We were interested how relevant are the observed differences between distributions of the morphological measures from the tissue and random packings. As already mentioned, for each tissue image we do 9 independent random packings of the same tissue nuclei and those packings slightly differ between each other. Therefore, we pick one representative image for low, high, and steady state density and calculate standard deviations of cell fractions in each bar for nine independent random packings of the same cell nuclei. For low and high cell density, we see that the probability distributions of cell area and cell perimeter in the tissue are mostly outside the errorbars of a random packing (Figure 5.13 A). Rescaled perimeter and cell elongation are mostly within the errorbars for high cell density, but cell area and perimeter differ for very high cell density class. Tissue organization at steady state density is the most similar to a random packing of its nuclei in respect to all four morphological measures (Figure 5.13 B). In mid-range cell densities (class two and three), most of the images have very similar probability distributions as random packing of their nuclei; however, large variations from image to image are present.

The Kolmogorow-Smirnow-test (KS-test) is a standard method used to test if two samples come from the same distribution. Comparing cell area and perimeter distributions obtained from the tissue and from a random packing with the KS-test gives density dependent results. For densities smaller that 3,000 cells/mm² (9 images), the equality of area distributions was not rejected in 89% of cases, while for cell densities larger than 6,700 cells/mm² (15 images) the equality of real and random distributions was rejected in 99.3 % of cases. These results contradict the visual comparison of the distributions, which revealed the best overlap of distributions at steady state densities.

The reason for this disparity lies in different statistics of the cells, because we do not have the same number of cells in the images of the tissue at different cell densities (the size of the image is fixed). The typical number of cells at lower cell densities is about 300, while at high cell densities there are more than 900 cells, and this alone influences the result of KS-test, which is sensitive to the number of cells in a sample. Therefore, conclusions based on KS-test can be misleading, due to sensitivity of this test on
Figure 5.10: Probability distributions of cell areas and perimeters for five classes of cell density constructed from several fields of view. Distributions from nine independent random packings of cell nuclei (full line) are plotted together with distributions from segmented images of the tissue (dashed lines).
Figure 5.11: Probability distributions of cell elongations for five classes of cell density. Distributions from nine independent random packings of cell nuclei (full line) are plotted together with distributions from segmented images of tissue (dashed lines and diamond scatter).
Figure 5.12: Probability distributions of the rescaled cell perimeter for five classes of cell density. Distributions from nine independent random packings of cell nuclei (full line) are plotted together with distributions from segmented images of the tissue (dashed lines and diamond scatter).
Figure 5.13: Probability distributions of cell areas, perimeters, elongations and the rescaled cell perimeter for cells from a monolayer which is still active (very low and very high cell densities) (A) and in steady state (B) and for Voronoi cells from nine random packings of the same cell nuclei.
5.3.3 Origin of the high correlation between NCR and cell density

During growth, epithelial monolayers simultaneously densify and invade surrounding space. This growth implies that forces are both exerted by a cell and transmitted to a cell by surrounding cells. As more and more cells occupy the same space, the pressure increases and each cell in the tissue is compressed. This leads to increase of nuclei-cell area ratio (NCR) since cell nuclei are more rigid than the cytoplasm.

To estimate how much stiffer cell nuclei are than cell cytoplasm we assume that the uncompressed shape of a cell is the one which cells have when a monolayer just forms (cell density of approx. 1,300 cells/mm$^2$) and that cells still do not exert stress on each other at such low cell densities. We use a set of 57 images with a density larger than 2,000 cells/mm$^2$. We also assume that the stress on each cell is axially symmetric.

We follow the changes of the average cell shape as cell density increases. We characterize the average shape of a cell and its nuclei with the average cell and nuclei area (see Figure 5.14). Furthermore, for cell nuclei we characterize the average length of two moments of inertia. If stresses are indeed axially symmetric, the strain of two moments of inertia must be the same. In the Figure 5.15 we see that average strain values really are very similar in both directions.

The mean of the total stress is a negative pressure, and we assume that
Figure 5.15: Cell density dependent average strain of cell nuclei in the direction of major (red) and minor axis (blue) of inertia moments reveals isotropic stress.

Figure 5.16: Strains as a function of cell density. (A) Average area strain of cell nuclei. (B) Average area strain of cell cytoplasm.
the pressure is a linear function of the cell density. For the elastic material the pressure can be related to the total area strain $\varepsilon = \varepsilon_{xx} + \varepsilon_{yy}$ with a single constant called the bulk modulus. If we have plane stresses the equations in 2D read:

$$\sigma = \frac{E}{2(1 - \nu)} \varepsilon_{area}$$

Where $\varepsilon$ is strain, $\sigma$ stress, $E$ elastic modulus and $\nu$ the Poisson’s ratio. This representation is invariant to the axis of rotation in the plane direction.

If MDCK-II cells are elastic material the strain has to linearly depend on the stress. The strain of the cell nuclei changes approximately linearly with the cell density (Figure 5.16 A), while the cell cytoplasm does not show linear behavior (Figure 5.16 B).

Data published on other cell lines reports that nucleus is 2 to 10 times stiffer than the cytoplasm [158]. In our data set, the ratio of the “bulk modulus” of cell nucleus and the “bulk modulus” of cytoplasm estimates that cell nucleus is $2.8 \pm 0.8$ times stiffer than cytoplasm. However, our estimate is based on the assumption that cells constitute a linearly elastic material, which is not true for the cytoplasm and is only approximately true for the cell nuclei.

### 5.3.4 Discussion and Conclusions

We measure reorganization of MDCK-II tissue with 17 morphological measures. We find that some measures strongly correlate with the cell density, while others have very weak correlation with the average cell density. To understand this result we compare tissue organization with organization of different random packings of ellipses.

As the tissue densifies, stresses on each cell increase. Consequently, cell nuclei occupy a larger fraction of cell area (NCR increases). If we consider the tissue as a packing problem, where cell nuclei represent the packed particles, NCR would correspond to area packing fraction. Therefore, we investigate reorganization in a random packing of ellipses with the increase of NCR. The reorganization is quantified with the morphological measures of the Voronoi cells constructed from the ellipse shapes.

In the case of ellipses with uniform size and elongation, all morphological measures strongly correlate with NCR. Specifically, all measures decrease with an increase of NCR except the mean number of neighbors which increases. Therefore, reorganization of the tissue should not surprise us since this is a necessary consequence of the fact that the mean NCR increases as the tissue densifies.
However, in tissues, cell nuclei have a distribution of areas and elongations and this needs to be taken into consideration. To test the effect of poly-disperse sizes and elongation of the packed ellipses, we randomly pack the ellipses with the same area and elongation as the cell nuclei in a given image of a real tissue. We find that the correlations with NCR are much weaker for randomly packed poly-disperse ellipses.

However, correlations with NCR in the real tissue, in the case of some morphological measures, are even weaker than in random packing of cell nuclei. An example of such measure is cell and nuclei elongation. For very low cell densities (less than approx 2,300 cells/mm$^2$), cell nuclei are more round and the cells themselves are more elongated. However, for all cell organizations in the monolayers with density larger that 2,300 cells/mm$^2$, this measure has the same value. Cells always have an elongation of $1.37 \pm 0.03$, while Voronoi cells from a random packing of tissue cell nuclei have strong correlation between cell elongation and mean NCR. This suggests an active mechanism which keeps the cell elongation constant during the tissue growth. It is to be explained why the cells actively organize in such a way that some of the morphological measures, like cell elongation, are kept constant.

Probability distributions of cell areas, perimeters, rescaled perimeters, and elongations revealed that on all cell densities the tissue has noticeably similar organization as the random packing of tissue nuclei. However, cell organization differs on different cell densities. In the case of lower cell densities the difference between tissue and random packing is the most pronounced. In the high density ring between the bulk and the edge region which is still a dynamic region, cells have very similar elongation and rescaled perimeter as the random packing but cell area and perimeter differ. Interestingly, cell organization in the steady state on the hard gels is the most similar to the random packing of the cell nuclei. The steady state is also the only configuration where the cells are immobile.

Even though we observe remarkable similarities between the random systems and the real epithelial tissue, we are still left with the question why the cell nuclei areas and elongations are the way they are. The distribution of the cell areas and elongations must come from the elastic properties of the cell and the forces that cells feel and exert. Therefore, understanding their origin means understanding the complexity of the cell’s mechanical properties and forces in a tissue.
5.4 Appendix: Image segmentation errors and performance of CMVT method on other substrates

Comparison with other nuclei segmentation routines on all substrates
We compare our segmentation procedure to ImageJ and CellProfiler, which are the two commonly used open-source software toolboxes for image analysis of cells: ImageJ officially recommends a specific procedure for the segmentation of DAPI stained nuclei. Accordingly, first a Gaussian blur filter should be used, before generating a binary image. Then, watershed is applied to a distance-transformed binary image. Since the Gaussian blur filter did not yield satisfactory results, we changed this approach to first subtracting the background (rolling ball radius of 60 pixels). The resulting image is binarized and watershed is applied as suggested by the developers. CellProfiler offers a segmentation tool that can be self-tailored to the properties of the objects expected in the image. This segmentation procedure includes methods for distinguishing clumped objects (segment objects) and methods for drawing lines between clumped objects (properly separate segmented objects). We obtain the best segmentation results with the so called “Intensity” method in both cases. In the case of distinguishing clumped objects, the “Intensity” algorithm counts each intensity peak as a separate object and is particularly suitable for objects that are brighter towards their interiors. For drawing lines between clumped objects, the ”Intensity” method performs a watershed on the original image starting from previously identified local maxima as seeds. Final thresholding of the image is performed using the Otsu method. The performance
of these two segmentation approaches is compared to the self-developed procedure described in the manuscript in the section 2.1.3 on images of Hoechst stained images (325.3 x 434.2 µm, pixel size 0.31 µm) of nuclei. The comparison is performed on images of cells grown on hard polyacrylamide (PA) gels (E > 3kPa), soft PA gels (E = 0.6 ± 0.2 kPa), and on glass. Over-segmentation and under-segmentation errors (Figure 5.17) are counted manually.

**Cells on hard PA gels** The tested set of data consists of ten images with various cell densities. The set is drawn from the original set of twenty-three images. Approximately 10,000 cells are analyzed with our method, ImageJ, and Cell Profiler. Oversegmentation and undersegmentation errors (Supplementary Figure 1) were manually counted. Results show that our method is prone to less errors than ImageJ and CellProfiler (Supplementary table 1) with a total error of 0.65%. The two watershed based algorithms had errors of 1.53% and 2.28%. ImageJ is significantly more prone to oversegmentation errors, which results in overestimation of an overall cell density. CellProfiler on the other hand is much slower than our method and ImageJ and tends to undersegment cells.
Cells on soft PA gels We analyzed the images used in the analysis. Those three images contained about 5,000 cells. Since the cells and their respective nuclei are very densely packed on the soft PA gels (E = 0.6 ± 0.2 kPa), all three methods have the largest errors for these conditions. However, our segmentation method was superior to ImageJ and CellProfiler Table 5.19 with a total error of 1.02% in comparison to 3.54% and 2.51%. CellProfiler tended to do more under segmentation errors same as in case of the main data set. ImageJ, on high densities, is equally prone to under and over segmentation of the cells’ nuclei.

Cells on glass substrate Six images used for analysis were tested for segmentation errors. In total, approx. 5,000 cells are analyzed. Those images had the smallest cell density of 5,200 ± 400 cells/mm², and therefore all three methods performed the best on those images. Our method was again better than CellProfiler and ImageJ (Table 5.20). CellProfiler tends to over segment images of cells grown on glass, in contrast to results on hard PA gels. ImageJ has the same tendency to over and under segment images as on soft Pa gels. Our method, on the other hand, has a tendency to over and under segment cell nuclei on all substrates leading to very small error in density estimation on all images.
Error of the membrane segmentation  Counting segmentation errors
Same as in the case of the nuclei segmentation, we have counted manually
the number of under and over segmentation cases in the segmentation of
our β-catenin images. The counting is done on all images used in the anal-
ysis: twenty-three in the case of hard PA gels, three on soft Pa gels, and
six images of the cells grown on glass substrate. Data is presented in the
Table 5.21. We see that for our main set of data, the segmentation error
was 1.61%. On soft PA gels, segmentation is much more precise (0.73% of
errors), while on glass substrates segmentation of β-catenin images leads
very often to over segmentation errors (6.01% of over segmentation er-
rors).

<table>
<thead>
<tr>
<th>substrate</th>
<th>Undersegmentation</th>
<th>Oversegmentation</th>
<th>Total error</th>
</tr>
</thead>
<tbody>
<tr>
<td>soft PA gels</td>
<td>0.42</td>
<td>0.31</td>
<td>0.73</td>
</tr>
<tr>
<td>hard PA gels</td>
<td>0.68</td>
<td>0.93</td>
<td>1.61</td>
</tr>
<tr>
<td>glass</td>
<td>0.91</td>
<td>6.1</td>
<td>7.01</td>
</tr>
</tbody>
</table>

Figure 5.21: Membrane segmentation errors on all substrates.

Consistency of the error estimations We compare consistency of man-
ually counted estimates of the segmentation errors with automatic count-
ing of the cells removed from the statistics because of our criteria (see sec-
tion 3.1 in the manuscript), and find that numbers match. In the case of
our main data set (hard PA gels), from total number of cells removed from
the statistic, 70% had one to one correspondence between membrane and
cell’s nucleus, but they did not satisfy the criteria that 95% of the nucleus
is in the membrane and the membrane is not occupied by other nucleus by
more than 5% of the nucleus area. The remaining 30% of membranes re-
moved from the statistic did not have one associated cell nucleus (20% had
no nucleus inside and 10% had more than 1). Since in total 15.6% of mem-
branes are eliminated, we see that 4.7% of total number of membranes are
eliminated because one to one correspondence could not be established.
From that number, 1% of the membranes are eliminated due to the differ-
ence in the statistics of the segmented membranes and nuclei, which arises
from slightly different elimination criteria of the cells at the edge. The rest
of the 3.7% is consistent with our table of manually counted errors in over
and under segmentation of cells’ nuclei and cells’ membranes. The total
number of errors is 3.3% (nuclei over segmentation and membrane un-
der segmentation need to be counted twice), but slightly lower number is
expected since cell divisions are not counted as segmentation errors.
Choosing the correct value of the suppressed minima in h-minima transform Since the watershed method is known to be prone to over segmentation, we preprocess the images with the h-minima transform to suppress all local minima smaller than the critical value. This critical value is chosen by visual inspection, but we check for small variations around this value. We find that, with the critical value of suppressed minima being 35, the segmentation procedure finds 374 cells more than with a critical value of 40. If critical value of suppressed minima is set to be 45, the segmentation procedure finds 272 cells less than with our chosen value of 40. The total statistic of recognized cells with a critical value of suppressed minima set to 40 is 17,850 cells. We compare the distribution of cell areas gained from those three different segmentation results and we find that those data sets differ significantly (Figure 5.22). For example, choosing 35 as a critical value of the suppressed minima leads to recognition of much smaller cells, and Kolmogorov-Smirnov-Test (KS test) rejects the null hypothesis that those cell areas come from the same distribution as the cell areas obtained with choosing 40 as a critical value of the suppressed minima. However, if we eliminate segmentation errors with our criteria (see section 2.1.5), the K-S test fully accepts that those distributions are the same. It is our criteria that ensures elimination of all the segmentation errors (segmentation errors due to our image analysis methods and inherent errors due to relative positions of β-catenin and nuclei signal) in the set of data used for our
comparison between morphological features of the cells’ membranes and Voronoi tessellations based on the centers of mass of the nuclei (CMVT). Therefore, the data set is not sensitive to the small variations around chosen value of 40 for the critical value of the suppressed minima, because all the segmentation errors are eliminated by our criteria.
Segmentation procedure for MDCK II monolayer grown on hard PA gels  Segmented β-catenin images on hard PA gels.

Figure 5.23: Segmentation of the representative picture (out of 23 pictures) showing the whole 325.3x434.2 µm field of view. (A) Segmentation of the cells’ nuclei. Red outlines represent edges of the cells’ nuclei. Big white-red dots mark the cells eliminated from presented data set because they did not satisfy the criteria. Small dots represent centers of mass of the cell nucleus and the corresponding cell membrane. Those cells enter the data set presented in the 2.1. (B) Segmented membrane. White-red dots represent cells eliminated due to our criteria. Bars are 20 µm.

Figure 5.24: Zoom in the regions of interest. (A) Zoom in upper middle part of the representative picture. Red lines represent edges of segmented cell nuclei, and cyan lines represented segmented membrane. Big white-red dots are indicating cells eliminated from statistics due to the criteria that minimally 95% of the cell’s nucleus belongs to the corresponding cell’s membrane, which is not occupied by more than 5% of the neighboring nucleus. (B) Zoom in down right corner of the representative picture where all the cells have satisfied the criteria. Yellow lines represent CMVT approximation of the real cell shape, while white lines represent the cell’s membranes on the set of the cells that contribute to the data set analyzed in the paper. In total 15,000 of membranes have been compared with the corresponding CMVT in respect to six morphological measures. Bars are 20 µm.
Performance of CMVT on soft 0.6 kPa polyacrylamide (PA) gels  Comparison between CMVT and segmented β-catenin images on soft PA gels.

Figure 5.25: Representative picture of MDCK II monolayer on soft (0.6 ± 0.2 kPa) PA gels. A) Hoechst stained cell nuclei. B) β-catenin stained cell membrane. Scale bars are 20 µm.

Figure 5.26: Histogram of cell area of MDCK II on soft PA gels. Total number of cells is 3,097 while 2,937 cells belong to subsets. Average cell size is 80 ± 18 µm².
Figure 5.27: Comparison of CMVT (yellow squares) and directly extracted morphological measures (blue circles). The graphs associated with cell areas, perimeters and number of neighbors are shown in the first, second and third column, respectively. Top graphs (A-C) show the probability distribution generated by direct measurement and estimated from the tessellation. (D-F) Second row is the average dependence of morphological characteristics on the cell area. The associated deviations of CMVT are shown in the third row (G-I). The distributions of errors for small, intermediate-sized and large cells are shown in the bottom row (J-L).
Figure 5.28: Comparison of CMVT (yellow squares) and directly extracted morphological measures (blue circles). The graphs associated with cell elongation, standard deviation of boundary length and co-alignment of the nuclei and the cell are shown in the first, second and third column, respectively. Top graphs (A-C) show the probability distribution generated by direct measurement and estimated from the tessellation. (D-F) Second row is the average dependence of morphological characteristics on the cell area. The associated deviations of CMVT are shown in the third row (G-I). The distributions of errors for small, intermediate-sized and large cells are shown in the bottom row (J-L).
Performance of CMVT on glass substrate  Comparison between CMVT and segmented $\beta$-catenin images on glass substrates.

Figure 5.29: Representative picture of MDCK II monolayer on glass substrates. A) Hoechst stained cell nuclei. B) $\beta$-catenin stained cell membrane. Scale bars are 20 $\mu$m.

Figure 5.30: Histogram of cell area of MDCK II on glass. Total number of cells is 2,576, while 2,335 cells belong to subsets. Average cell size is $196 \pm 52 \, \mu m^2$. 
Figure 5.31: Comparison of CMVT (yellow squares) and directly extracted morphological measures (blue circles). The graphs associated with cell areas, perimeters and number of neighbors are shown in the first, second and third column, respectively. Top graphs (A-C) show the probability distribution generated by direct measurement and estimated from the tessellation. (D-F) Second row is the average dependence of morphological characteristics on the cell area. The associated deviations of CMVT are shown in the third row (G-I). The distributions of errors for small, intermediate-sized and large cells are shown in the bottom row (J-L).
Figure 5.32: Comparison of CMVT (yellow squares) and directly extracted morphological measures (blue circles). The graphs associated with cell elongation, standard deviation of boundary length and co-alignment of the nuclei and the cell are shown in the first, second and third column, respectively. Top graphs (A-C) show the probability distribution generated by direct measurement and estimated from the tessellation. (D-F) Second row is the average dependence of morphological characteristics on the cell area. The associated deviations of CMVT are shown in the third row (G-I). The distributions of errors for small, intermediate-sized and large cells are shown in the bottom row (J-L).


179


