DISSERTATION ZUR ERLANGUNG DES AKADEMISCHEN GRADES

DOKTOR DER NATURWISSENSCHAFTEN

(DR. RER. NAT.)

Dynamics and Structure of Cellular Aggregation

Dem Fachbereich Physik der Philipps-Universität Marburg vorgelegt von

Patrick Bitter (geb. 19.03.1986 in Arolsen)

am 17.12.2014

Zusammenfassung

Die vorliegende Arbeit behandelt die Dynamik und Struktur der Aggregation von Zellen. Dabei spannt sie den Bogen von der Beweglichkeit auf der Ebene von Einzelzellen als Vorstufe von Aggregation zu deren Ergebnis, der Struktur der Zellaggregate.

Im Bereich der Beweglichkeit von Mikroorganismen gibt es viele gut untersuchte Systeme. Als Beispiele seien hier das Schwimmen des Bakteriums *E. coli* und das Gleiten der Amöbe *D. discoideum* genannt. Allerdings gibt es nur wenige Arbeiten zur Beweglichkeit von Säugerzellen, insbesondere zu deren Reaktion auf externe Einflüsse. Um den Mechanismus dieser Zellmigration besser zu verstehen, habe ich ein generisches Modell entwickelt, das ausschließlich auf der mechanischen Wechselwirkung zwischen Zelle und Substrat basiert. Das Modell macht sich dabei die diskrete Natur des Bewegungszyklus von Säugerzellen zu Nutze. Die Grundidee ist ein zufälliges Auswerfen von Armen und ein Zusammenziehen, das von den externen Einflüssen, wie z.B. der Konzentration eines Pheromons abhängt. Dieses Modell ist in der Lage die meisten experimentellen Beobachtungen zu reproduzieren, dabei zeigt es insbesondere gute Ergebnisse für das Übertreten von scharfen Stufen in einer heterogenen Umgebung. Außerdem erklärt es das Festhalten des hinteren Zellpols an der Substrato-

berfläche: Die Fähigkeit einem Gradienten zu folgen wird dadurch stark verbessert.

Der zweite Teil der Arbeit beschäftigt sich mit der quantitativen Analyse von Aggregaten. Dafür präsentiere ich neue experimentelle Methoden zur Visualisierung von Zellaggregaten und eine Zusammenstellung von Werkzeugen zu ihrer Analyse. Diese Ansätze wurden anhand von zwei Beispielen getestet: der Aggregation von Fibroblasten und der zuckersensitiven Flokkulation von *S. cerevisiae* .

Für die zweidimensionale Aggregation der Fibroblasten war es möglich durch Videos auch die Dynamik des Aggregationsprozesses zu untersuchen. Dabei kristallisierte sich eine Zweiteilung des Prozesses heraus: Zuerst gibt es eine Phase, während der die Zellen sich strecken und die nach einiger Zeit in eine Kontraktionsphase übergeht. Außerdem zeigte sich, dass die Struktur der Aggregate (gemessen an deren fraktaler Dimension) abhängig ist von ihrer Größe. Dabei gibt es einen Übergang von einem linearen Anstieg der fraktalen Dimension mit der Aggregatsgröße zu einer konstanten fraktalen Dimension.

Im Fall der Flokkulation von *S. cerevisiae* gibt läuft die Dynamik dieses drei dreidimensionalen Prozesses zu schnell ab um sie durch Videoaufnahmen zu bestimmen. Jedoch war es möglich durch konfokale Laser-Scan-Mikroskopie dreidimensionale Bilder der fixierten Flocken zu produzieren und ihre Struktur zu untersuchen. Die wichtigsten Ergebnisse hier sind die fraktale, selbstähnliche Struktur der Flocken und die starke Benachteiligung von Nicht Produzenten gegenüber von Produzenten des für die Flokkulation nötigen Bindeproteins Flo5. Dadurch werden diese sogenannten "Cheater" stark benachteiligt und ihr Wachstumsvorteil durch einen Nachteil bei der Positionierung in der Flocke abgemildert oder sogar kompensiert. Dieses Ergebnis legt nahe, dass auch wenn flo5 nicht die wörtliche Definition eines "Green Beard" Gens erfüllt, doch eine ähnliche Bevorzugung von Trägern des gleichen Gens (bzw. von Produzenten des Bindeproteins) vorliegt.

Abstract

This work provides new insights into the dynamics and structure of cellular aggregation. Starting from cell motility which is necessary to get the cells into close proximity it presents new tools for visualization, analysis and modeling of aggregation processes.

While a lot of work has been done in the field of microbial and amoeboid motility, there is a lack in theoretical understanding of mammalian cell motion, especially concerning directed migration stirred by external cues. To close this gap I developed a two-dimensional generic model based on mechanical cell-substrate interactions. This model facilitates the discrete nature of the motion cycle of mammalian cells by a randomized growth of protrusions and their retraction depending on the strength of an external cue. This model is capable of reproducing most experimental observations, especially the behavior at sharp changes in strength of the external cues, and provides an explanation for the attachment of the lagging cell pole as it increases the efficiency of gradient sensing.

Furthermore, I introduce new experimental methods to visualize and analytical toolkits to analyze the structure of the highly irregular cell aggregates. These approaches were tested in two example cases: the two dimensional aggregation of mouse embryonic fibroblast (MEF)

cells and the flocculation of *S. cerevisiae* mediated by the sugar-dependent adhesion protein Flo5.

While it was possible to achieve temporal information of the MEF cell aggregation, the flocculation of *S. cerevisiae* is not accessible in this way.

The time-lapse microscopy series indicate a subdivision of MEF cell aggregation into a spreading and a contraction phase. In addition, the data shows that there is a dependency of the aggregate's structure on its size with a sharp transition from a linear dependency to a constant structure.

The three-dimensional imaging of immobilized flocs using a confocal laser scanning microscope provided information about the structural properties of yeast flocs. The most important findings are that the flocs are self similar fractal structures and that cheater cells, i.e. cells that do not produce the necessary binding proteins but benefit from the altruistic behavior of producing cells, are largely underprivileged in the process. This indicates that, even though flo5 does not qualify as a "green beard gene" by definition, the benefits of the resulting altruistic behavior are strongly shifted in favor of the producing cells by the aggregation mechanism.

Für Anna

Contents

Chapter 1

Introduction

Social behavior of organisms is a wide studied field in biology. Our everyday experience – that is based on higher animals and humans – teaches us that social interaction and altruistic behavior can be beneficial on many levels, reaching from individual to species wide benefits (Wilson and Wilson, 2007) and is an increasingly important factor in our strongly interconnected community (Perc et al., 2013). However, it lacks insights into the sociobiology of microorganisms that are detracted from our own senses.

A very important example of social behavior on a microscopic level is aggregation. The formation of biofilms as reviewed by Nadell et al. (2009) for example is a phenomenon of great medical importance, since it possibly mediates drug and other resistances. Flocculation of *S. cerevisiae* on the other hand is a form of aggregation with a huge biotechnological impact (Bauer et al., 2010).

In order to form aggregates the cells need a way to get into close proximity so that they can

bind and aggregate (Flierl et al., 1999). To achieve this cells either rely on passive transport in a liquid medium – like *S. cerevisiae* does – or evolved means of active motion (Lenz (2008) provides an overview of different types of cell motility). In higher organisms the spatial organization and interaction of cells is an important factor during embryonic development as it is a precursor for cell differentiation (Melton, 1991; Hardikar et al., 2003; Chandra et al., 2011; Govindasamy et al., 2011).

This work is structured as follows: Chapter 2 introduces well studied forms of cell motility. I developed a model that is capable of explaining and characterizing several details of the motility of mammalian cells. It is introduced in section 2.3.1. Having established the forms of cell motility that precede cellular aggregation, chapter 3 introduces a set of tools that make it possible to quantify patterns (see section 3.1). In order to analyze the physical mechanisms and interactions leading to aggregation, I developed an individual based model capable of simulating the dynamics and structure of aggregation patterns (see section 3.2). I applied the previously mentioned tools and the model to two different example cases: the aggregation of mouse embryonic fibroblast (MEF) cells (see section 3.3) and the flocculation of *S. cerevisiae* (see section 3.4). In close collaboration with out colleagues from the department of biochemistry at the University of Hong Kong, we were able to implement automated tools to analyze their time-lapse data on the two dimensional aggregation of MEF cells. In a SYNMIKRO[1] cooperation with working groups from physics, molecular genetics and computer sciences, I developed a new method for three dimensional imaging of *S. cerevisiae* flocs and the analysis of those images. We were able to implement the individual based model on a general purpose graphics processing unit (GPGPU) system which boosts the computational efficiency (see Leinweber et al. (2014)) and offers the possibility for realistic simulations. Chapter 4 summarizes the work and provides an outlook on further studies and questions arising from this work.

---

[1]LOEWE Center for Synthetic Microbiology, see http://www.synmikro.com

1.1   Social Behavior on the Microscopic Level

As it turns out microorganisms show many variations of social behavior. Zak et al. (1994) analyzed the functional diversity provided in microbial communities, where genetic diversity translates to a functional diversity and in this way mediates a cooperation across the species border. A very prominent example for microbial social behavior at a single species level is *D. discoideum*. Strassmann et al. (2000) described the altruistic life cycle of this social amoeba. If nutrition is depleted the cells start aggregation and sporulation in order to spread their spores as far as possible from the position of the depleted nutrition.

The first theoretical concept for the evolution of social behavior on a genetic level was developed by Hamilton (1964a,b). The basic idea of his model is that the genetic relation between two individuals is important for the appearance of social behavior. Performing an altruistic action to help a close relative can be beneficial on a genetic level and so ensure the persistence of the genetic heritage.

Dawkins (1976) introduced the term green beard gene for genes that mediate altruistic social behavior. The essence of his definition is the recognition and preferential treatment of organisms that carry the same gene. Foster et al. (2007) reviewed the fields of sociogenomics and sociomicrobiology and showed the importance of the genetic background on the emergence of social behavior. Queller et al. (2003) proved that the gene responsible to mediate the social aggregation of *D. discoideum* is one of these green beard genes.

Another well studied example of green beard genes is the gene flo11 present in the bakers yeast *S. cerevisiae* (Lo and Dranginis, 1996). This membrane associated protein encoded by this gene binds to other Flo11 proteins and mediates flocculation of gene carriers while prohibiting the inclusion of non-carriers (Brückner and Mösch, 2012).

1.2 Cell Motility

Motility is the ability of an organism to travel through the environment. Because it enables the organism to react to environmental cues, such as threads or food supply, it is a very important feature on all scales of life.

While our every day experience is guided by inertia as stated by Newton's first law, life on the microscopic level is dominated by viscous effects. The Reynolds number *Re* is a measure for the fraction of inertial forces to viscous forces. According to Purcell (1977), the Reynolds number for macroscopic objects is in the order of 10^2 to 10^4, so inertial forces dominate the motion. On a cellular level on the other hand organisms experience Reynolds numbers in the order of 10^{-4} to 10^{-5}, so inertia is completely irrelevant. Since the Navier-Stokes equation becomes independent of time in the limit of small Reynolds numbers, Purcell (1977) argued that motion on this scales requires the organisms to propel themselves with a persistent cyclic but non-reciprocal motion, a motion that is different when time is reversed.

If the cell's motion is guided by external cues into a defined direction rather than a random motion, this state of motility is called directed cell motion or cell migration (Macnab and Koshland, 1972; Tindall et al., 2008a; Levine and Rappel, 2008). According to the extracellular cues guiding the motion, the motility can be categorized into one of three different groups: (1) Directed motion guided by solvent chemical cues is called chemotaxis (Macnab and Koshland, 1972). (2) Migration biased by substrate-bound cues influencing the cell-substratum adhesiveness is called haptotaxis (Carter, 1965). (3) If the cell motility is guided by mechanical cues such as substrate rigidity the directed motion is called durotaxis (Lo et al., 2000).

In case of mammals cell motility is of great importance in tissue invasion of pathogenic microorganisms, inflammation and immune response (Chicurel, 2002; Gray et al., 2003; Prost et al., 2008; Parsons et al., 2010), cancer progression, invasion and metastasis (Lo et al., 2000;

Chicurel, 2002; Gerisch and Chaplain, 2008; Bordeleau and Galarneau, 2010; Parsons et al., 2010), embryonic development (Chicurel, 2002; Gray et al., 2003; Huttenlocher and Horwitz, 2011), and other diseases (Wong et al., 2003).

1.2.1 Microorganisms

Microorganisms use a broad bandwidth of methods for active propulsion that are of the cyclic non-reciprocal kind Purcell (1977) proposed. Most of them use flagella, pili or cilia (see section 2.2.1) to swim through liquid media[2]. On solid surfaces according to Keren and Theriot (2008) the crawling modes are used by cells to propel themselves (see section 2.2.2).

The most prominent example of bacterial propulsion is the use of rotating flagella that take on the shape of a rotating corkscrew and in this way carry out a persistent cyclic and non-reciprocal movement (Purcell, 1977). Section 2.2.1 reviews this mode of propulsion.

The social amoeba *D. discoideum* is a well studied model system for motility and migration of eukaryotic cells (Fisher et al., 1989; Fuller et al., 2010). Small and Resch (2005) and Pollard and Borisy (2003) reviewed the molecular principles driving the motion in *D. discoideum*. After cell polarization actin polymerization and contraction causes the growth of protrusions at the front and at the same time retraction of the back of the cell (see figure 2.5) resulting in a cell speed of about $10\,\mu\mathrm{m/min}$. A detailed review of the amoeboid motion can be found in section 2.2.2.

1.2.2 Mammalian Cells

Mogilner (2009) described mammalian cell motility as a complex process integrating biochemical kinetics, transport, and physical forces that makes it necessary to use combinations of continuum mechanics and biochemical kinetics to model it.

---

[2]An overview of propulsion methods can be found e.g. in Alberts et al. (2002)

5

As in amoeba, the predominant mode of migration in mammalian cells is the crawling motion (Keren and Theriot, 2008). In contrast to amoeboid-like migration which is more floating, mesenchymal migration consist of a cycle of five discrete steps. The amoeboid-like migration is very similar to the migration of *D. discoideum* . The mesenchymal mode of migration of epithelial cells, e.g. mouse embryonic fibroblasts, is influenced by a variety of chemical and physical signals, in particular by different mechanical forces (Butcher et al., 2009) but shows a very high efficiency in following all kinds of external gradients (Theveneau et al., 2010).

The steps of the migration cycle are carried out within about ten minutes (Mogilner, 2009; Lauffenburger and Horwitz, 1996; Pathak and Kumar, 2011; Ridley et al., 2003; DiMilla et al., 1991; Zaman et al., 2005; Parsons et al., 2010; Huttenlocher and Horwitz, 2011; Chicurel, 2002; Keren and Theriot, 2008). In general, crawling is propelled by the active lamellipodeum at the leading edge that pulls the passive cell body forward (Mogilner, 2009).

In addition to chemical cues, stiffness and inhomogeneity of the substrate or the extracellular matrix (ECM) directly influence the direction of motion through durotaxtis as first proposed by Lo et al. (2000) and confirmed by Gray et al. (2003), Yeung et al. (2005), Provenzano et al. (2008), and Trichet et al. (2012). Besides the impact on migration itself, Holmes and Edelstein-Keshet (2012) showed that these factors also strongly influence cell shape.

According to Pathak and Kumar (2011) the interaction of the cell and the substrate or the ECM can be viewed as a bilinear sequential binding that exerts forces on the cell, predominantly on the nano scale. Through mechanosensing mechanisms as mentioned by Butcher et al. (2009) and Trichet et al. (2012) the cell is able to actively react to these mechanical cues. However, cells can also actively manipulate the ECM or the substrate, e. g. in case of cancer, where tumor cells stiffen the surrounding ECM through extracellular actin polymerization, and build a rigid stroma around the tumor[3]. Bordeleau et al. (2013) observed that this step in ECM stiffness promotes cells from the outside moving inside, but prevents cells from the

---

[3]for a review see (Butcher et al., 2009; Provenzano et al., 2008)

inside to migrate outside.

In general, mammalian cells actively react to external forces gradually by pulling or pushing the ECM or the substratum instead of being passive recipients of mechanical forces (Lauffenburger and Horwitz, 1996; Lo et al., 2000; Butcher et al., 2009).

The great complexity of cell motility makes it an interesting field of study for both experimental and theoretical scientists alike. Mogilner (2009) reviews a number of models published on this topic. So far, these models are focused on cell shape during migration (Rubinstein et al., 2005), continuum descriptions for cell populations (Gerisch and Chaplain, 2008; Häcker, 2012) or on special short-term aspects of motion like migration speed (DiMilla et al., 1991) or effective adhesiveness (Zaman et al., 2005).

However, there are a lot of unanswered question. In chapter 2 I introduce a simple, generic model for environment sensitive motion of fibroblast-like cells that is solely based on mechanics. It covers chemotaxis, haptotaxis, and durotaxis as modes of motility.

1.3   Cell Aggregation and Differentiation in Mammals

Cellular aggregation is one way to establish spatial patterns and therefore a first step towards cell differentiation. Pattern formation is an important feature at all scales of life and even beyond. The smallest biologically relevant level are protein folding and the establishment of spatial concentration patterns inside single cells like the spatio-temporal gradient of the Min-System in *E. coli* as reviewed by Lutkenhaus (2007) and the MipZ-System in *C. crescentus* as described by Thanbichler and Shapiro (2006) that guide the position of cell division.

In embryonic development it is very challenging to explain how a relatively small number of genes can result in such different animals, as Melton (1991) put it. This is where studying pattern formation (and hence cellular aggregation) helps to get a better understanding of the formation of life.

Pattern formation on the level of cell colonies is often triggered by cell-cell interactions. Liu et al. (2011) transformed a synthetic gene circuit into *E. coli* cells that connects the motility system of *E. coli* to a quorum sensing module. In this way the motility of the cells is inhibited if the cell concentration is higher than a threshold value. This coupling leads to stripe patterns that can be tuned experimentally as predicted in models introduced by Fu et al. (2012).

These model systems are less complex than mammals, but they help to study the basic principles under laboratory conditions. One outcome of these kind of studies is the finding that the majority of spatial concentration patterns can be explained by models based on a two-player system as introduced by Turing (1952). Patterns arising from this system based on diffusion of and reactions between the two players are called Turing patterns. Wolpert (1969) pointed out that the key component of this system is the combination of a short range activation (e.g. achieved by a self activation) and a long range inhibition (e.g. achieved by a fast diffusing inhibitor). Kondo and Miura (2010) mentioned the difficulty in proving a Turing mechanism is responsible for a biological pattern. In living cells interactions are very complicated (and

sometimes unknown). However, Turing mechanisms are able to reproduce the formation of stripes and dots in the animal skin studied by Kondo (2002), Koch and Meinhardt (1994) reviewed a number of systems in which they are used to simulate and explain patterns occurring in very complex biological structures.

This indicates that the underlying mechanisms in pattern formation can be relatively simple and at the same time produce a very complex spatial structure. It is also a very good example for the benefits of the interlock of experimental and theoretical methods to approach this topic.

As mentioned above, a very important type of pattern formation is aggregation. From a theoretical perspective diffusion limited aggregation (DLA) which is mainly used to describe crystallization processes is one of the best studied aggregation processes. Witten Jr. and Sander (1981) developed a model, where a cluster is seeded at the origin of a lattice. A particle is added to a random lattice site at a large distance from the origin and allowed to perform a random walk until it visits a lattice site adjacent to an occupied site. The particle becomes part of the cluster and a new particle is added at large distance and allowed to perform a random walk and so forth[4]. Vicsek (1984) showed that it is possible to generate various kinds of regular patterns using a slightly modified DLA model.

In biology, aggregating cells are an important example for pattern formation in mammals. Islet like cell aggregation as described by Chandra et al. (2011), Govindasamy et al. (2011), and Hardikar et al. (2003) is a very important example of cellular aggregation, since the resulting cell clusters can be used as a substitute for the damaged insulin producing cell islets in diabetes type one patients.

Schlessinger et al. (1978) was one of the first to use 3T3 cells (Todaro and Green, 1963) as a model system to understand the basic principles of the aggregation of fibroblast cells which is an important step in the formation of islet-like cell aggregates. In this work we use the similar STO strain (Martin and Evans, 1975) to study the aggregation behavior of Fibroblast cells.

---

[4]for a detailed analysis of this model see Witten and Sander (1983)

As will be introduced in chapter 2, MEF cells perform a crawling motion. After they drew near to each other they form bonds and contract to islet-like aggregates. As reviewed by Hynes (1987, 1992), these bonds are mediated via cell surface proteins called integrins, that are also of great importance in many other interactions of fibroblast cells, e.g. in cell motility (see chapter 2) or cell-surface interaction in general. In a more recent study Toret et al. (2014) elucidated the high genetic complexity of these integrins involved in cell-cell adhesions.

After introducing tools to quantify aggregation patterns in section 3.1 and a generic mathematical model for cellular aggregation of any type in section 3.2, section 3.3 shows the experimental results from time-lapse microscopy and the modeling results for MEF cell aggregation.

1.4 *S. cerevisiae* : A Model Organism to Study Social Behavior

S. cerevisiae is a standard model organism to study eukaryotic genetics. Commonly known as bakers yeast it has been used in technological applications for a long time (Legras et al., 2007). While flocculation is the most relevant lifestyle with respect to biotechnological application (Bauer et al., 2010) *S. cerevisiae* has two other important lifestyles besides its vegetative state.

1.4.1 Biofilms, Filaments and Flors: Different Lifestyles of *S. cerevisiae*

S. cerevisiae cells are capable of expressing a number of adhesive proteins that mediate different types of aggregation that can be characterized by the possible binding partners of the cell: foreign (e.g. cell-substratum), self (e.g. cell-cell) and mixed adhesions (see Brückner and Mösch (2012) for a review of the related adhesins). Besides flocculation, which is described in the subsequent section, these adhesions can lead to a variety of different lifestyles.

A prominent example are biofilms of *S. cerevisiae* cells that are mediated by a mixture of self and foreign adhesions (Reynolds and Fink, 2001) and occur at the interfaces between solid substrates and air or solid substrates and liquid. Another lifestyle that is mediated by a mixture of self and foreign adhesions is the filamentation of *S. cerevisiae* cells (Gancedo, 2001). In this state the cells build long filaments that adhere to the surface of a solid substrate and are able to invade this substrate.

Flor formation is different to biofilm formation and filamentous growth because it is mediated only by self adhesion instead of a mixture. Brückner and Mösch (2012) defined a *S. cerevisiae* flor as "an air- liquid interfacial layer of floating cells that are attached to each other and form a biofilm". This definition is based on the works of Cruess et al. (1938); Hohl and Cruess (1939); Freiberg and Cruess (1955); Martinez et al. (1995); Zara et al. (2005) that represent seven decades of research on the adhesive lifestyles of *S. cerevisiae*.

1.4.2 Flocculation

Flocculating *S. cerevisiae* cells have no means of active motion and are drawn near to each other by external flows. As reviewed by Soares (2011), Brückner and Mösch (2012) flocculation is defined as the asexual, reversible and Ca^{2+} dependent aggregation of typically thousands of vegetative cells into flocs. The biotechnological impact of *S. cerevisiae* strains makes flocculation an important field of study, because it is a cheap and effective way of removing cells from a desired product after fermentation processes. The adhesins mediating the bonds between vegetative yeast cells belong to the family of fungal glycosylphosphatidylinositol-linked cell-wall proteins, and are studied extensively, e.g. by Verstrepen et al. (2004); Verstrepen and Klis (2006); Dranginis et al. (2007); Linder and Gustafsson (2008). Dranginis et al. (2007); Verstrepen and Klis (2006) classified the adhesins into two general groups: they mediate either sugar-sensitive (called lectin-like) or sugar-insensitive adhesion.

Flo11 is an example for the sugar-insensitive adhesion, where the cell-surface protein is ligand and receptor at the same time. In this way, only cells expressing the protein can adhere to each other, mixed adhesions between expressing and non-expressing cells are not possible (Lo and Dranginis, 1996; Brückner and Mösch, 2012).

In this work I focus on the lectin-like adhesions mediated by the Flo5-protein. These sugar-sensitive adhesins bind to mannose-oligosaccharid chains on the cell surface and mediate adhesions between expressing (bearer) cells as well as adhesions between bearer and non-expressing (cheater) cells (Veelders et al., 2010). However, adhesions between bearer and cheater cells contain less molecular bonds than bearer-bearer adhesions and are thus thought to be less stable. The question at hand is if this difference in binding strength is enough to result in a preferential treatment of cells that are producing the protein over non-producers.

Chapter 2

Cell Motility

The ability to travel through the environment is an important aspect of life on all scales. It is necessary to get towards or away from environmental cues such as food sources, local climate, toxic components, or predators.

Our every day experience is mainly determined by effects of inertia. Newton's first law states, that a body will stay at rest or in uniform motion unless external forces act on it. On the macroscopic scale, swimming objects experience a relatively low viscous drag: If they stop their active propulsion, they will move on for some time.

However, on a microscopic level viscous effects become much more important. The ratio Re of inertial forces to viscous forces is called Reynolds number. For macroscopic objects it is in the order of 10^2 to 10^4, so the inertial forces dominate the motion. Microorganisms and cells on the other hand experience Reynolds numbers in the order of 10^{-4} to 10^{-5}, so inertia is completely irrelevant as Purcell (1977) pointed out. He also showed that the Navier-Stokes

equation in this limit requires the organisms to swim by a cyclic but non-reciprocal motion, a motion that is different when time is reversed.

Microorganisms developed different kinds of active motions of this cyclic and non-reciprocal kind. A good part of them use flagella, pili or cilia (see section 2.2.1) to swim through liquid media. On solid surfaces they facilitate crawling modes to propel themselves (see section 2.2.2).

Directed cell motion (or cell migration) is a mode of motion in which extracellular cues regulate the decision in which direction the cell will move (Macnab and Koshland, 1972; Tindall et al., 2008a; Levine and Rappel, 2008). The motility mode can be classified into three groups according to the extracellular cues: (1) Directed motion guided by solvent chemical cues is called chemotaxis (Macnab and Koshland, 1972). (2) Migration biased by substrate-bound cues influencing the cell-substratum adhesiveness is called haptotaxis (Carter, 1965). (3) If the cell motility is guided by mechanical cues such as substrate rigidity the directed motion is called durotaxis (Lo et al., 2000).

In case of mammals, cell motility is of great importance in embryonic development (Chicurel, 2002; Gray et al., 2003; Huttenlocher and Horwitz, 2011), tissue invasion of pathogenic microorganisms, inflammation and immune response (Chicurel, 2002; Gray et al., 2003; Prost et al., 2008; Parsons et al., 2010), cancer progression, invasion and metastasis (Lo et al., 2000; Chicurel, 2002; Gerisch and Chaplain, 2008; Bordeleau and Galarneau, 2010; Parsons et al., 2010), and other diseases (Wong et al., 2003).

This chapter gives an overview of the typical modes of motion throughout the various cell types. In the first part I will discuss the passive motion of *S. cerevisiae* that uses aggregation as a way to overcome random movement. Second, I will give a short overview of active motion of microorganisms to introduce a basis for the comparison with mammalian cell motility for which I will introduce a novel generic mechanical model.

2.1 Passive Motion of Yeast Cells

S. cerevisiae cells have no means of self propelled motion and just float in liquid media. They completely rely on external forces to propel them, e.g. for mixing purposes. Because the cell shape of *S. cerevisiae* is an almost perfect sphere, the force driving the motion of the cells is the Stokes friction in approximation for small Reynolds numbers $Re = \frac{uR\rho}{\eta}$

$$\vec{F}_{\text{Stokes}} = 6\pi R\eta \left(\vec{u} - \vec{v} \cdot \vec{u} \frac{\vec{u}}{|\vec{u}|^2} \right),$$ (2.1)

where ρ is the density of the medium, R is the cell radius, η is the dynamic viscosity, $\vec{u}$ is the flow field at the position $\vec{x}$ of the cell and $\vec{v}$ is the speed of the cell.

As mentioned above, inertia is irrelevant for the motion at the microscopic level, so the cells are almost instantly accelerated to the velocity of the surrounding flow.

While these arguments are true for individual cells, they do not hold for larger cell clusters such as flocs, because they are so big that inertia becomes relevant. Especially gravity and buoyancy influence the motion of the clusters. They either settle down or float upwards, so *S. cerevisiae* cells do have some kind of active motion as a collective because they can actively tune their flocculation behavior, but this does not qualify as an active motion on a cellular level.

2.2 Modes of Active Motion of Microorganisms

2.2.1 Bacterial Motility: The example of *E. coli*

The most prominent example of bacterial propulsion is the use of rotating flagella that consist of a basal body containing a molecular motor propelling a rotation, a long filament that reaches out in the medium and a hook connecting the two (see figure 2.1).
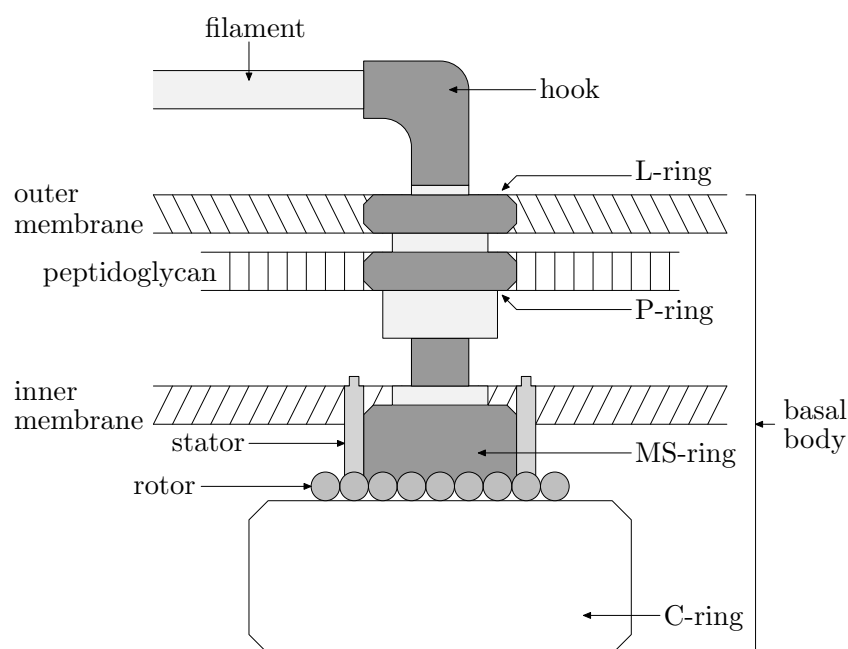

Figure 2.1: Schematics of a flagellum. The rotor is propelled by the membrane potential through a flux of protons from the periplasm into the cytoplasm. The basal body is held in place via rings that are anchored into the layers of the prokaryotic membrane. Through molecular signaling, the cell is able to change the orientation of the rotation.[1]

When the flagellum rotates it takes on the shape of a rotating corkscrew helix that fulfills the requirement of being a non-reciprocal cyclic motion (Purcell, 1977) and therefore results in a propulsion of the cell.

The best studied microorganism with respect to flagellar motion is *E. coli*, a rod-shaped

---

[1]Adapted from (Alberts et al., 2002) p. 891

prokaryote. It is the model organism for a huge number of genetic and microbiological systems. In its motile phenotype a number of flagella grow out of its sides (see figure 2.2 A). Adler (1966), Macnab and Koshland (1972) described the motion of *E. coli* as a sequence of straight swimming interrupted by tumbling that changes the direction for the next swimming period. If all the flagellar motors rotate counter-clockwise (CCW) the flagella merge and the bacterium moves forward (see figure 2.2 B). If one or more of the flagellar motors change direction and rotate clockwise (CW), the flagella leave the bulk and undergo a series of molecular configuration changes (see figure 2.2 C). This results in the cell stopping its forward motion and starting to tumble until the motors change their direction again and the flagella merge again propelling the cells in the new direction of motion (see figure 2.2 D)
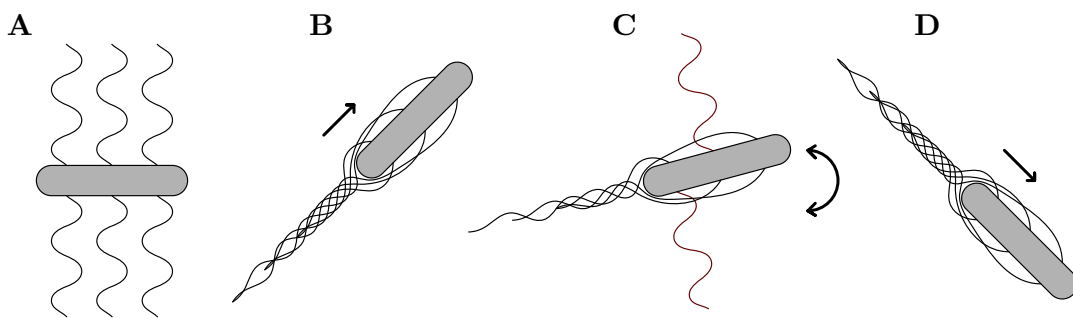

Figure 2.2: Flagella of *E. coli*: Principle of the swim and tumble. **A** *E. coli* has a number of flagella located at the sides of the rod-shaped bacterium. **B** If all flagella rotate in counter-clockwise direction they merge to propelling a straight swim in the direction indicated by the arrow. **C** If one or more flagella change the direction of their rotation to clockwise they will leave the bulk and undergo a series of confirmational changes preventing them to merge with the bulk. This causes the cell to tumble. **D** Because of the tumble the cell will swim in another direction after all flagella go back to counter-clockwise rotation.[2]

Adler (1966, 1975), Macnab and Koshland (1972) observed, that *E. coli* can modulate the timing of switching between the swimming and the tumbling motion. This enables them to react to environmental cues and perform chemotaxis. This system is one of the most studied

---

[2]Adapted from (Alon, 2007) p. 139

models for cellular migration – experimentally and theoretically[3]. The chemotaxis is guided by a temporal sensing mechanism, which can be thought of as a memory (Macnab and Koshland, 1972; Tu et al., 2008).

When *E. coli* encounters an increase in the concentration of an attractant (or a decrease in repellent) it will decrease the tumbling frequency and keep swimming longer distances. On the other hand, if it experiences a decrease in attractant (or an increase in repellent) it increases the tumbling frequency until it moves only the distance necessary to out-swim diffusion, as Purcell (1977) argued.

For my simulations I model a cell of length L_c (all occurring lengths will be measured in units of L_c) that swims a normally distributed distance l, which represents the minimum distance to out-swim diffusion, into a random direction. The cell measures the change in concentration of the cue, which is in our case an attractant, and continues to move into that direction with a probability

$$p\left(\Delta c\right) = \begin{cases} 1 - \dfrac{\frac{p_0}{\Delta c}}{kp_0\left(\frac{1}{\left(\frac{1}{p_0}-1\right)^2} + \frac{2}{\frac{1}{p_0}-1} + 1\right)} & , \Delta c \leq 0 \\[2em] \dfrac{1}{1 + \dfrac{k}{\Delta c + \frac{k}{\frac{1}{p_0}-1}}} & , \Delta c > 0 \end{cases} \tag{2.2}$$

that depends on the change of concentration and is high for larger positive changes ($\Delta c > 0$) and low for negative changes ($\Delta c < 0$) with a base probability of $p\left(\Delta c = 0\right) = p_0$ and a stiffness-parameter k (see figure 2.3). This probability is chosen in order to satisfy three requirements: $\lim_{\Delta c \to -\infty} = 0$, $p\left(\Delta c = 0\right) = p_0$, and $\lim_{\Delta c \to \infty} = 1$ with a tunable slope in between.

In the simulations the cell experiences a linear gradient parallel to the *x*-axis
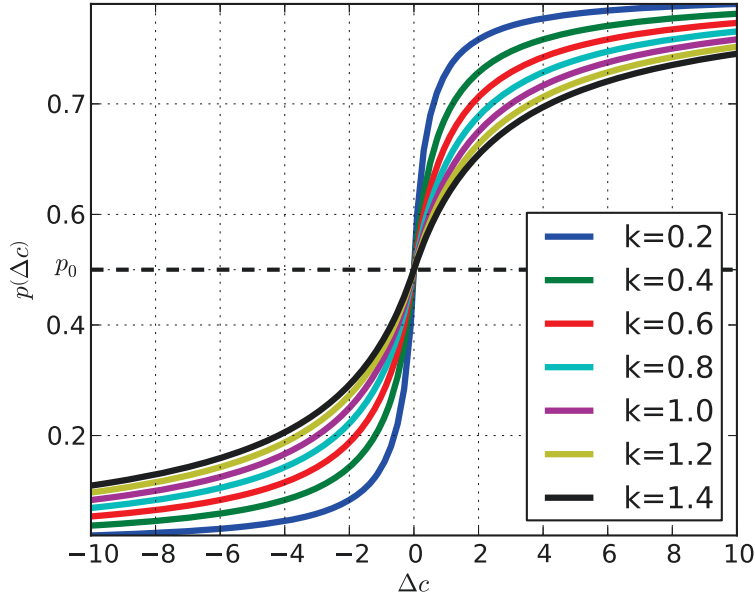
---

[3]See Tindall et al. (2008a,b) for review

Figure 2.3: Probability to continue swimming as a function of the change in concentration Δc for varying stiffness parameters k. In this case the base probability was set to $p_0 = 0.5$.

$$c(\vec{r}) = c_0(1 + r_x),\tag{2.3}$$

where c_0 parameterizes the slope.

The chemotaxis factor CF is a common measure for the efficiency of chemotaxis. Like Fuller et al. (2010), I define it as the ratio of the distance traveled in gradient direction to the total distance traveled.

As shown in figure 2.4 the CF increases with the gradient slope c_0 until it reaches a saturation. While the saturation level itself is almost independent of the parameters the slope required to reach it strongly depends on the stiffness parameter k of the probability $p(\Delta c)$.
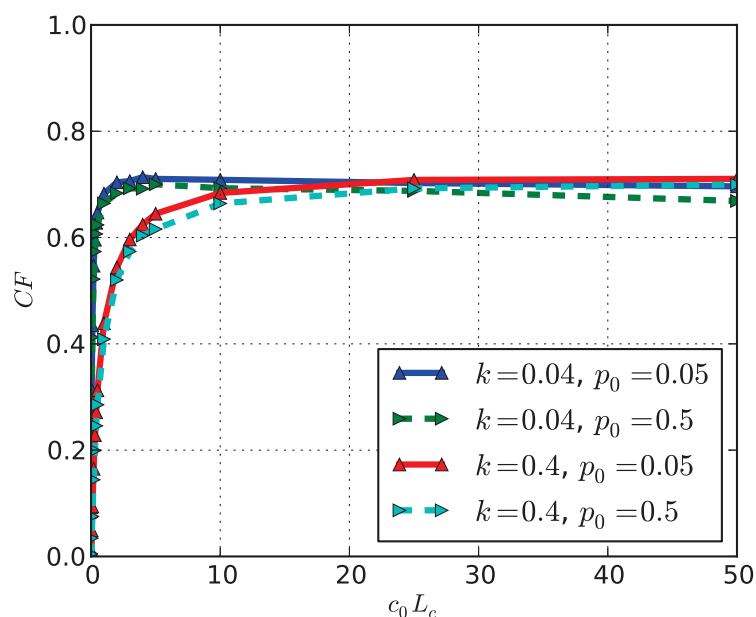
Figure 2.4: Parameter study of the swim-and-tumble model of bacterial chemotaxis. While the basal probability p_0 has very little effect on the chemotaxis factor CF, the stiffness parameter k is very important. The stiffer the probability of continuing the motion is as a function of the change in concentration, the higher gets the CF. In general steeper gradients lead to higher values for the CF until a saturation is reached.

2.2.2 Amoeboid Motility: The example of *D. discoideum*

The social amoeba *D. discoideum* is often used as a model system for motility and migration in eukaryotic cells (Fisher et al., 1989; Fuller et al., 2010). Small and Resch (2005) and Pollard and Borisy (2003) reviewed the molecular principles of the crawling mechanism of *D. discoideum* where after cell polarization actin polymerization and contraction cause the growth of protrusions at the front and at the same time retraction of the back of the cell (see figure 2.5). In this way the cells reach a speed of about $10 \, \mu m/min$.

Fuller et al. (2010) quantified the chemotaxis properties of *D. discoideum* cells and observed a sigmoidal dependence of the chemotaxis factor CF on the gradient stiffness, saturating at a maximum of $CF \approx 0.63$.

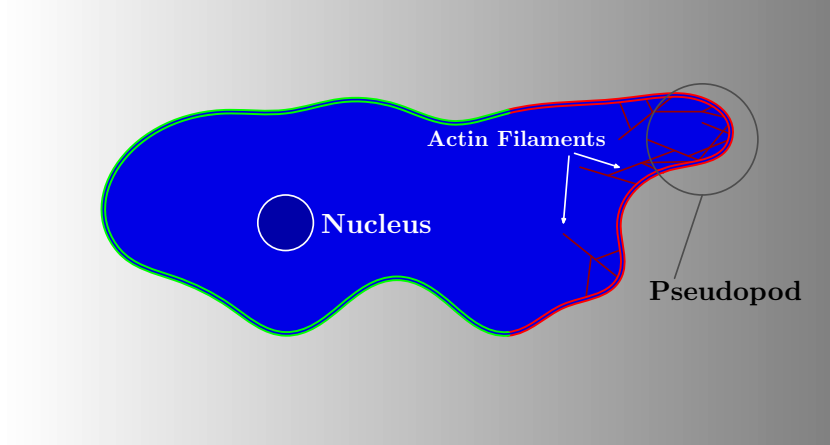[4]Adapted from (Levine and Rappel, 2013)

Figure 2.5: Scheme of a *D. discoideum* cell moving to the right. Pseudopods grow at the front (red part) of the polarized cell. At the same time the back of the cell (green part) contracts resulting in an overall motion. The polarization of the cell and therefore the direction in which the pseudopods grow to is determined by the presence of chemoattractants (represented by the gradient) that trigger a self-organization of actin filaments at the front.[4]

Levine and Rappel (2013) proposed a model for initiating the cell polarization based on the membrane bound chemoattractant receptors. The model cell has N evenly distributed receptors S that are either on ($S = 1$) or off ($S = 0$). The probability

$$p\left(S_j = 1\right) = \frac{c_j}{c_j + K_d},\tag{2.4}$$

of a receptor being "on" depends on the chemoattractant concentration c_j at the position of the receptor S_j and the dissociation constant K_d. Assuming a fairly circular cell and receptors equidistantly positioned at angles θ, the polarization angle can be calculated:

$$\Phi = \text{Arg}\left(\frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j} S_j\right),\tag{2.5}$$

where Arg is the phase of the complex number. This results in a Gaussian shaped distribution of angles (see figure 2.6).

Buenemann et al. (2010) presented a model that is focused on the contraction part of the
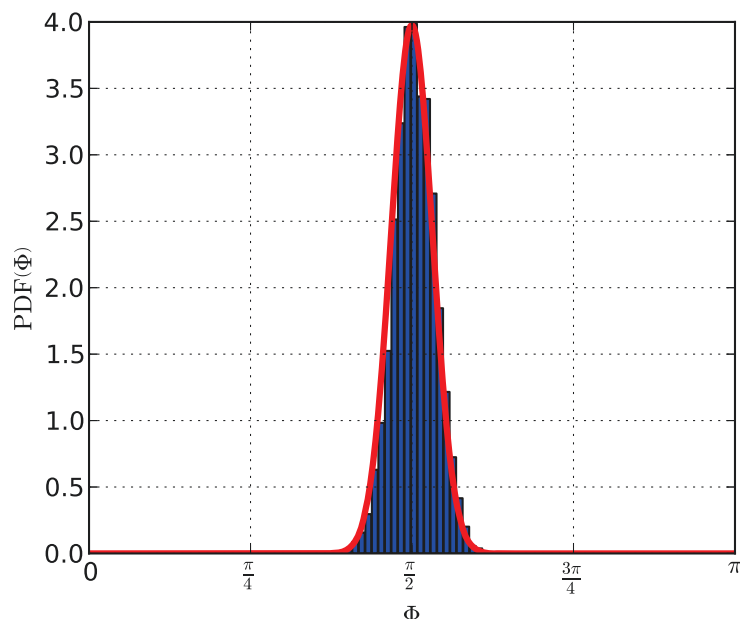
Figure 2.6: The distribution of polarization angles of the model by Levine and Rappel (2013) for a linear concentration gradient in $\Phi = \frac{\pi}{2}$ direction shows an approximate Gaussian distribution in gradient direction (the red line shows a fit with $\mu = 0.500\pi$ and $\sigma = 0.032\pi$). For an increasing number of receptors and steeper gradients the distribution will narrow, for a decreasing number of receptors and flatter gradients it will broaden.

cellular motion, which they assume to be uniformly and at constant speed over a contraction phase. They represented the adhesive area by an ellipse with N randomly distributed attachment points. This rather simple model already has some experimentally testable predictions the strongest of which is that the cell speed is largely independent of the value of the adhesive forces between the cell and the substratum.

Hecht et al. (2011) introduced a more realistic model, that represents the membrane of the cell in two dimensions as a set of nodes. The motion is triggered by the concentration fields of an activator and an inhibitor that is calculated using reaction-diffusion equations. Each node of the membrane experiences a force normal to the membrane that depends on the value of the activator concentration at the node's position. With this detailed model they were able to qualitatively reproduce the amoeboid motion, especially the membrane dynamics at the cell

front.

The model presented by Camley et al. (2013) describes the cell's cytoskeleton as a viscous, compressible fluid driven by active stresses resulting from actin polymerization. The coupling of cell shape and cell polarity results in a robust periodic migration that requires a balance between contraction and protrusion.

In conclusion, models of amoeboid motion and migration so far are focused on details of the motion process like polarization (Levine and Rappel, 2013) or contraction (Buenemann et al., 2010), on special modes of migration (Camley et al., 2013) or on the shape of the cell (Hecht et al., 2011).

2.3 Mammalian Cell Migration

Mammalian cell motility is a complex process integrating biochemical kinetics, transport, and physical forces. Mogilner (2009) reviewed the often used combinations of continuum mechanics and biochemical kinetics to model this migration.

The predominant mode of migration in mammalian cells is the crawling motion (Keren and Theriot, 2008) that can either be classified as amoeboid-like migration or mesenchymal migration. The amoeboid-like migration is very similar to the migration of amoeba described in the previous section. Therefore, I will focus on the mesenchymal mode of migration of epithelial cells, e.g. mouse embryonic fibroblasts, which is influenced by a variety of chemical and physical signals, in particular by different mechanical forces (Butcher et al., 2009) but shows a very high efficiency in following all kinds of external gradients (Theveneau et al., 2010).

The migration cycle consists of five discrete steps carried out within about ten minutes (Mogilner, 2009; Lauffenburger and Horwitz, 1996; Pathak and Kumar, 2011; Ridley et al., 2003; DiMilla et al., 1991; Zaman et al., 2005; Parsons et al., 2010; Huttenlocher and Horwitz, 2011; Chicurel, 2002; Keren and Theriot, 2008): (1) polarization of the cell, resulting in a defined leading and a defined lagging pole; (2) formation of protrusions at the leading pole and attachment of the lagging pole; (3) stabilization of these protrusions by adhesion to the substrate or the extracellular matrix (ECM); (4) trans-location of the cell-body by myosin-mediated contraction; (5) retraction of the rear by loosening the adhesions at the lagging pole. Mogilner (2009) stated that the crawling is propelled by the active lamellipodeum at the leading edge that pulls the passive cell body forward.

The stiffness and inhomogeneity of the substrate or the ECM are very important factors for cell motility because they influence the direction of motion through durotaxis as first discovered by Lo et al. (2000) and confirmed by Gray et al. (2003), Yeung et al. (2005), Provenzano

et al. (2008), and Trichet et al. (2012). Besides the impact on migration itself Holmes and Edelstein-Keshet (2012) showed that these factors also strongly influence cell shape.

Pathak and Kumar (2011) describe the interaction of the cell and the substrate or the ECM as a bilinear sequential binding which affects the cell predominantly on the nano scale through mechanosensing mechanisms as mentioned by Butcher et al. (2009) and Trichet et al. (2012).

However, cells can also chemically manipulate the ECM, e. g. in case of cancer, where tumor cells stiffen the surrounding ECM, and build a rigid stroma around the tumor[5]. Bordeleau et al. (2013) observed that this step in ECM stiffness promotes cells from the outside moving inside, but prevents cells from the inside to migrate outside.

In general, mammalian cells are not passive recipients of mechanical forces, but actively react to them gradually by pulling or pushing the ECM or the substratum (Lauffenburger and Horwitz, 1996; Lo et al., 2000; Butcher et al., 2009).

2.3.1 The "Octopus Model"

The great complexity of cell motility makes it an interesting field of study for both experimental and theoretical scientists alike. Mogilner (2009) reviews a number of models trying to describe different modules of the cell motility system or to simplify the system to some extent.

So far, the models are focused on cell shape during migration (Rubinstein et al., 2005), continuum descriptions for cell populations (Gerisch and Chaplain, 2008; Häcker, 2012) or on special short-term aspects of migration like migration speed (DiMilla et al., 1991) or effective adhesiveness (Zaman et al., 2005).

However, there are a lot of unanswered questions. My approach is a simple, generic model for environment sensitive motion of fibroblast-like cells that is solely based on mechanics. It covers chemotaxis, haptotaxis, and durotaxis as modes of motility, to answer general questions on the kind of motion and taxis potential of the cells. My model is able to explain the

---

[5]for a review see (Butcher et al., 2009; Provenzano et al., 2008)

high efficiencies of mesenchymal-like motion observed by Theveneau et al. (2010) independent of cellular morphology. It also covers large-scale dynamics and describes the statistical properties at environmental discontinuities, e.g. a step in stiffness of the substratum or a step in concentration of a chemoattractant as observed by Bordeleau et al. (2013).

The model facilitates the discrete nature of the cell motility in a modular way, which enables it to act as platform for testing different ideas on regulation, polarization, protrusion and contraction.
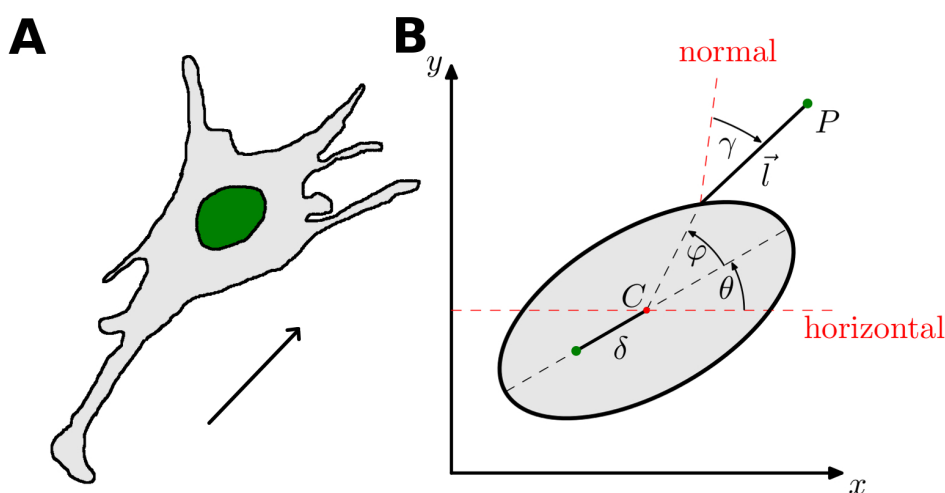

Figure 2.7: Motivation and schematics of the octopus-model. **A** Typical cell shape as observed in Lo et al. (2000) for a cell polarized in the direction of the arrow. The protrusions mainly grow into forward direction, only a few in backward direction. **B** The position of an attachment point is defined through the length of a protrusion and through three angles: the rotation angle of the cell θ, the angle that defines the attachment position on the cell membrane φ and the angle between the central line and the arm itself γ.

Figure 2.7 motivates the parameterization of the model. A cell as observed by Lo et al. (2000) (shown in figure 2.7 **A**) is represented by an ellipse with major radius R_l and minor radius R_s (see figure 2.7 **B**). All lengths will be measured in units of the major radius. The angle θ parameterizes the rotation of the long-axis of the ellipse with respect to the x-axis around the ellipse center C. The rotational center of the ellipse, i.e. the fixed counterpart of the

protrusions is defined by the shift δ along the major axis.

The gradient in concentration or stiffness that the cell follows is calculated using the position. For simplicity this gradient is set to be linear and parallel to the x-axis:

$$c(\vec{r}) = c_0(1 + r_x).\tag{2.6}$$

For the initial polarization of the cells I assume a molecular mechanism that effects the cell similar to the model presented in section 2.2.2 in presence of a gradient in concentration or stiffness. For the purpose of this work, I do not explicitly model the polarization mechanism, but I replace it by a normally distributed initial angle $\theta(t=0) = \theta_0$ around the gradient direction with the standard deviation σ_{θ_0} as a parameter. The value of this parameter can now be calculated from a fit to the model (see figure 2.6). For a gradient of $c_0 = 1/R_l$ this fit results in a standard deviation of $\sigma_{\theta_0} \approx 1$.
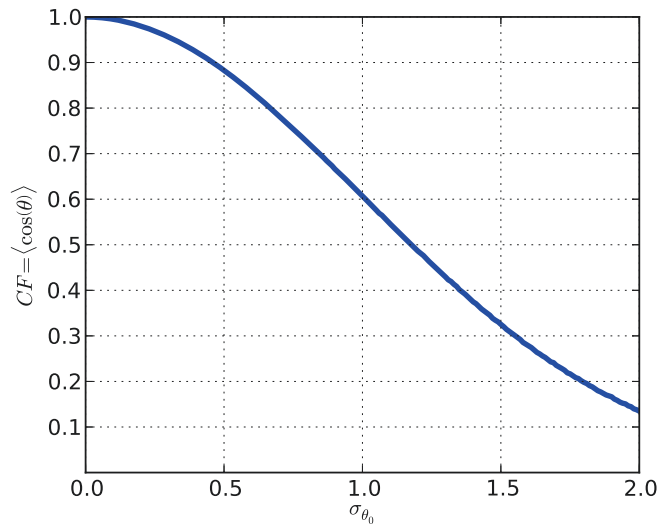

Figure 2.8: Initial chemotactic factor as obtained by a normally distributed initial angle θ_0 with a standard deviation of σ_{θ_0}

The chemotaxis factor for a step taken in θ_0 direction as a function of the standard deviation ($CF = \langle \cos(\theta_0) \rangle$) is an almost linear decrease (see figure 2.8). This initial CF will be used to

benchmark the model parameters towards an increase of efficiency.

The protrusions by which the cells pull themselves forward are represented by adhesive arms that grow out of the ellipse at a random angle φ, which is distributed according to the polarization of the cell

$$
\begin{aligned}
p_{\text{side}}(\varphi) = \quad & \frac{1}{\sqrt{2\pi}} \left(\frac{p_{\text{front}}}{\sigma_{\varphi_{\text{front}}}} \exp\left[-\frac{1}{2} \left(\frac{\varphi}{\sigma_{\varphi_{\text{front}}}} \right)^2 \right] \right. \\
+ \quad & \left. \frac{p_{\text{back}}}{\sigma_{\varphi_{\text{back}}}} \exp\left[-\frac{1}{2} \left(\frac{\varphi - \pi}{\sigma_{\varphi_{\text{back}}}} \right)^2 \right] \right).
\end{aligned}
\tag{2.7}
$$

The weights p_{front} of the leading pole and $p_{\text{back}} = 1 - p_{\text{front}}$ of the lagging pole result in a polarization of the cell by shifting the arm distribution towards one pole.

The length of the arms is randomly chosen from a normal distribution with variable mean μ_l and standard deviation σ_l. For simplicity, the angle γ that defines the direction in which the arm grows in relation to the ellipse's surface normal is drawn from a Gaussian distribution with mean $\mu_\gamma = 0$ and (variable) standard deviation σ_γ.

Arm formation occurs at a constant rate. Every arm applies a linear force on the cell that is proportional to the concentration or rigidity of the substrate at the attachment point of the arm. Thus, if at time t there are N arms of lengths l_i at angles φ_i the total force applied on the cell is given by

$$
\vec{F}(\vec{r}_{\text{cell}}, \theta, t) \quad = \quad \sum_{i=1}^{N} \vec{F}_i = \sum_{i=1}^{N} k(c) \vec{l}_i (\vec{r}_{\text{cell}}, \theta, \varphi_i),
\tag{2.8}
$$

where $k(c) = k_0 c(\vec{r}_{arm})$. This results in a trans-location of the cell. Furthermore, the arms exert a total torque

$$
M(\theta, t) \quad = \quad \sum_{i=1}^{N} x(\varphi_i, \theta, \delta) F_{i,y} - y(\varphi_i, \theta, \delta) F_{i,x},
\tag{2.9}
$$

that leads to a rotation of cell. We interpret this rotation of our rigid body model as a gradual repolarization of the cell which in reality is pliant.

To apply a gradient to our model, e.g. for a gradient parallel to the *x*-axis, we modify the concentration

$$c\left(\vec{r}_{\text{arm}}\right) = c_0\left(r_{\text{arm},x}+1\right), \tag{2.10}$$

where c_0 is the increase over the major radius R_l of the ellipse. Figure 2.9 shows the statistics of force and torque with and without an external gradient. It indicates a shift in the torque-distribution ultimately resulting in an alignment with the gradient. Also the *x*-component of force distribution is shifted to larger values which leads to the cell moving along the gradient.
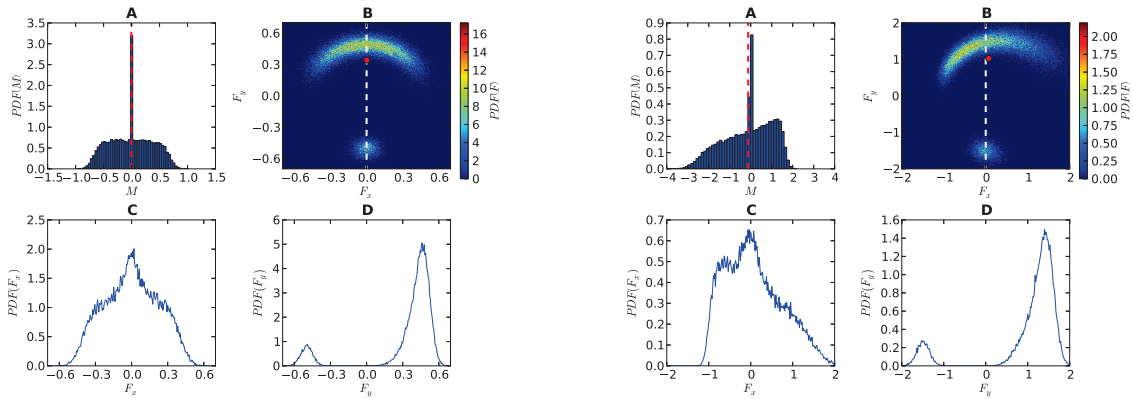

Figure 2.9: Statistical analysis of force and torque for a cell oriented towards $\theta = \frac{\pi}{2}$ without a gradient (left) and with a linear gradient in $\theta = 0$-direction (right). **A**: The torque distribution becomes asymmetrical if the system is influenced by a gradient. This leads to an effective torque pushing the cell towards an alignment with the external gradient. **B** The two-dimensional force distribution also becomes asymmetrical and shifts in gradient direction resulting in a movement along the gradient. **C** and **D**: The integrated force distributions in *x* and *y* direction indicate, that the shift is only affecting the forces in gradient direction and not in the perpendicular direction.

To numerically analyze the model I let the cell grow N new arms of lengths l_i at angles φ_i in every iteration. The new position and orientation of the cell is then calculated by solv-

ing $\vec{F}(\vec{r}_{\text{cell}}, \theta, t) = \vec{0}$ and $M(\theta, t) = 0$ independently. Then, all arms are removed, the time is increased and the iteration starts over.

2.4 Modeling Results

To classify the planar cell motion (in the x-y plane) produced by the model, I numerically calculated the mean displacements ($\langle x \rangle$, $\langle y \rangle$) and the mean squared displacements ($\langle x^2 \rangle$, $\langle y^2 \rangle$) as a function of time t (represented by the number of iteration steps), as an average over 500 independent runs. Figure 2.10 (left) shows the evaluation of these values in presence and absence of a gradient[6]. For better comparison the plots show $\log \langle x^2 \rangle$ and $\log \langle y^2 \rangle$ instead of $\langle x^2 \rangle$ and $\langle y^2 \rangle$.

In absence of a gradient the cells perform an isotropic random walk, indicated in figure 2.10 (left) by the equal linear increase in $\langle x^2 \rangle$ and $\langle y^2 \rangle$ while the mean displacement is zero. For $N = 10$ the effective diffusion coefficients $D_x \approx D_y = 43.77 \, R_l^2 / \delta t$ were measured by fitting linear functions to $\langle x^2 \rangle$ and $\langle y^2 \rangle$.

If a linear gradient of $c_0 = 1/R_l$ parallel to the x-axis is applied to the simulation, the cells perform a biased random walk in gradient direction while the motion in perpendicular direction is suppressed by a factor of 135: $D_y = 0.32 \, R_l^2 / \delta t$ in presence of a gradient compared to $D_y = 43.77 \, R_l^2 / \delta t$ in absence of a gradient.

To quantify the efficiency in following the applied gradient I measured the chemotactic factor of the cellular motion as introduced in section 2.2.1. In the presence of a gradient with $c_0 = 1/R_l$ we find $CF = 92.4\%$ for $N = 10$.

The model is very robust to varying gradients c_0. There is a strong increase in efficiency with increasing gradient strength for smaller values of c_0 but the efficiency saturates fast to its maximum value of around $CF = 95\%$. For small gradients the model shows slightly worse efficiencies than the initial polarization while it shows a clear increase in efficiency for gradients above $c_0 = 0.02/R_l$ (see figure 2.11).

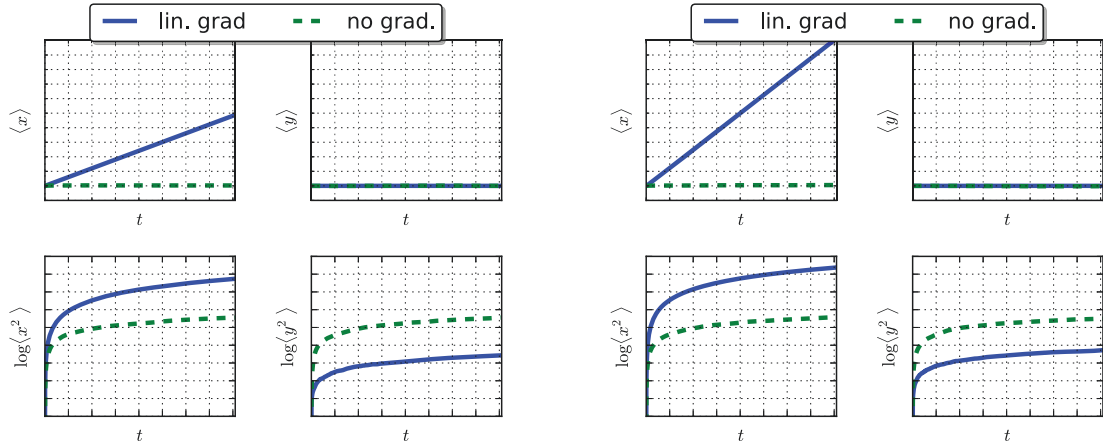[6]a detailed parameter listing can be found in table 2.1

Figure 2.10: Classification of the migration dynamics: The cell performs a random walk which is biased in the presence of a gradient (left images). The motion in *y*-direction is suppressed compared to the situation without gradient. If the arms are regulated (right images) the motion in absence of a gradient stays the same, whereas the motion in presence of a gradient is boosted, resulting in an increase in speed while keeping a constant efficiency.

Increasing the number of arms results in a speed up of the motion and an increase in efficiency. However, there is saturation in efficiency and speed for large numbers of arms. If taken into account that 5 to 10 arms with an average length of $7\,\mu$m would roughly cover 10% to

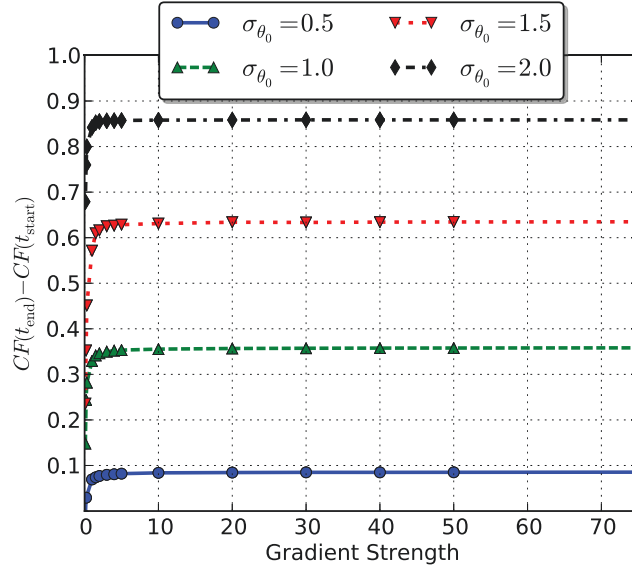| Parameter	Fig. 2.12	Fig. 2.13**A**	Fig. 2.13**B**	Fig. 2.15	Fig. 2.10
Iter.	500	500	500	500	81000
Runs	500	500	500	500	500
a	1	1	1	1	1
b	0.5	0.5	0.5	0.5	0.5
δ	–	-1	-1	-1	-1
σ_θ	–	–	0	1.047	1.047
k	1	1	1	1	1
c_0	1	1	1	1	–
N	10	10	10	10	10
p_{Front}	0.9	0.9	0.9	0.9	0.9
μ_l	0.5	0.5	0.5	0.5	0.5
σ_l	0.05	0.05	0.05	0.05	0.05
$\sigma_{\varphi_{\text{Front}}}$	0.5	0.5	0.5	0.5	0.5
$\sigma_{\varphi_{\text{Back}}}$	0.05	0.05	0.05	0.05	0.05
σ_γ	0.1	0.1	0.1	0.1	0.7

Table 2.1: Parameter sets for the Figures

Figure 2.11: Gains for increasing gradient strengths. The increase of efficiency only depends on the gradient strength for very small gradients.

20% of the cell's surface (each arm has about $22\,\mu\text{m}^2$ surface while according to Baxter et al. (2002), Raucher and Sheetz (1999), the cell body has about $1000\,\mu\text{m}^2$ membrane area) the model reaches a good balance between increase of membrane area and efficiency within this range of N. This is comparable to the number of arms observed experimentally, e.g. by Lo et al. (2000).

Lo et al. (2000) have shown that cells spread further on stiff substrates than on soft substrates. To account for this effect in the model, I introduce a regulated additive elongation

$$\lambda\left(c\right) \quad = \quad \frac{\lambda_{\text{max}}}{1 + \frac{k_c}{c}} \tag{2.11}$$

of each arm that depends on the absolute stiffness (or concentration or adhesiveness) $c = c\left(\vec{r}_{arm}\right)$.

Increasing the average arm length entails an increase in efficiency that saturates for longer arms. This rise in efficiency comes with an increase in speed along the gradient in terms of distance traveled per iteration. If the concentration dependent regulation of arm lengths is

used a comparable increase in efficiency can be seen but at an average arm length that is 40% to 60% shorter compared to the unregulated system. The speed also slightly increases with regulation – the same speed as the unregulated system is obtained at 7% shorter average arm lengths.
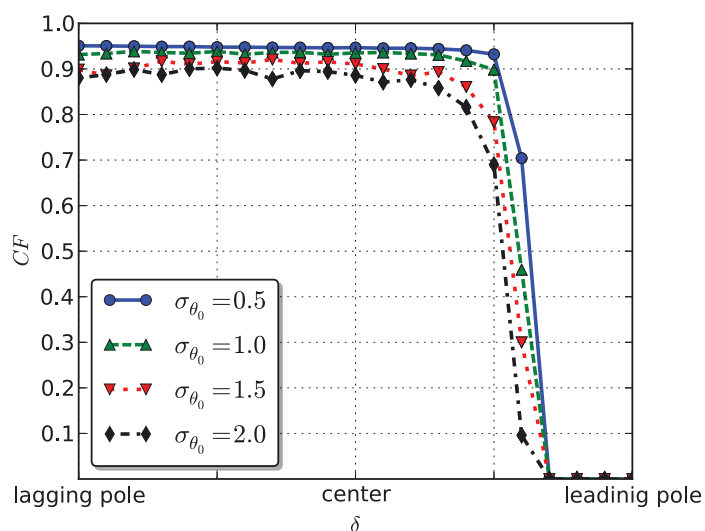

Figure 2.12: Dependency on the shift δ. If we shift the rotational center from the center of the ellipse towards the leading pole, we see a stiff drop in efficiency as indicated by the chemotactic factor CF. Here the final efficiency is even below the initial efficiency that results from the normally distributed θ (see FIG. 2.8). The reason for this is the reduced lever of the leading compared to the lagging pole. This result emphasizes the need for the cells to have a defined leading pole and an attachment point further away from that pole.

The above results are robust with respect to variations in the standard deviations of the distributions for φ and γ. For range from about 0.01π to about 0.2π for σ_φ and σ_γ the CF remains almost constant. For even broader distributions the efficiency decreases as a result of insufficient polarization of the cell. Especially the high efficiencies of distributions with small standard deviations might explain the experimental observations, that the protrusions almost grow in normal direction out of the surface of the cells, and that the protrusions are located around a narrow region at the leading pole (Maly and Borisy, 2001; Svitkina et al., 2003;

Mogilner and Rubinstein, 2005; Parsons et al., 2010).

Furthermore, cell motion shows an interesting dependence on position of the attachment point of the cell to the substratum quantified by the parameter δ, see Fig. 2.7. If this point is shifted towards the leading pole, the efficiency drops sharply to zero indicating that the cell is not able to follow the gradient at all (see figure 2.12).

It is a quite common scheme for crawling cells to detach the adhesions in a region close to their lagging poles (Lauffenburger and Horwitz, 1996; Ridley et al., 2003). My model gives an explanation for this, as shows the achieved increase in taxis efficiency. In this way the cell not only compensates for the additional energy costs and the slowing effect of an attached back but also directly benefits from the attachment.

Butcher et al. (2009) and Bordeleau et al. (2013) describe the movement towards a step in substrate rigidity as a typical scenario of inhomogeneous substrates encountered by the cells. The model shows that the resulting behavior heavily depends on the direction of movement, i.e. if the cell moves from softer to stiffer substrate or vice versa, see figure 2.13.

Cells that move from a soft to a stiffer substrate just keep moving on. However, their trajectory gets bent towards the direction perpendicular to the step. I observed that the relation between the angle of the cell before and after the step obeys a refraction law similar to that of light. In this way we can characterize the motion by refraction indices. The ratio of refraction indices $\frac{n_{\text{Soft}}}{n_{\text{Stiff}}}$ decreases with increasing step height (see Table 2.2) which indicates a stronger alignment to the direction perpendicular to the step.

On the other hand a step from a stiff substrate to a softer substrate is a barrier for the cells. The passing probability depends on the step height. From the distribution of the maximal x-positions – i.e. how far the cells invaded the softer side of the step – encountered by the cells during 500 iterations (see figure 2.14) one can calculate the probability of a cell moving across the step, which I define as transmission coefficient $T_{\text{C}} = \int_{-\infty}^{x_{\text{step}}} p(x) = p(x < x_{\text{step}})$, where $p(x)$
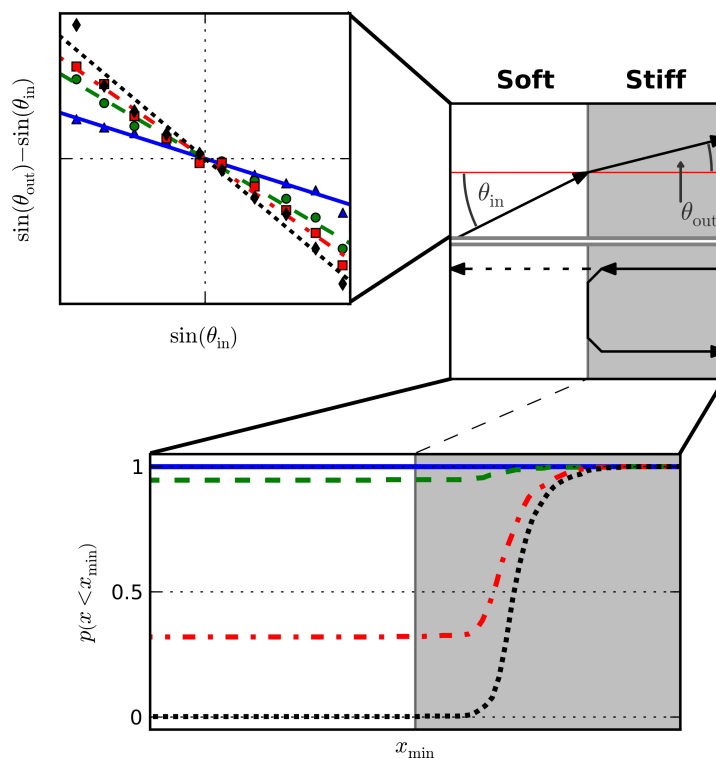
Figure 2.13: Transition properties at a sharp concentration/stiffness step. If cells travel from the soft region to the stiff region we see a refractive behavior with a step size dependent ratio of the refractive indices (see Table 2.2, ▲ $s = 2$, ● $s = 5$, ■ $s = 10$, ♦ $s = 100$). Cells moving from the stiff to the soft region will experience a barrier, that they overcome with a step size dependent transmission coefficient. For large step sizes no cells will be able to break the barrier whereas at small step sizes almost all cells will cross it (see Table 2.2).

is the probability to find a cell that traveled to position x and x_{step} is the position of the step.

This coefficient decreases as the step size increases (see Table 2.2) showing that the barrier-

Step height	$\frac{n_{\text{Soft}}}{n_{\text{Stiff}}}$	T_{C}
2	0.9374	1.0000
5	0.8843	0.9480
10	0.8622	0.3220
100	0.8330	0.0020

Table 2.2: Values for the ratio of the refraction indexes $\frac{n_{\text{Soft}}}{n_{\text{Stiff}}}$ and transmission coefficient T_{C}. The refractive effect becomes stronger with increasing step size whereas the transmission decreases.
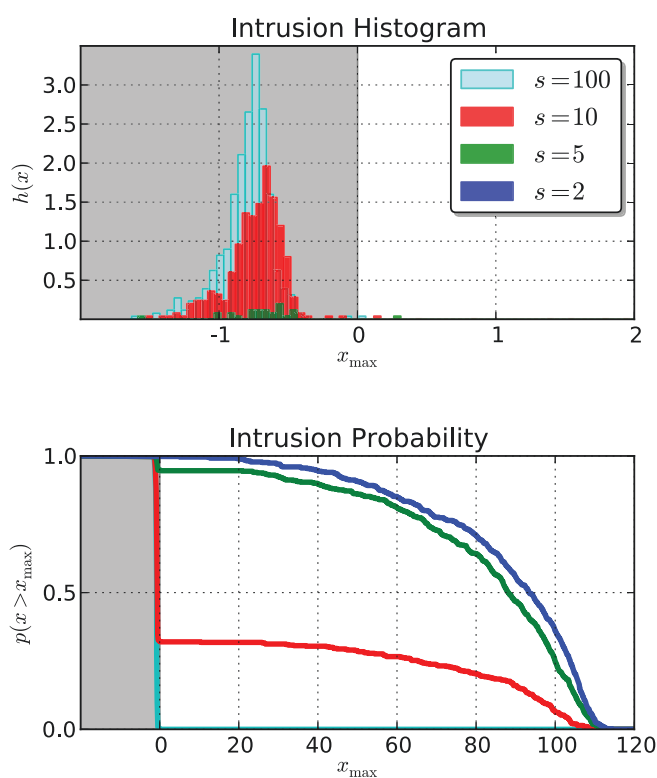
Figure 2.14: Histogram and probability of intrusion from the right onto a stiff-soft interface. The Histogram shows the frequency of the maximal (most right) x value reached during the runs. For higher steps not all the cells will cross the interface. The maximal distance traveled in positive x-direction peaks closer to the interface with decreasing step size. If the step size becomes to small almost all cells breach the interface as indicated by the intrusion probability that is defined as the probability that a cell travels further than x_{max}.

effect becomes much stronger for higher steps.

If we compare the results to a two dimensional swim-tumble model of bacterial chemotaxis (see 2.2.1 for details) with a parameter set optimized for efficiency, the results of the crawling model shows a much faster saturation of the efficiency on a higher efficiency level than the swim-tumble model (see Fig. 2.15).
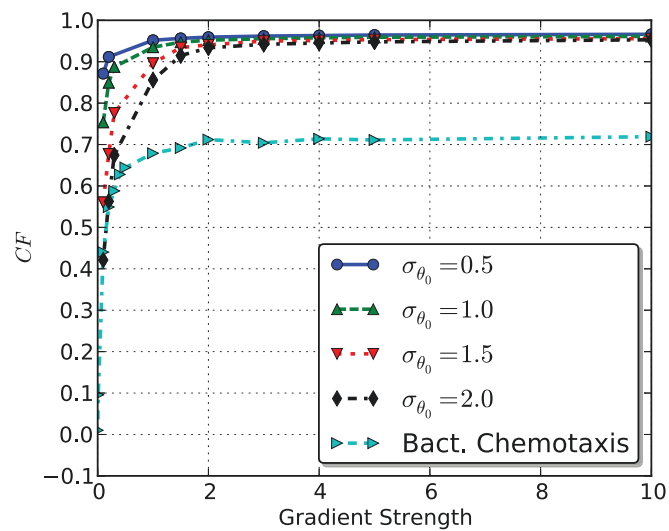
Figure 2.15: Variation of the efficiency with the gradient strength. The cells reach their maximum efficiency at rather small gradients compared to a swim-tumble model of bacterial chemotaxis (for details of that model see 2.2.1).

2.5 Discussion and Summary

This chapter gave an overview of the different modes of motility at the microscopic level. Especially the low Reynolds number has a huge impact on how cells and microorganisms move, because it makes it necessary to perform non-reciprocal motion.

Cell migration is a directed movement of a cell or microorganism. Environmental cues lead to a bias in the otherwise random motion directing the cells towards good or away from bad environments.

The best studied system for bacterial motility is *E. coli*. This prokaryote uses flagella to swim through its environment and a swim and tumble technique to find gradients (Macnab and Koshland, 1972). With a very simple model, I showed that in this way an efficiency of about 70% to 75% can be reached.

In eukaryotic motility *D. discoideum* is the model system for amoeboid like migration. Because of its size it can use membrane based sensors to find the direction of a local gradient. Models often simulate local membrane deformations by active internal stresses that are guided by external cues (Buenemann et al., 2010; Hecht et al., 2011; Camley et al., 2013).

Another important mode of eukaryotic motility is the mesenchymal mode of migration of epithelial cells. For this cyclic mode of migration I introduced a generic two dimensional model that attempts to simulate large scale crawling dynamics while avoiding deformation effects of a three dimensional generalization. The simulation results are in good agreement with the experimental observations at steps in substrate rigidity observed by Bordeleau et al. (2013), cellular morphology observed by Lo et al. (2000), and indicate that the attachment of the lagging pole increases the effectiveness of crawling.

Furthermore, the refractive indices introduced in section 2.4 can be used to characterize substrate rigidity including the biological impact of the material compared to a reference substrate.

So far I have neglected the mechanical effects that the cells have on the substrate. If cells attach their protrusions to the substrate and contract they locally stiffen the substrate. This might lead to an indirect attraction between two cells in proximity and thus promote aggregation.

All in all, my model produces reasonable results that fit many experimental observations qualitatively. To further quantify the motion one could measure the chemotactic properties – especially the averages $\langle x \rangle$, $\langle x^2 \rangle$, $\langle y \rangle$ and $\langle y^2 \rangle$ – using a Dunn-Chamber (e.g. as described by Zicha et al. (1991)) with time-lapse microscopy. This should be possible for a large number of cells using automation techniques (Fotos et al., 2006) and cell-tracking methods (Chen et al., 2006). To check the results concerning steps in concentration, adhesiveness and stiffness one could use the methods presented by Bordeleau et al. (2013) to produce flat substrates of different stiffnesses and measure the polarization axes of the cells before and after the interface and the transmission coefficients with automatized time lapse microscopy.

Chapter 3

Cellular Aggregation

Pattern formation is relevant at all scales of life from protein folding over spatial gradients inside a single cell to the formation of complex lifeforms. Examples for the importance of spacial concentration patterns are the roles of the Min-System in *E. coli* as reviewed by Lutkenhaus (2007) and the MipZ-System in *C. crescentus* as described by Thanbichler and Shapiro (2006).

The majority of concentration patterns can be explained by models based on a two-player system as introduced by Turing (1952). This system is based on diffusion of and reactions between the two players, i.e. activations and inhibitions. Wolpert (1969) pointed out that the key component of this system is the combination of a short range activation (e.g. achieved by a self activation) and a long range inhibition (e.g. achieved by a fast diffusing inhibitor).

Kondo and Miura (2010) mentioned the difficulty in proving that a Turing mechanism is responsible for a biological pattern, because of the complicated (and sometimes unknown) interactions within living cells. However, it is a good tool to understand the basic principles

of pattern formation in biological systems like the formation of stripes and dots in the animal skin studied by Kondo (2002). Koch and Meinhardt (1994) reviewed a number of systems in which Turing mechanisms are used to simulate and explain patterns occurring in complex biological structures.

Pattern formation is very important in embryonic development because, as Melton (1991) put it, it is challenging to explain how a relatively small number of genes and mechanisms can result in such different animals.

On a higher level, cell-cell interactions can trigger the formation of patterns on a colonial level. Liu et al. (2011) transformed a synthetic gene circuit into *E. coli* cells that connects the motility system of *E. coli* to a quorum sensing module. In this way the cells motility is inhibited if the cell concentration is higher. This coupling leads to stripe patterns that can be tuned experimentally as predicted in models introduced by Fu et al. (2012).

This indicates that the underlying mechanisms in pattern formation can be relatively simple. It is also a very good example for the benefits of the interlock of experimental and theoretical methods to approach this topic.

Another very important type of pattern formation is aggregation. From a theoretical perspective, the best studied form of aggregation is the diffusion limited aggregation (DLA) which is mainly used to describe crystallization processes. Witten Jr. and Sander (1981) developed a model, where a cluster is seeded at the origin of a lattice. A particle is added to a random lattice site at a large distance from the origin and allowed to perform a random walk until it visits a lattice site adjacent to an occupied site. The particle becomes part of the cluster and a new particle is added at large distance and allowed to perform a random walk and so forth[1]. Vicsek (1984) showed that it is possible to generate various kinds of regular patterns using a slightly modified DLA model.

In biology, aggregating cells are an important example for pattern formation through aggre-

---

[1] for a detailed analysis of this model see Witten and Sander (1983)

gation. In my work I will focus on two examples: the two dimensional aggregation of mouse embryonic fibroblast (MEF) cells and the three dimensional flocculation of *S. cerevisiae*.

Established by Todaro and Green (1963), the so called 3T3 MEF cell line has become the standard fibroblast cell line in laboratories. The original fibroblast cells are transferred to a fresh medium every three days and inoculated at a cell concentration of $3 \cdot 10^5$ cells per petri dish.

Schlessinger et al. (1978) was one of the first to use this model system to understand the basic principles of the aggregation of fibroblast cells which is an important step in the formation of islet-like cell aggregates that could substitute the damaged insulin producing cell islets in diabetes type 1 patients (see the works of Chandra et al. (2011), Govindasamy et al. (2011), and Hardikar et al. (2003)).

As introduced in chapter 2, MEF cells show an active form of motility in the similar STO strain developed by Martin and Evans (1975). After they drew near to each other they form bonds and contract to islet-like aggregates. As reviewed by Hynes (1987, 1992), these bonds are mediated via a family of cell surface proteins named integrins, which also play a role in many other interactions of fibroblast like cell motility (see chapter 2) or cell-surface interaction in general. In a more recent study Toret et al. (2014) elucidated the high genetic complexity of these integrins involved in cell-cell adhesions. For the purpose of this work it is sufficient to consider the integrin mediated interactions as a simple binding while neglecting possible interactions on a regulatory level.

Flocculating *S. cerevisiae* cells on the other hand have no means of active motion and are drawn near to each other by external flows. As reviewed by Soares (2011), Brückner and Mösch (2012) flocculation is defined as the asexual, reversible and Ca^{2+} dependent aggregation of typically thousands of vegetative cells into flocs. The biotechnological impact of *S. cerevisiae* strains makes flocculation an important field of study, because it is a cheap and effective way of removing cells from a desired product after fermentation pro-

cesses. The adhesins mediating the bonds between vegetative yeast cells belong to the family of fungal glycosylphosphatidylinositol-linked cell-wall proteins, and are studied extensively, e.g. by Verstrepen et al. (2004); Verstrepen and Klis (2006); Dranginis et al. (2007); Linder and Gustafsson (2008). Dranginis et al. (2007); Verstrepen and Klis (2006) classified the adhesins into two general groups: they mediate either sugar-sensitive (called lectin-like) or sugar-insensitive adhesion. As mentioned in the introduction, the focus of my work lies on the lectin-like adhesions, that in contrast to sugar-insensitive adhesions allow for adhesions between expressing and non-expressing cells.

I collected a set of tools from fractal and graph theory to quantify and compare aggregation patterns and developed models to simulate the aggregation dynamics in both cases.

In cooperation with Jianjiang Hu and Wei Huang from the Department of Biochemistry of the Hong Kong University we performed time-lapse microscopy of MEF cell aggregation and engineered tools for automated evaluation of the images.

In cooperation with Stefan Brückner and Hans-Ulrich Mösch from the Department of Biology of the Philipps-University Marburg we developed new experimental approaches for the three dimensional structural analysis of yeast flocs using confocal laser scanning microscopy.

This chapter summarizes the efforts towards a quantitative understanding of cellular aggregation in all kinds of biological systems. It reviews existing and introduces new techniques to generate a universal picture of the aggregation process.

3.1 Quantification of Aggregation Patterns

Like many natural spatial patterns cellular aggregates are so highly irregular or fragmented that concepts of classical geometry are insufficient to describe their structure. To quantify complex structures like these Mandelbrot (1977) introduced the concepts of fractal geometry.

Bunde and Havlin (1994) collected a number of examples and application for these concepts in natural sciences. In a strict mathematical sense fractals do not have a characteristical length scale. This means if one sees a picture of a fractal where $1\,cm$ represents $10\,km$ and another picture of the same fractal, where $1\,cm$ represents $0.01\,cm$ one can not relate either of the pictures to the used scale. However, in most natural patterns this is only true for all scales but for a good range of scales, because all natural things have a finite maximal size while the smallest (biologically relevant) scale is reached with molecules. Cellular aggregates have a finite size, so for lengths larger than this size the single cell yields a characteristic length-scale. On scales smaller than the size of a single cell the fractal properties are not applicable.

Nevertheless, characterizing cellular aggregates using the tools of fractal geometry yields good methods to quantify and compare those patterns like the fractal dimension and the pair-correlation function I will introduce in sections 3.1.2 and 3.1.3.

A number of other good tools to characterize cellular aggregates can be derived from the theory of complex networks. Albert and Barabási (2002) reviewed the commonly used techniques to quantify networks. From their work I adopted the clustering coefficient (section 3.1.6), the average path length (section 3.1.7), and the degree distribution (section 3.1.8) as measures to quantify and compare the (complete) data obtained through simulations.[2]

---

[2]The measures cannot be directly adapted for experimental data, see section 3.1.6.

3.1.1 Preparing Fluorescent Microscopy Data for Evaluation

Aggregation of cells and patterns of cellular aggregates are observed using different kinds of microscopy techniques. Most of them use fluorescence to label cells or parts of cells and the data to be analyzed consists of a series of microscopy images. The pixels of these images typically are encoded with either 8 bit or 16 bit for each color channel, meaning that each pixel can have 256 or 65536 shades of gray between completely dark and bright.
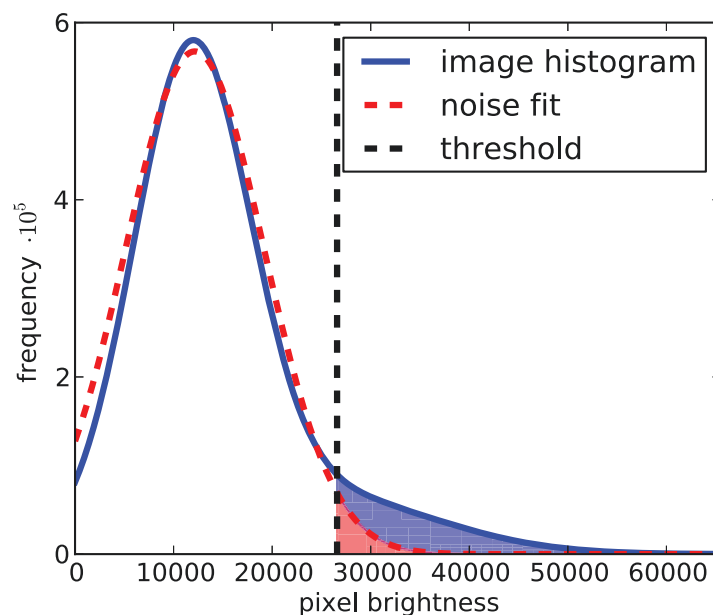
Figure 3.1: Preparation of microscopy data. The method is based on a histogram of pixel brightnesses shown as the solid blue line. Assuming that most pixels can be assigned to the background of the image (and therefore identified as noise) a fit of a Gaussian distribution to the data around the peak (shown as the red dashed line) yields a good approximation for the distribution of pixels without cells. Choosing a probability to assign a background pixel as a cell pixel (red area) one can directly calculate the threshold value for identifying pixels belonging to a fluorescent signal (blue area).

The task is to automatically discriminate pixels that belong to cells from those that belong to the background of the image. If the image mainly consist of background, i.e. most of the pixels do not belong to cells, one can estimate a threshold value as shown in figure 3.1. The

peak in the brightness histogram of the image belongs to the background. Assuming it to be a Gaussian noise one can fit this peak with a Gaussian distribution to get an estimate for the pixels that belong to the background. The choice of a probability p_{fp} to falsely assign a background pixel as a cell pixel directly leads to an estimate of the threshold by solving

$$p_{fp} - \int_{-\infty}^{x_{th}} ds \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left[\frac{s-\mu}{\sigma}\right]^2\right) = 0, \tag{3.1}$$

where μ and σ are obtained by the fit. Every pixel that is brighter than x_{th} belongs to a cell with a probability of $1 - p_{fp}$.

This method works good for images with little areas of fluorescent signals. If the image is mostly occupied by pixels belonging to a signal this method will fail to yield a good threshold. However, the images used in this thesis all fulfill the requirements for this method.

3.1.2 Boxcount Dimension

The first important parameter used in fractal theory is the fractal dimension. This can be calculated from the experimental data using a box-counting method.

The box-counting-dimension D_f is given by the limit of number $N(\varepsilon)$ of boxes belonging to the fractal divided by the invert of the box-size ε for box-size $\varepsilon \to 0$.

$$D_f = \lim_{\varepsilon \to 0} \frac{\log(N(\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)} \tag{3.2}$$

Bunde and Havlin (1994) introduced the box counting method as a method to measure the fractal dimension of a computer generated or digitalized fractal picture. Practically one draws a grid on the fractal with some lattice constant ε_1 and counts the number of boxes $N(\varepsilon_1)$ that are necessary to cover the entire fractal. Increasing the lattice constant in a series $\varepsilon_1 < \varepsilon_2 < ... < \varepsilon_n$ and calculating the numbers of boxes each time, one can calculate the fractal dimension by plotting $N(\varepsilon)$ versus $\frac{1}{\varepsilon}$ in a double logarithmic plot because $N(\varepsilon)$ scales as

$$N(\varepsilon) \sim \varepsilon^{-D_f}. \tag{3.3}$$

The minimum and maximum lattice constants must be chosen according to the fractal. The minimum lattice constant should be larger than the smallest unit of the quasi fractal, while the maximum lattice constant must be sufficiently small to avoid all boxes being necessary to cover the fractal.

3.1.3 Pair-correlation Function

The density-density autocorrelation $c(\vec{r})$ can be used as a tool to analyze the scaling of mass in a fractal object. The derived pair-correlation function $g(r)$ represents the probability to find a particle at distance r to a reference particle.

I define the density-density autocorrelation according to ben-Avraham and Havlin (2004) as

$$c(\vec{r}) = \frac{1}{\Phi} \sum_{\vec{r}_i} \rho(\vec{r}_i)\, \rho(\vec{r} - \vec{r}_i), \tag{3.4}$$

where the density $\rho(\vec{r}) = 1$ if $\vec{r}$ is part of the fractal and $\Phi = \sum_{\vec{r}_i} \rho(\vec{r}_i)$ is a normalization factor. If we consider a set of N particles the density can be rewritten as $\rho(\vec{r}) = \sum_{i=j}^{N} \delta(\vec{r} - \vec{r}_j)$, where $\delta(\vec{r}) = 1$ if $\vec{r} = \vec{0}$ and $\Phi = N$. Plugging all this back into (3.4) and dividing by the overall particle density number $\rho_0 = \frac{N}{V}$ we get

$$c'(\vec{r}) = \frac{V}{N^2} \sum_{i} \sum_{j \neq i} \delta(\vec{r} - \vec{r}_{ij}) \tag{3.5}$$

Averaging this function over the angles and dividing by the volume of a spheric shell of radius r and thickness Δr yields a probability, the pair-correlation function for a three dimensional fractal

$$g_{3D}(r) = \frac{V}{N^2 4\pi r^2 \Delta r} \left\langle \sum_i \sum_{j \neq i} \delta(\vec{r} - \vec{r}_{ij}) \right\rangle_{\text{angles}}. \tag{3.6}$$

For a two-dimensional fractal we have to consider areas instead of volumes and can show analogously that in this case the pair-correlation is

$$g_{2D}(r) = \frac{A}{N^2 2\pi r \Delta r} \left\langle \sum_i \sum_{j \neq i} \delta(\vec{r} - \vec{r}_{ij}) \right\rangle_{\text{angles}}. \tag{3.7}$$

To calculate $g(r)$ for experimental data, i.e. digitalized images, I used the digitalized and binarized images (see section 3.1.1), extracted the positions $\vec{r}_i$ of each pixel and ran a brute force algorithm.

Veatch et al. (2012) introduced a much faster way of approximately calculating the pair-correlation for two dimensional data using Fourier transformations. It uses the binary image $I(\vec{r})$ which is extended with zeros in every direction for a distance of r_{valid} that defines the range in which the approximation is valid. They define a window function $W(\vec{r})$ that is 1 if $\vec{r}$ is part of the original image and 0 else.

The two dimensional pair-correlation function then can be calculated using simple Fourier transformations $\mathscr{F}$ and inverse transformations $\mathscr{F}^{-1}$:

$$g(r) = \frac{\mathscr{F}^{-1}\left(|\mathscr{F}(I(\vec{r}))|^2\right)}{\rho^2 \mathscr{F}^{-1}\left(|\mathscr{F}(W(\vec{r}))|^2\right)}, \tag{3.8}$$

where ρ is the original image density. This considerably reduces the computational complexity from $\mathscr{O}(N^2)$ to $\mathscr{O}(N \log N)$, where N is the number of bright pixels.

The pair-correlation function yields a good measure for the size of aggregates. Because it decays almost exponentially for fractal objects, the slope of this decay can be used as a measure for the aggregate size by calculating the full width at half maximum (FWHM) of $g(r)$ numerically by fitting an exponential function to the data.

3.1.4 Simulated Attack

Beside the fractal structure of the aggregates I am also interested in the distribution of the different specimen inside, especially the relation between the fraction of two specimen inside the aggregate and their fraction on the surface. This is important to analyze the advantage or disadvantage of cheater cells inside a floc. The overall fraction of specimen ξ can be calculated straight forward by dividing the number of cells of one specimen $N_{\text{spec.1}}$ by the number of cells of the total number of cells $N_{\text{tot.}} = N_{\text{spec.1}} + N_{\text{spec.2}}$:

$$\xi_{\text{aggr.}} = \frac{N_{\text{spec.1}}}{N_{\text{tot.}}}. \tag{3.9}$$

To measure the fraction on the surface I perform a "simulated attack" on the aggregate by putting a random walker on the edge of the experimental microscopy data. This random walker represents a diffusing particle and will at some point hit the aggregate. Tracking the specimen hits by the walkers yields the fraction of specimen on the surface of the aggregate by dividing the number of hits $H_{\text{spec.1}}$ and $H_{\text{spec.2}}$ of each specimen:

$$\xi_{\text{surface}} = \frac{H_{\text{spec.1}}}{H_{\text{tot.}}}. \tag{3.10}$$

Keeping in mind, that a raise in concentration of species one results in an equal decline of concentration in species two, the shift $\Delta\xi_{\text{spec.1}} = 2\left(\xi_{\text{aggr.}} - \xi_{\text{surface}}\right)$ indicates how much more species one is exposed to environmental effects compared to species two.

3.1.5 Aggregation Degree

The aggregation degree is simply defined as the fraction of cells in clusters. To apply this measure to microscopy data one has to discriminate between single cells and aggregates, preferably with an automated method. Because this is computationally complicated I decided to apply this method of measuring the aggregation degree only to simulation data.

In the case of yeast flocculation I used another experimental way apart from microscopy to measure the aggregation degree. Since flocs settle much faster than single cells, one can measure the aggregation degree by performing settling experiments. Starting with a defined number of cells in a solution, one starts by inducing the flocculation and vortexing. After the flocks had time to settle one pulls a sample of the supernatant and counts the cells inside it. This number divided by the original number of cells yields the aggregation degree

$$G_{\text{Aggr.}} = \frac{N_{\text{supern.}}}{N_{\text{total}}}. \tag{3.11}$$

3.1.6 Clustering Coefficient

The clustering coefficient C like the measures described in the following sections is based on characterization of complex networks as reviewed by Albert and Barabási (2002). To apply these measures to a system, one has to provide the full graph data, i.e. the nodes and the connections which represent the cells and the bonds between cells in the studied aggregates.

Albert and Barabási (2002) defined the clustering coefficient C_i for an individual network node i. If we consider node i to be connected with k_i edges to k_i neighboring nodes, the total number of possible edges between these nodes would be $\frac{k_i(k_i-1)}{2}$. The individual clustering coefficient is then defined as the ratio of existing edges in this neighborhood E_i and the possible edges:

$$C_i = \frac{2E_i}{k_i(k_i-1)}. \tag{3.12}$$

The total clustering coefficient of the network is then defined as the average over the individual cluster coefficients:

$$C = \langle C_i \rangle_i = \frac{1}{N} \sum_{i=1}^{N} \frac{2E_i}{k_i(k_i-1)}. \tag{3.13}$$

For a random graph, where two nodes have a probability of p of being connected, the clustering coefficient is equal to this probability, and therefore

$$C_{\text{rand}} = p = \frac{\langle k_i \rangle_i}{N}. \tag{3.14}$$

This value can be used as a reference value to characterize the cellular aggregates.

3.1.7 Cluster sizes and Average Path Lengths

The first step in characterizing individual cell clusters is to identify them within the simulation data. The necessary data structure for this provides the cells and the bonds between cells. Ideally these are represented in a graph-like or tree-like structure where cells are represented by nodes and bonds are represented by edges.

To find clusters within these kinds of structures Tarjan (1972) proposed an algorithm[3] that applies a depth first search on the data starting at a random node storing all visited nodes and removing them from the initial dataset. After all nodes of a cluster have been visited, the cluster is identified. Algorithm 1 shows the algorithm in pseudo code.

Algorithm 1: Tarjan's algorithm in pseudo code. This is the simplified version to find clusters instead of strongly connected components.

   Data: node i, pointers to nodes connected to i, node-stack
   Result: filled node-stack
1 node-stack $\rightarrow$ add node i
2 **for** $j \in$ *neighbors of i* **do**
3   **if** *node $j \notin$ node-stack* **then**
4   call Tarjan's algorithm for node j
5 **return** *node-stack*

If the algorithm identified all clusters in the dataset one can easily analyze the statistics of clusters, i.e. the size distributions.

---

[3]The original algorithm describes a method to find strongly connected components within a graph. I use a simplified version to find clusters within graphs.

Another quantity that can be used to quantify the clusters is the average path length $\langle l \rangle$. For each cluster it is defined as the average number of nodes to pass when going from one node to another without visiting the same node twice.

3.1.8 Degree Distribution

Albert and Barabási (2002) define the degree of a node i as the number of edges k_i connecting it to other nodes. The distribution of degrees $P(k)$ is the relative occurrence of nodes with exactly k connections and therefore equals the probability that a randomly selected node has exactly k direct neighbors.

For a lot of complex networks this distribution scales for large k as a power law:

$$P(k) \sim k^{-\gamma}. \tag{3.15}$$

Networks that show this behavior are called scale free networks and their power law parameter γ can be used for comparison. Using this scaling one can also characterize networks between random graphs that show a Poisson distributed degree and these scale free networks.

3.2 An Individual-Based Model for Cellular Aggregation

Individual-based modeling – also called agent-based modeling – is a bottom-up modeling technique that has its roots in social sciences and ecology. Bonabeau (2002) and Grimm et al. (2005) defined individual-based modeling as a technique that uses a collection of autonomous decision-making entities called agents.

The first step in individual-based modeling is to collect and compile the information on the individuals and to formulate theories about their behavior and their interactions. After implementing this set of individuals and rules in a computer simulation one can observe the emergence of system-level properties. Therefore, individual-based models can be used to understand the connection between the properties of an individual and interactions and the behavior of a complex system.

In case of cellular aggregation the natural choice for the individual to base the models upon are the cells themselves. The major task in building an individual-based model is to find the properties and interactions in a quantitative way.

Thorne et al. (2007) nicely described the challenges that one encounters using this bottom-up approach to simulate multicellular dynamics and patterns. The biggest challenge is to determine the simulation parameters. Obviously the literature is the first source to acquire the necessary parameters, but in many cases these parameters have not been measured (or published) or they are conflicting in different publications. The two ways to approach this problem are performing the experiments either in collaboration or, if possible, by oneself or use an estimate of the parameters and perform a stability analysis.

Another important challenge is the computational complexity of individual based models that rapidly increases with the number of individuals and the number of rules used to model a system. However, a lot of new possibilities arise because the technology grows in power very fast.

In this study I used the principles of molecular dynamics simulations described in section 3.2.1 as a starting point to define a set of rules for the interactions between cells.

To overcome the issues with the computational complexity in the simulation of yeast flocculation, I used the highly parallel implementation of the algorithm introduced in section 3.2.3 and published by Leinweber et al. (2014) that facilitates graphic processors to perform the necessary calculations.

3.2.1 Principles of Molecular Dynamics Simulations

Molecular dynamics simulations are a special kind of individual-based simulations, where the core individuals are molecules or atoms, which are referred to as particles. The particles are defined by their physical properties like position, velocity, or forces applied to them.[4]

The rules that define the interactions are derived from the physical interactions between particles and translated to forces acting on the individuals.

The final step is to formulate the equations of motion for each particle and to solve them. Newtonian mechanics seems to be the fitting tool to formulate these equations, since I interpret the rules as forces that drive the behavior of a particle.

If all relevant forces are identified and the equations are set up in case of a three dimensional system of N particles the task is to solve $3N$ coupled differential equations numerically.

To do this I used the Störmer-Verlet integration method, that discretizes time and integrates the equations in a way that all current properties are available for evaluation in every iteration step (see section 3.2.1.1 for details).

A computationally very expensive step is the one-to-one comparison that is necessary to calculate distance dependent forces, e.g. for a Coulomb interaction. The complexity for these calculations is $\mathscr{O}\left(N^2\right)$. To reduce this complexity I use a linked cell method and a cutoff radius as described by Griebel et al. (2004). The idea of this method is to cover the simulation space

---

[4]For an introduction to molecular dynamics simulations see "Molecular Modelling" by Leach (2001)

with a grid and use the cutoff radius as the lattice constant. All cells are assigned to a lattice site according to their position. If the interaction of cells further apart than the cutoff radius is neglected, the number of comparisons is reduced to the number of cells at the lattice site of the reference cell plus the number of cells in the neighboring lattice sites.

3.2.1.1 The Störmer-Verlet Integration Method

To numerically integrate Newton's equation of motion one has to discretize time first. The Störmer-Verlet method of discretization is an efficient and stable algorithm to achieve this goal. In this section I will shortly describe the derivation of this algorithm by Griebel et al. (2004) to provide the reader with a better understanding of the simulations.

Starting with Newton's classical equation of motion

$$m\frac{d^2}{dt^2}\vec{r}(t) = \vec{F}\left(\vec{r}, \frac{d}{dt}\vec{r}, t\right) \tag{3.16}$$

where $r(t)$ is the position, m is the mass and $\vec{F}$ is the force acting on the particle, and using the time discretization $t_{n+1} = t_n + \delta t$ the second derivative is approximately

$$\frac{d^2}{dt^2}\vec{r} \approx \frac{1}{\delta t^2}\left(\vec{r}(t_{n+1}) - 2\vec{r}(t_n) + \vec{r}(t_{n-1})\right). \tag{3.17}$$

If one discretizes the velocity $\vec{v} = \frac{d}{dt}\vec{r}$ in time using the central difference

$$\vec{v}(t_n) = \frac{\vec{r}(t_{n+1}) - \vec{r}(t_{n-1})}{2\delta t} \tag{3.18}$$

and plugs this into the discretized form of (3.16) one can derive two equations to calculate $\vec{r}(t_{n+1})$ and $\vec{v}(t_{n+1})$:

$$\vec{r}(t_{n+1}) = \vec{r}(t_n) + \delta t\vec{v} + \frac{\vec{F}(t_n)\delta t^2}{2m} \tag{3.19}$$

$$\vec{v}(t_{n+1}) = \vec{v}(t_n) + \frac{\left(\vec{F}(t_n) + \vec{F}(t_{n+1})\right)\delta t}{2m}. \tag{3.20}$$

These equations are solved in the sequence indicated in algorithm 2. From this sequence one can easily see, that all necessary parameters are available for further evaluations in every time-step.

Algorithm 2: Störmer-Verlet integration Method, adapted from (Griebel et al., 2004)

Data: Initial conditions $\vec{r}_{i,0}$ and $\vec{v}_{i,0}$ for each particle i, time-step size δt and ending time t_{end}

1 initialize vectors $\vec{F}_{i,\text{old}}$ and $\vec{F}_i$
2 calculate forces $\vec{F}_i$
3 **while** $t < t_{\text{end}}$ **do**
4 $\quad$ $t = t + \delta t$ **for** *every particle i* **do**
5 $\quad\quad$ calculate new position $\vec{r}_i$
6 $\quad\quad$ set $\vec{F}_{i,\text{old}} = \vec{F}_i$
7 $\quad$ calculate forces $\vec{F}_i$
8 $\quad$ **for** *every particle i* **do**
9 $\quad\quad$ calculate new velocity $\vec{v}_i$ using $\vec{F}_i$ and $\vec{F}_{i,\text{old}}$
10 $\quad$ calculate system parameters (e.g. kinetic energy, potential energy)
11 $\quad$ output of t, $\vec{r}_i$, $\vec{v}_i$, $\vec{F}_i$ and all system parameters

---

3.2.2 Model Description

As motivated above, the first step in building an individual-based model is to identify the individuals and the rules they are obeying. I already mentioned the cells being the natural choice for the individuals when it comes to cellular aggregation. This also comes with the big advantage that single cells are traceable through the whole aggregation process which allows for studying the dynamics and the internal structure of cellular aggregation and cellular aggregates.

The rules that the cells obey during aggregation partly depend on the system. While rules like "two cells can not be in the same place" and "bonds result in attractive forces between cells" apply for all systems there are some rules that are specific to a special system. A good example for these specific rules is the interaction with gravity: while it has an impact in yeast flocculation, it is not important for the two dimensional aggregation of MEF-Cells.

Most of the rules can be derived as forces and plugged into an equation of motion. However, rules like "cells grow at a constant rate" have to be handled in a fitting manner.

Another important issue is the numerical division by very small values. These kind of divisions occur quite often in forces, that point into the direction of the difference vector $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$ of two cells. These numerical singularities can be avoided in two ways: (i) rewriting the forces to avoid the division or (ii) add a small singularity parameter ε in the calculation of the norm of the vector (which I did to avoid numerical problems):

$$\left| \vec{r}_{ij} \right| = \sqrt{\left(\vec{r}_i - \vec{r}_j \right)^2} \approx \sqrt{\left(\vec{r}_i - \vec{r}_j \right)^2 + \varepsilon^2}, \tag{3.21}$$

$$\varepsilon \ll 1. \tag{3.22}$$

3.2.2.1 A 2D Model for MEF-Cell Aggregation

In the case of MEF-cell aggregation one deals with one cell type, which will be the individual for the model. The properties of each cell are its position, its velocity, its bonds to other cells, its interaction radius, and the growth rate of its interaction radius.

The first simplification I made here is that the cells do not move by themselves but rather increase their interaction radius. This is done linearly with a growth rate ρ_i that characterizes the motility of the cell i. If the interaction radii of two cells i and j overlap they form a bond with probability p_{bond} and exert an attractive force on each other represented by a spring:

$$\vec{F}_{ij} = -k \frac{r_{ij} - \left(R_i + R_j \right)}{\sqrt{r_{ij}^2 + \varepsilon^2}} \vec{r}_{ij}. \tag{3.23}$$

The total force exerted on a cell by bonds is calculated by summing equation (3.23) over all bound cells j. Bonds break if the two cells move apart to far or it may break spontaneously with a probability p_{break}.

Legend:
—— cell diameter
······ interaction range
—— bond (atractive force)
—— collision distance (repulsive force)
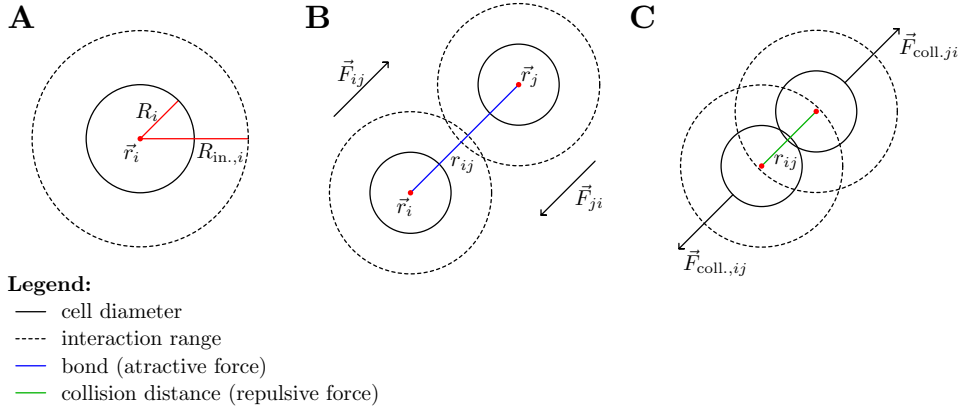
Figure 3.2: Schematic of the MEF-cell aggregation model. **A** The cell is characterized by its radius R_i, its interaction radius $R_{\text{in},i}$, its position $\vec{r}_i$, and its velocity $\vec{v}_i$ (not shown). **B** If two cells interaction radii overlap, they will form a bond (blue line) which results in an attractive force between the two cells. The force on each cell is equal except for their sign $\vec{F}_{ij} = -\vec{F}_{ji}$.**C** If two cells approach each other and get closer than the sum of their individual radii they will experience a repulsive force. This prevents two cells from occupying the same space.

The second force a cell experiences is the drag of the substratum. It should be proportional to the area covered by the cell and decelerate it's motion. I defined it as a friction force

$$\vec{F}_{\text{drag},i} = -\pi R_i^2 \gamma \vec{v}_i, \tag{3.24}$$

where γ is a friction parameter normalized to a unit area.

To prevent cells from moving through space blocked by other cells I introduced a repulsive force that is only exerted on cells that are closer to each other than the combined cell radius $R_{ij} = R_i + R_j$

$$\vec{F}_{\text{coll.},ij} = \begin{cases} k_{\text{coll.}} \dfrac{r_{ij} - l_0}{\sqrt{r_{ij}^2 + \varepsilon^2}} \vec{r}_{ij} & \text{if} \quad r_{ij} < R_{ij} \\ 0 & \text{else} \end{cases}. \tag{3.25}$$

Again the total force experienced by a cell because of collisions is the sum over all other cells j that are in range.

Module	Equation	Explanation
radius of cell i	R_i	$\approx 25\,\mu\mathrm{m}$ Rishal et al. (2012)
interaction radius of cell i	$R_{\mathrm{in.},i}$	$R_{\mathrm{in.},i}\,(t=0)=R_i$
growth rate of $R_{\mathrm{in.},i}$	ρ_i	drawn from normal distribution
total force on cell i caused by bonds	$\vec{F}_{\mathrm{bond},i} = \displaystyle\sum_{j\in\mathrm{bonds}_i} -k\frac{r_{ij}-(R_i+R_j)}{\sqrt{r_{ij}^2+\varepsilon^2}}\vec{r}_{ij}$	k defines the strength of bonds
drag of cell i	$\vec{F}_{\mathrm{drag},i} = -\pi R_i^2\gamma\vec{v}_i$	γ defines the strength of the drag per unit area
total force on cell i caused by collisions	$\vec{F}_{\mathrm{coll.},i} = \displaystyle\sum_{j=1}^{N} k_{\mathrm{coll.}}\frac{r_{ij}-l_0}{\sqrt{r_{ij}^2+\varepsilon^2}}\vec{r}_{ij}\Theta\left(R_{ij}-r_{ij}\right)$	$k_{\mathrm{coll.}}$ is the spring constant of a cell, Θ is the Heaviside function

Table 3.1: Summary of the MEF-cell aggregation model.

Table 3.1 summarizes the model components. The implementation of the model requires additional parameters that control the integration (ending time t_{end} of the integration, time-step δt), the simulation space (length in x and y direction), and the initialization of cells (distribution of cell size, growth rate, number of cells).

3.2.2.2 A 3D Model for Yeast Flocculation

In case of aggregating yeast cells there are two types of cells represented by two different kinds of individuals: bearer cells that express the Flo5 gene and produce binding proteins and cheater cells that do not produce binding proteins. This difference is implemented by three different

binding types T_{ij}: cheater-cheater bonds (no attractive interaction) $T_{ij} = \{\text{cheater} - \text{cheater}\}$, bearer-bearer bonds $T_{ij} = \{\text{bearer} - \text{bearer}\}$, and mixed bonds $T_{ij} = \{\text{bearer} - \text{cheater}\}$ that characterize the attractive interaction.

The properties of each individual are essentially the same as in the MEF-cell aggregation model except that now a cell type has to be included and the reaction radius is fixed. The properties are position, velocity, bonds to other cells, reaction radius and cell type (bearer or cheater).

Under laboratory conditions the cells are mixed by a vortex after the flocculation was initiated. This rapid vortexing leads to turbulent flows and eddies in the liquid media causing the cells to move almost randomly.

I implemented this random movement using two random forces given by random vectors with a uniformly distributed direction and a normally distributed length. One random force is directly applied to the individual cells while the other is applied a whole spatial region in order to move larger aggregates. Another small force that depends on the distance to the z-axis was applied to mimic an eddie resulting from mixing the cells.

If the distance $\left|\vec{r}_{ij}\right|$ between two cells i and j is smaller or equal to the sum of their radii $R_i + R_j$, then the cells interact repulsively. The associated repulsive collision force is given by a damping term, that accounts for inelastic effects, a shear term that incorporates a possible scattering, and a direct elastic repulsion:

$$
\begin{aligned}
\vec{F}_{\text{coll.},ij}(t) \;=\; & \left[\gamma_{\text{coll.}} \vec{v}_{ji}(t) + \sigma_{\text{coll.}} \left(\vec{v}_{ji}(t) - \frac{\vec{r}_{ij}(t)}{\left|\vec{r}_{ij}(t)\right|} \left(\vec{v}_{ji}(t) \cdot \frac{\vec{r}_{ij}(t)}{\left|\vec{r}_{ij}(t)\right|} \right) \right) \right. \\
& \left. - \; k_{\text{coll.}} \left(R_j + R_i - \left|\vec{r}_{ij}(t)\right| \right) \frac{\vec{r}_{ij}(t)}{\left|\vec{r}_{ij}(t)\right|} \right] \Theta \left(R_j + R_i - \left|\vec{r}_{ij}\right| \right).
\end{aligned}
\tag{3.26}
$$

The spring factor $k_{\text{coll.}}$, the damping $\gamma_{\text{coll.}}$ and the shear factor $\sigma_{\text{coll.}}$ are adjustable variables, Θ is the Heaviside function and $\vec{v}_{ij} = \vec{v}_i - \vec{v}_j$ denotes the velocity difference.

If the distance between two cells is smaller or equal to the sum of their reaction radii $R_{\mathrm{react},i} + R_{\mathrm{react},j}$, then the cells will bind to each other with a probability depending on the types of the two cells. Two bearer cells will have a high probability to bind while two cheater cells can not form bonds since they do not produce the necessary protein (see Figure 3.3).
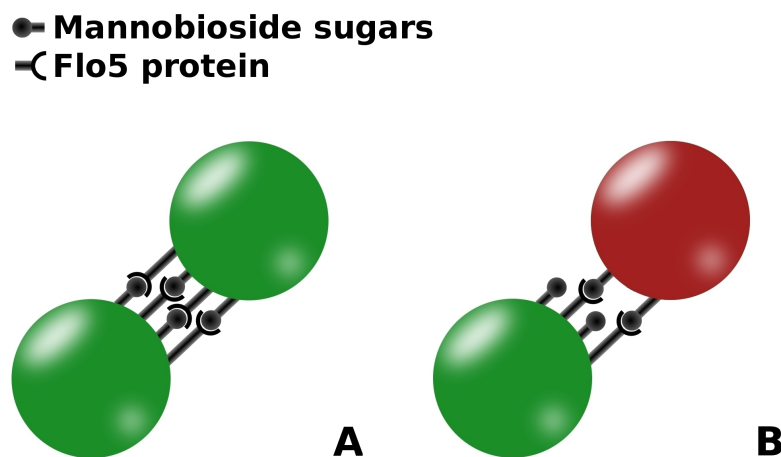
Figure 3.3: Illustration of the two different types of bonds. **A** shows the bonds between two bearer cells while **B** shows the bonds between cheater (red) and bearer cells. Cheaters do not produce the binding protein but they do have the anchors on their cell surface

.

Again, the binding forces were approximated by harmonic oscillators. Thus, the bonds carry out spring forces with spring constant k_{ij} and free length l_{ij}. According to Hooke's law, the force is

$$\vec{F}_{\mathrm{bond},ij}(t) = -k_{ij}\left(\left|\vec{r}_{ij}(t)\right| - l_{ij}\right)\frac{\vec{r}_{ij}(t)}{\left|\vec{r}_{ij}(t)\right|} \tag{3.27}$$

In principle, k_{ij} and l_{ij} may differ for each pair of cells i and j, but there are only two important types of bonds, those between two bearers and those between a bearer and a cheater cell. Because the spring forces of parallel springs add up and the number of bonds of a mixed pair is approximately half the number of bonds of pair of bearer cells, we can set $k_{ij} = k$ for bearer-bearer bindings and $k_{ij} = \frac{k}{2}$ for bearer-cheater bindings (see Figure 3.3).

To avoid a gain of energy in the system through the formation of bonds and the exertion of random forces I introduced a friction force

$$\vec{F}_{\text{fric},i}(t) = -\Gamma \vec{v}_i(t), \tag{3.28}$$

where Γ is the friction parameter. The total force acting on cell i at time t is obtained by summing over all the forces:

$$\vec{F}_i(t) = \vec{F}_{\text{rand},i}(t) + \vec{F}_{\text{fric},i}(t) + \sum_{j \in B_i} \vec{F}_{\text{bond},ij}(t) + \sum_{j=1, j \neq i}^{N} \vec{F}_{\text{coll},ij}(t) \tag{3.29}$$

where the first sum extends over all cells B_i that are bound to cell i, i.e. their distance is smaller than the sum of the reaction radii.

Table 3.2 summarizes the model components. The implementation of the model requires additional parameters that control the integration (e.g. ending time t_{end} of the integration, time-step δt), the simulation space (length in x, y and z direction), and the initialization of cells (e.g. distribution of cell sizes, number of cells).

3.2.2.3 Parameter Estimation

For the yeast flocculation model it is possible to estimate some parameters especially the amplitude of the random force which is hard to measure experimentally.

In principle cells are spheres with a certain elasticity. An encounter of two cells can be described as an inelastic collision. During the time this collision takes place there is a certain time depended contact area $A_{\text{cont}}(t)$ of the two cells. Given a constant protein length l_P this yields a reaction volume $V(t) = l_P A_{\text{cont}}(t)$.

If the molecular forming rate k_{mol} of bonds is known, calculating the collision time t_c and the average reaction volume one can estimate the number of bonds formed during a collision.

Module	Equation	Explanation
radius of cell i	R_i	$\approx 4\,\mu\mathrm{m}$
reaction radius of cell i	$R_{\mathrm{react},i}$	a fixed multiple of the cell radius R_i
binding probabilities	$p_{\mathrm{b-b}}$ and $p_{\mathrm{b-c}}$	the binding probabilities are very similar
unbinding probabilities	$p_{\mathrm{break,b-b}}$ and $p_{\mathrm{break,b-c}}$	$p_{\mathrm{break,b-b}} < p_{\mathrm{break,b-c}}$
total force on cell i caused by collisions	$\begin{aligned} \vec{F}_{\mathrm{col.},i} = & \sum_{j\neq i} \Bigg[\gamma_{\mathrm{col}}\vec{v}_{ji}(t) + \sigma_{\mathrm{col}}\Bigg(\vec{v}_{ji}(t) \\ & - \frac{\vec{r}_{ij}(t)}{\left|\vec{r}_{ij}(t)\right|}\left(\vec{v}_{ji}(t)\cdot\frac{\vec{r}_{ij}(t)}{\left|\vec{r}_{ij}(t)\right|}\right)\Bigg) \\ & - k_{\mathrm{col}}\left(R_j+R_i-\left|\vec{r}_{ij}(t)\right|\right)\times \\ & \times\frac{\vec{r}_{ij}(t)}{\left|\vec{r}_{ij}(t)\right|}\Bigg]\Theta\left(R_j+R_i-\left|\vec{r}_{ij}\right|\right) \end{aligned}$	$k_{\mathrm{coll.}}$ is the spring constant of a cell, Θ is the Heaviside function
bond strengths	$k\left(T_{ij}\right) = \begin{cases} k & \text{if } T_{ij} = \{\text{bearer} - \text{bearer}\} \\ \frac{k}{2} & \text{if } T_{ij} = \{\text{bearer} - \text{cheater}\} \end{cases}$	T_{ij} denotes the combination of cell types involved in the bond
total force on cell i caused by bonds	$\vec{F}_{\mathrm{bond},i} = -\sum_{j\in B_i} k\left(T_{ij}\right)\left(\left|\vec{r}_{ij}(t)\right|-l_{ij}\right)\frac{\vec{r}_{ij}(t)}{\left|\vec{r}_{ij}(t)\right|}$	B_i denotes the list of cells bound to cell i
drag of cell i	$\vec{F}_{\mathrm{fric},i} = -\Gamma\vec{v}_i$	Γ defines the rate at which kinetic energy is removed from the system

Table 3.2: Summary of the yeast flocculation model.

If the resulting binding energy is larger than the kinetic energy, the cells bind, otherwise they will separate further away and stay unbound.

As a simplification of the system I only considered central collisions and reduced it to one dimension. The starting point is the differential equation of a dampened harmonic oscillator

$$\ddot{x} + \frac{\Gamma}{m}\dot{x} + \frac{k_{\text{cell}}}{m}x = 0, \tag{3.30}$$

where Γ is the friction parameter, k_{cell} is the spring constant of the cell, and m is the cell mass.

Integrating equation (3.30) from time t_0 where the two cells are in contact in one single point ($x(t_0) = l_0$) with an initial collision speed $\dot{x}(t_0) = v_0$ one calculates the time dependent reaction volume

$$V(t) = \left(\frac{\pi l_B}{4}\right)\left(l_0^2 - \left(l_0 - \frac{v_0}{\omega}\exp(-\gamma t)\sin(\omega t)\right)^2\right), \tag{3.31}$$

where $\gamma = \frac{\Gamma}{2m}$ and $\omega = \sqrt{\frac{k_{\text{cell}}}{m} - \gamma^2}$.

To calculate the average reaction volume as a function of the initial collision speed one has to average $V(t)$ over the collision time $t_{\text{c}} = \frac{\pi}{\omega}$:

$$V_{\text{av}} = \frac{l_B v_0 \omega \left[8\left(1 + \exp\left(-\frac{\gamma\pi}{\omega}\right)\right)\gamma l_0 + \left(\exp\left(-\frac{2\gamma\pi}{\omega}\right) - 1\right)v_0\right]}{16\gamma(\gamma^2 + \omega^2)}. \tag{3.32}$$

With this averaged volume and the collision time one can calculate the approximate number of bonds using a simple rate equation. This leads to a total number of

$$T = c_{\text{u},t_0}\exp\left(\frac{k_{\text{mol}}\pi}{\omega}\right)V_{\text{av}} \tag{3.33}$$

bonds, where c_{u,t_0} is the concentration of unbound proteins at time t_0 and k_{mol} is the rate constant.

If there is an energy of E_B per bond then the cells will stick together if $E_B T$ is smaller or equal the kinetic energy at time $t = \frac{\pi}{\omega}$. Equalizing binding energy and kinetic energy yields the maximum collision velocity for which the cells will bind:

$$
v_{\text{max}} = \frac{8 E_B \gamma l_0 l_B c_{u,t_0} \omega \exp\left(\frac{(\gamma+k)\pi}{\omega}\right) \cdot \left[1 + \exp\left(\frac{\gamma\pi}{\omega}\right)\right]}{8\gamma m \left(\gamma^2 + \omega^2\right) \exp\left(\frac{\gamma\pi}{\omega}\right) + E_B l_B c_{u,t_0} \omega \left[\exp\left(\frac{(2\gamma+k)\pi}{\omega}\right) - \exp\left(\frac{k\pi}{\omega}\right)\right]}. \tag{3.34}
$$

Plugging this into the Gaussian distribution with mean v_m and deviation σ for the velocities of the cells and integrating it from zero to v_{max} one obtains the probability of two colliding cells sticking together:

$$
p_{\text{glue}} = \frac{1}{2} \left[\text{erf}\left(\frac{v_m}{\sigma\sqrt{2}}\right) - \text{erf}\left(\frac{v_m - v_{\text{max}}}{\sigma\sqrt{2}}\right)\right]. \tag{3.35}
$$

Most of the used parameters are available in the literature (see table 3.3 for overview). Smith et al. (2000) studied the mechanical properties of yeast cells. From their data a spring-constant $k_{\text{cell}} \approx 5.4058\,\text{N/m}$ can be estimated in the linear regime of cell-deformation. The typical concentration of adhesive proteins on yeast measured by Dranginis et al. (2007) is about $c_{u,t_0} \approx 2.5 \cdot 10^4\,\text{proteins/cell}$.

Other parameters have been estimated by own experiments. The binding-energy of Flo5a and Mannose is about $6\,\text{kJ/mol}$, leading to an energy of $E_B \approx 8.33 \cdot 10^{-21}\,\text{J}$. I estimated the length of a Flo5a-Protein l_B to be about $120\,\text{nm}$. The diameter of a typical yeast cell as used is about $l_0 = 4\,\mu\text{m}$. From this the mass can be calculated as about 110% of the mass of a water-sphere with the same diameter. This leads to a mass $m = 3.686 \cdot 10^{-14}\,\text{kg}$ for the cell.

This leaves the friction parameter, the reaction-constant and the mean collision-speed as unknown parameters. To get an idea of the importance of these parameters, I analyzed the stability of v_{max} and p_{glue} to changes in these parameters.

As a starting point to estimate the friction parameter, one can use the viscosity of water and multiply it by a typical length-scale in the system, in our case the cell diameter $\Gamma = l_0 \eta_{\text{H}_2\text{O}}$.

This gives $10^{-9}\,\mathrm{kg/s}$ as an order of magnitude for Γ. Using this one obtains $10^{5}\,\mathrm{1/s}$ as an order of magnitude for γ.

As can be seen from equation (3.34) the rate constant k_{mol} mainly occurs in sums with γ. Considering a fast reaction in the order of $10^8/\mathrm{Ms}$ and given the initial concentration of proteins we obtain k_{mol} in the order of 10^2 which is much smaller than γ. So we can neglect it, and get the approximation for the maximal collision velocity:

$$v_{\mathrm{max}} = \frac{8E_{\mathrm{B}}\gamma l_0 l_B c_{\mathrm{u},t_0}\omega \exp\left(\frac{\gamma\pi}{\omega}\right)\cdot\left[1+\exp\left(\frac{\gamma\pi}{\omega}\right)\right]}{8\gamma m\left(\gamma^2+\omega^2\right)\exp\left(\frac{\gamma\pi}{\omega}\right)+E_{\mathrm{B}}l_B c_{\mathrm{u},t_0}\omega\left[\exp\left(\frac{2\gamma\pi}{\omega}\right)\right]}. \tag{3.36}$$

For the estimation of a good range for Γ and v_μ I analyzed the behavior of the binding probability p_{glue} for those parameters. If we increase the friction constant at some point all collisions will result in a binding of the two cells regardless of the collision speed (see figure 3.4). However, if we assume the friction caused by the collision to be less or equal 160 times the friction of water, we get mean velocities in the range of $0.1\,\mathrm{mm/s}$ and $10\,\mathrm{mm/s}$ for which there is a noticeable difference in the formation probabilities of bearer-bearer bonds and bearer-cheater bonds (see figure 3.5).

From the binding energy E_B and the bond length l_B one can calculate the spring constant for bonds. Starting from the protein concentration at the surface $c_{\mathrm{u},t_0} = 4.145\cdot10^{21}/\mathrm{m^3}$ and assuming densely packed cells (contact area is one twelfth of the whole cell surface) there is a total of roughly 2000 bonds between each two cells. This leads to a total binding energy of approximately $E_{\mathrm{total}} \approx 1.666\cdot10^{-17}\,\mathrm{J}$.

If one assumes that the proteins are capable of stretching twice their size before breaking, this happens at a distance of $2l_{\mathrm{B}}$ so the energy of a spring with spring constant k_{total} at this distance is $E_{\mathrm{Spring}} = 4k_{\mathrm{total}}l_{\mathrm{B}}^2$. Equalizing this with the total binding energy leads to a spring constant of $k_{\mathrm{total}} \approx 29\,\mathrm{dyn/m}$ for the cell-cell binding.
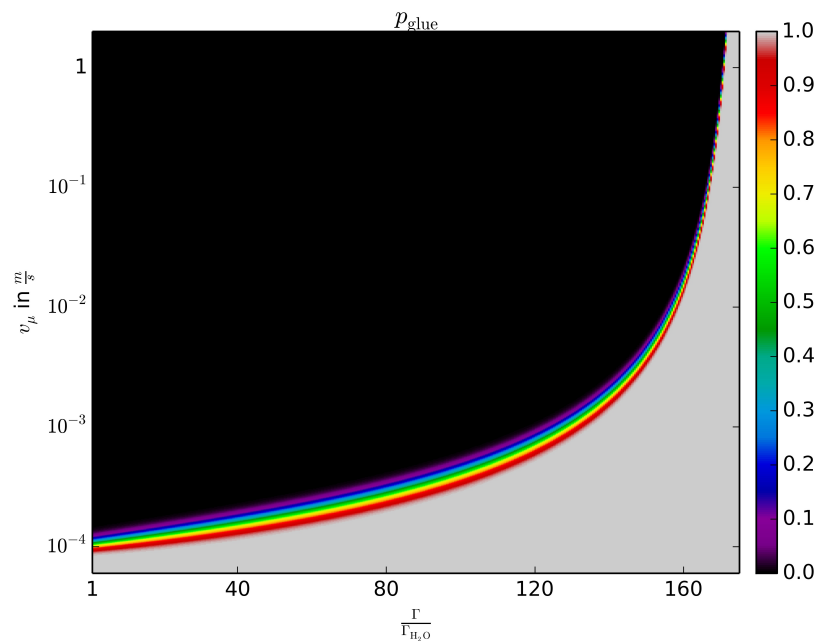
Figure 3.4: Binding probability as a function of the friction parameter Γ and the mean collision velocity v_μ. If the friction parameter is more than approximately 160 times bigger than $\Gamma_{H_2O} = 4\,\mu m \eta_{H_2O}$, the binding probability becomes independent of the mean collision velocity.

To convert the mean velocities into the used random forces, one can use $F_{rand} = \Gamma v_\mu$ as an estimate. This leads to a range for the amplitude of the random force from $F_{rand} \approx 0.3\,\mathrm{pN}$ to $F_{rand} \approx 6\,\mathrm{nN}$. Assuming it is enough to simulate with a time discretization of $\Delta t = 10\,\mu s$ per time-step one can estimate the mean amplitude of the random force to range from $F'_{rand} = 0.0002$ to $F'_{rand} = 4$ in system units using m_{cell} and l_0 for mass- and length-scale.

3.2.3 High Efficiency GPU Implementation of the Yeast Flocculation Model

My first attempt to implement the yeast aggregation model for validation was written in C/C++, based objects and object arrays for storing the cell data. Due to large runtimes starting at a few thousand simulated cells, the implementation was improved using OpenMP to utilize
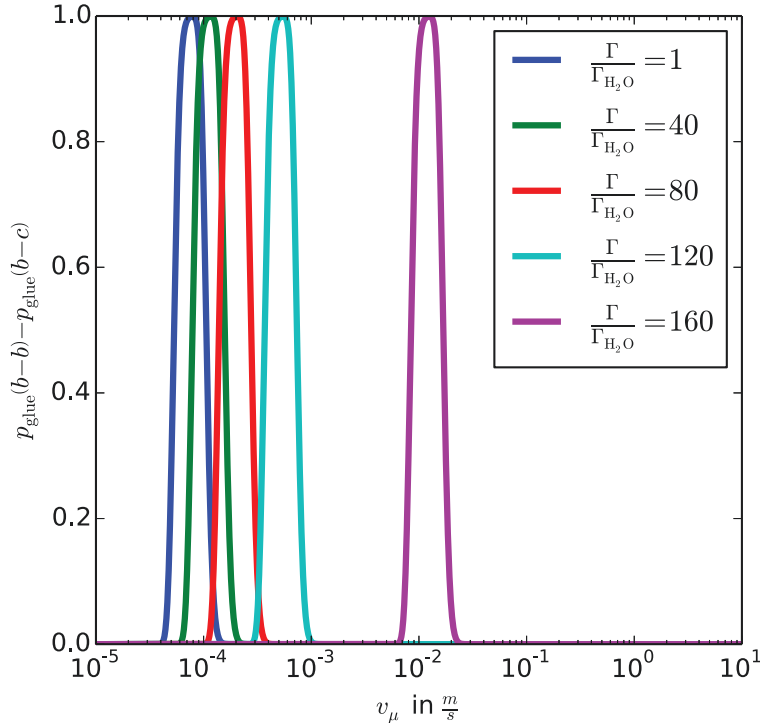
Figure 3.5: Difference in p_{glue} for bearer-bearer and bearer-cheater bonds. The peaks of the graphs give a good indication for the choice of a mean velocity $v_\mu \in [0.1\,\text{mm/s}, 10\,\text{mm/s}]$, because the forming probabilities of bearer-bearer and bearer-cheater bond should be discriminable at some point.

multi core architectures. Furthermore, I used space partitioning for reducing the computational overhead required for finding nearby cells. However, the time to simulate a reasonable amount of cells was enormous. This led to unwanted side effects. If the number of cells is to small, the size of the flocks is very limited. Consider for example a simulation of 500 cells. A flock in this simulation can never be larger than 500 cells; it is more likely that they will be substantially smaller. So to have a simulation with reasonable flock sizes one has to simulate a lot more cells than those that fit into one floc.

To solve this problem my collaborators from the department of computational sciences implemented the model using graphic processors. The methods I review in this section were published in (Leinweber et al., 2014). It was implemented it OpenCL to achieve a platform

Parameter	Value	Description	Source/Remarks
$c_{U,t=0}$	$4.145 \cdot 10^{21}/\mathrm{m^3}$	Flo5a conc.	*Draginis et. al. (2007)*
k_{cell}	$5.4058\,\mathrm{N/m}$	spring-constant	*Smith et. al. (2000)*
l_B	$120\,\mathrm{nm}$	bond-length	estimate
E_0	$8.33 \cdot 10^{-21}\,\mathrm{J}$	energy per bond	equivalent of $6\,\mathrm{kJ/mol}$, exp.
l_0	$4\,\mu\mathrm{m}$	cell-diameter	typical for haploid cells
m	$3.686 \cdot 10^{-14}\,\mathrm{kg}$	cell-mass	110% density of water
Γ	$\in [\Gamma_{\mathrm{H_2O}}, 160\Gamma_{\mathrm{H_2O}}]$	friction-parameter	p_{glue} independent of v_μ for higher values

Table 3.3: Used parameter values for the simulation.

independent program that will run on almost all available hardware.

The program flow of the simulation can be divided into four stages: 1) determine collisions and reactions 2) apply binding forces, 3) integrate the system and 4) determine cracked bonds. The first challenging computational problem in implementing the model is to find neighboring cells for each individual cell. This is needed to account for the cell-cell interactions, i.e. the reactions and collisions. The brute-force attempt to this would be to compare the position of each cell to the position of each other cell. For N cells, this would lead $\frac{N(N-1)}{2}$ comparisons to calculate all distances. For large n, both the required memory and the computational time are too high to be carried out on a GPU because of hardware limitations. To solve this problem, the simulation space has been divided into a 3-dimensional grid, as suggested by NVIDIA's Particle Collision Simulation[5].

To keep this program flow as efficient as possible all data structures have been adapted to fit the needs of a massively parallel algorithm. A programming challenge was the handling of

[5]https://developer.nvidia.com/opencl
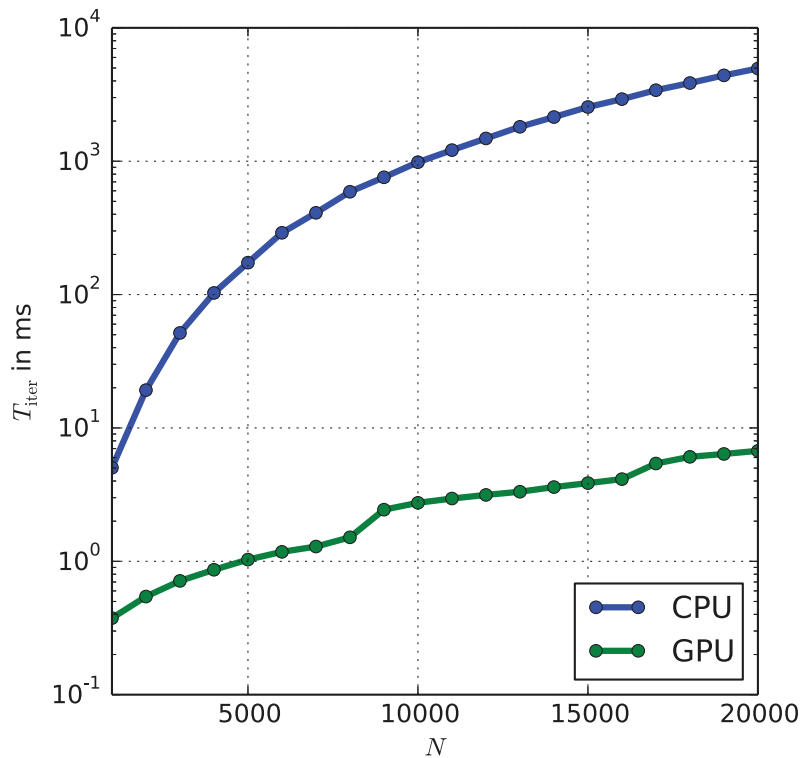
Figure 3.6: Speedup of the GPU implementation compared to the CPU implementation depending on the number of simulated cells N. The graph shows the time it takes to perform one iteration step. The increase in efficiency becomes larger for higher cell numbers. However, the number of cells is limited by the amount of memory build into the used GPU.

the grid, which was simplified to array sorting – which can easily be carried out in parallel – by a hashing technique, that maps the position of a cell to a unique grid index.

As a special feature of the simulation program a live visualization was incorporated, that allows for a live view of digital yeast cells flocculating. Despite slowing down the simulation considerably, this provides never seen (live) insights into the dynamics of flocculation and helps to adjust the parameter ranges.

The GPU implementation is much more time-efficient than the initial C/C++ prototypes. For $N = 20000$ cells the GPU algorithm is over 700 times faster than the parallel CPU version (see figure 3.6). This speedup enables me to simulate a reasonably large amount of cells to

prevent the previously described side effects of a limited cell number. Moreover, the GPU implementation only requires 25 ms to perform an iteration in simulations with up to 1.000.000 yeast cells.

3.3 An Experimental Approach to MEF-Cell Aggregation

In mammals cells have to differentiate according to their function in order to cooperate in forming and maintaining the higher organism. As Wolpert (1969) pointed out, a challenging problem is the differentiation according to spatial patterns at almost all stages of development.

Aggregation is one way of establishing spatial patterns such as islet-like aggregates that are able to differentiate later on (Hardikar et al., 2003; Chandra et al., 2011; Govindasamy et al., 2011). Therefore it is important to understand the regulatory and dynamical factors involved in cellular aggregation.

Martin and Evans (1975) established the STO mouse embryonic fibroblast (MEF) cell lines as the model system to study the biology of fibroblast cancer cells under laboratory conditions. The vast amount of tools to manipulate and work with these cell lines make them an ideal model system to study cellular aggregation. Another important aspect is the relatively slow movement of cells in the order of $20\,\mu\mathrm{m}/\mathrm{hr}$ (Pathak and Kumar, 2011) that makes the cells easily traceable through time-lapse microscopy.

The goal of the presented experiments is to quantify the spatial dynamics of this type of cellular aggregation and validate the model presented in section 3.2.

3.3.1 Experimental Setup

In order to analyze the dynamics of the cellular aggregation, the cells were tracked via automated time-lapse fluorescent microscopy. To increase the amount of data, the experiments were carried out with automation.

3.3.1.1 Cell Lines

The STO MEF cells were provided by Hu Jianjiang from the Department of Biochemistry of the Hong Kong University. The cell lines were stably transfected with enhanced green

fluorescent protein (eGFP). The eGFP protein labels the cell cytoplasm for our measurements of the cell shape, size and aggregate size and shape.

Additionally the cells were transfected with an H2B-mCherry fusion gene. This enables the cells to produce the fluorescent mCherry protein linked to histone H2B. Histone H2B is one of the chromatin components located in the cell nucleus. In this way, the H2B-mCherry fusion protein labels the cell nuclei which makes them traceable separately in fluorescent microscopy.

The used culture medium for MEF cells is Dulbecco's Modified Eagle Medium (DMEM) high glucose[6].

To evenly induce the cellular aggregation at all cell densities a conditioned medium has been used. This medium used for the MEF cell aggregation assays is the supernatant of a dense cell culture which has been transferred to another dish.

3.3.1.2 Time-Lapse Microscopy

In order to measure large samples of data in parallel, the people at Huang Wei's group at the Department of Biochemistry of the Hong Kong University developed an experimental setup using an automated time-lapse microscopy.

They used an auto-focus based on a laser-beam to focus the microscope to the bottom of a micro-titer plate. Since the thickness of the plates are roughly constant, the position of the object lens can be reset from there to focus the cells. The position of the plate has been set up using an *X-Y* table controlled by a computer program to sample the wells used for the experiment. Figure 3.7 shows a typical path for a 24 well micro-titer plate. If the program reaches the end of the path it is set to wait for a tunable time to ensure that the pictures of each sample are taken in equal length time steps.

The original image data was captured in 16 bit tiff format and consists of a differential interference contrast image, one fluorescent image from the mCherry channel to track the

---

[6]https://www.lifetechnologies.com/order/catalog/product/12100061?ICID=search-product
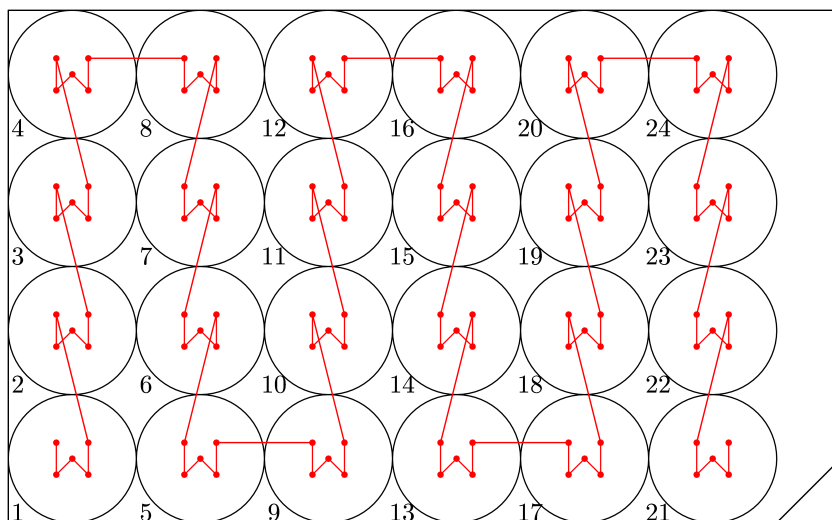
Figure 3.7: Typical path of a programmed time-lapse microscopy session. Five samples were taken per well. If the program reaches the end of the programmed path, it will wait for a defined time (might be zero) to ensure that the time shift between every image of the same sample stays constant.

cell nuclei, and one fluorescent image from the eGFP channel to track the cell positions and shapes. The two original fluorescent images were transformed to one bit images and analyzed using the tools introduced in section 3.1.

3.3.2 Experimental Results

Through time-lapse microscopy we were able to observe the dynamics of the cellular aggregation. A typical series of these images is shown in the top row of figure 3.8.

This figure indicates, that the primary aggregation process takes place within the first six hours of this experiment. After this primary aggregation the change in size and structure seems to become smaller.

To further quantify the aggregation process we converted the images into a binary format using the algorithm presented in section 3.1.1. We tuned the relative threshold (or the false-positive probability) for each series to account for the variable densities. This was done in a way that we achieve a linear dependency between measured initial cell density $\rho(t=0) = \rho_0$
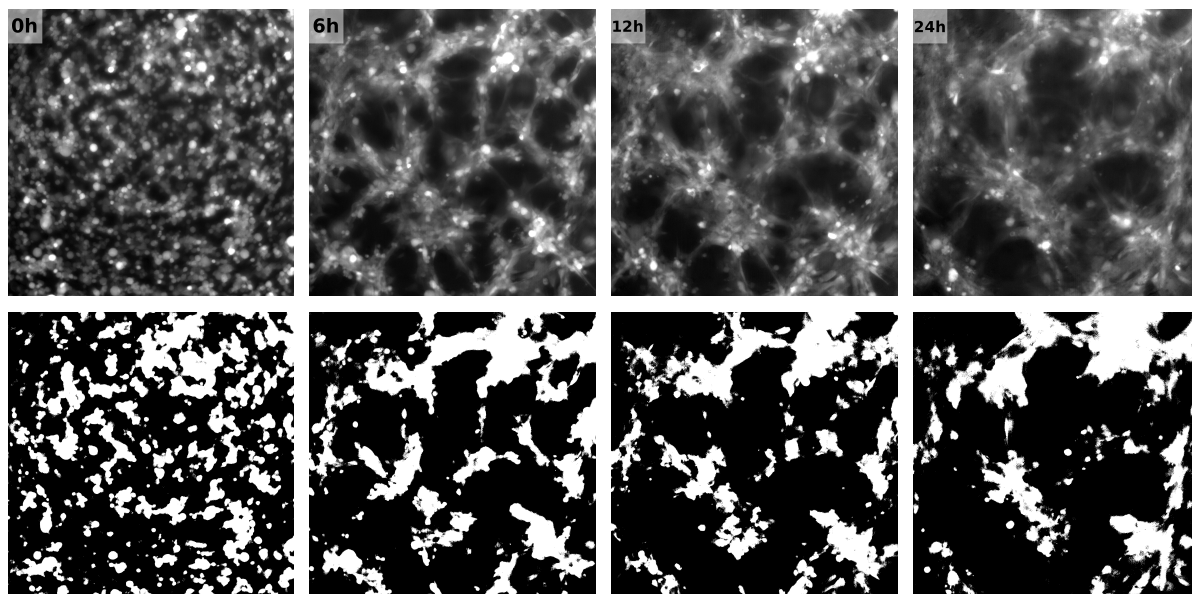
Figure 3.8: Time series of an aggregation experiment (top row) and corresponding binary images after conversion (bottom row). The experiment was done with an initial cell concentration of 160000 cells per well (160k). We used a false-positive probability of 0.2 to set the threshold for the binary conversion. In this time series one can clearly see the aggregation process is pretty fast in the beginning and then slows down at a certain aggregate size. The structures occurring in the first six hours seem to change only very slowly.

and the experimentally set initial cell density. An example result of this conversion is shown in the bottom row of figure 3.8. Even though we lose some image information through this conversion the images are able to represent the real states of the aggregates concerning size and structure of the respective original images.

We quantified and analyzed nine series with initial densities reaching from 10000 cells per well (10k) to 160000 cells per well (160k) using the measures introduced in section 3.1.

The analysis of the measured density $\rho(t) = \frac{\sum_{i=1}^{N_P} P_i}{N_P}$, where $P_i \in \{0, 1\}$ is the pixel value of the ith pixel and N_P is the total number of pixels in the image, shows that there are two phases of aggregation (see figure 3.9).

Right after the beginning of the experiment there is an increase in density indicating a spreading of the cells. For the lowest initial density this phase endures over the complete
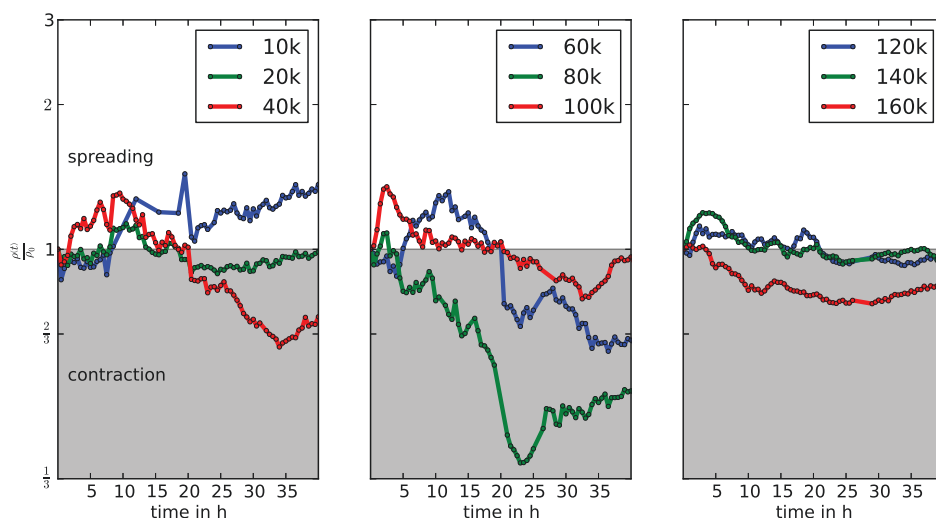
Figure 3.9: Measured density as a function of time for experiments with different initial cell concentrations. The curves are normalized to the measured initial density $\rho_0 = \rho\,(t=0)$. Leaving aside the 10k experiment, we can see two patterns in this graph. The cells from the 80k and 160k experiments start their contraction phase almost directly after starting the experiment, which might implicate that they already have been in an aggregated state before the experiment started. All other experiments show a transition between spreading and contraction after about 20 hours, independent of their initial cell concentration. This might imply a triggered mechanism that induces contraction a specified time after a certain threshold in the degree of aggregation is reached.

time the cells have been observed. Except for the experiments with initial concentrations of 80000 (80k) cells per well and 160000 (160k) cells per well all experiments show a transition from spreading to contraction after about 20 hours. The mentioned exceptions in the 80k and 160k experiments are likely caused by an aggregation taking place before the experiment was started.

The growth dynamics in figure 3.10 of the aggregates also show a difference for the three cases that differ in their density behavior (10k, 80k and 160k). In the experiments with 80000 and 160000 cells per well the average aggregate size – measured by the full width at half maximum r_{FWHM} of the pair correlation function – and the growth rate – defined as

77

$$k_g = \frac{r_{\text{FWHM}}\,(t = 40\,\text{h}) - r_{\text{FWHM}}\,(t = 0\,\text{h})}{40\,\text{h}}, \tag{3.37}$$

the linear growth over the experiment time of $40\,\text{h}$ – are far below expectations indicating that they reach their steady state earlier. This supports the argument of a premature aggregation before the experiment was started.
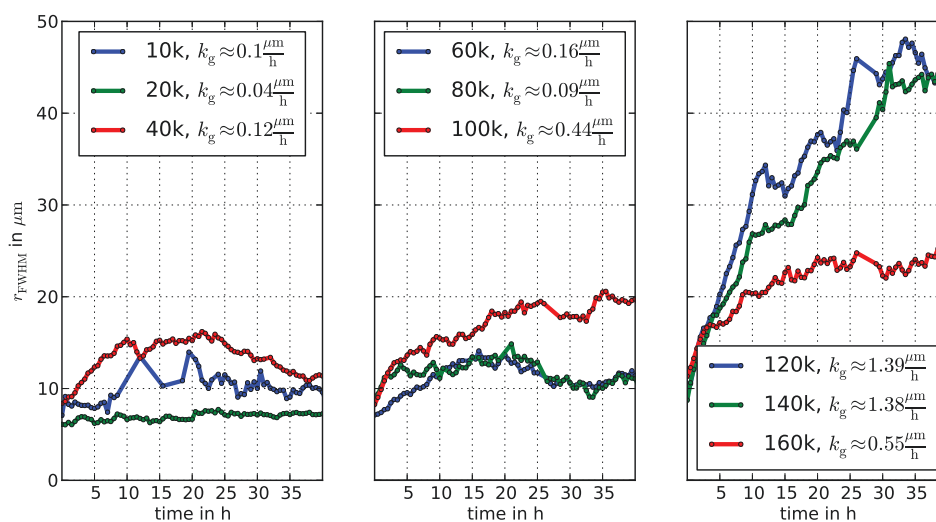
Figure 3.10: Time dependence of the aggregate size measured by pair correlation function. In general there is an increase in size r_{FWHM} and growth-rate k_{g} with increasing initial cell concentration. Like in figure 3.9 the experiments with 10k, 80k and 160k show a different behavior. While the difference in the 10k experiment can be explained by the relatively small number of aggregates (and therefore increased variance), the differences in the 80k and the 160k experiment support the argument of an aggregation taking place right before the experiment has been started.

In the other experiments we see a clear increase in aggregate size and growth-rate with increasing cell concentration indicating, that the aggregation is more efficient and faster for higher cell concentrations.

The time dependency of the fractal dimension shown in figure 3.11 indicates an increase in regularity, i.e. an approach to two dimensional objects, with time.
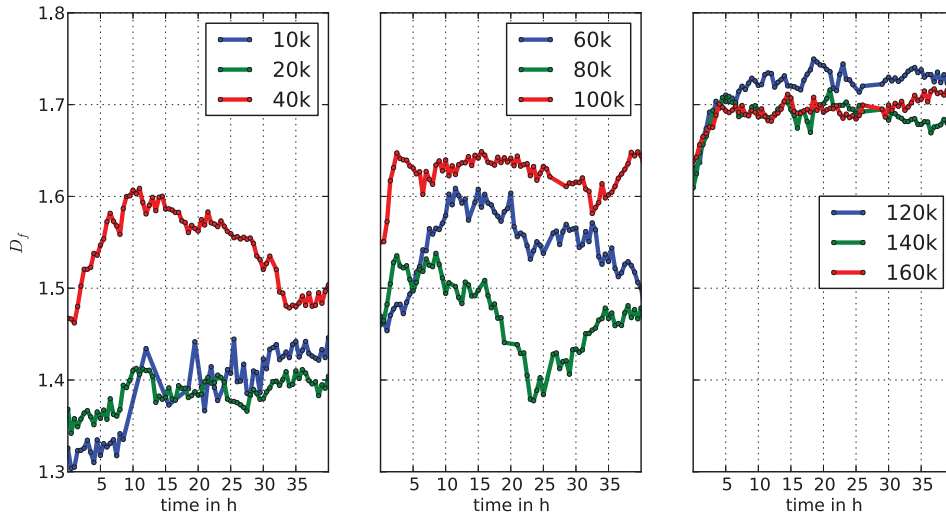
Figure 3.11: Fractal dimension of the experiments as a function of time. In most experiments the fractal dimension increases with time and increasing initial concentration until a saturation is reached. Again, the three experiments 10k, 80k and 160k differ from this observation. An increase in fractal dimension implies a more regular, i.e. more two dimensional structure. In general the aggregates get more and more regular with time (or with increasing aggregate size) until they reach a certain saturation.

This also reflects the size dependence of the structure, because the size of the aggregates simultaneously increases with time as shown in figure 3.10. Again and for the same reasons the only exceptions from this behavior are the experiments with 10k, 80k and 160k cells per well.

This simultaneous rise in fractal dimension and aggregate size indicates that the fractal dimension is related to the aggregate size in some way.

This relation is shown in figure 3.12. The fractal dimension increases with aggregate size before it reaches a saturation for $r_{\mathrm{FWHM}} \geq 20.33\,\mu\mathrm{m} = r_{\mathrm{trans.}}$. Interestingly, aggregates larger than this transition size have an average fractal dimension of $\langle D_f \rangle \approx 1.70$, which is very close to the fractal dimension resulting from a diffusion limited aggregation in two dimensions (Witten Jr. and Sander, 1981).
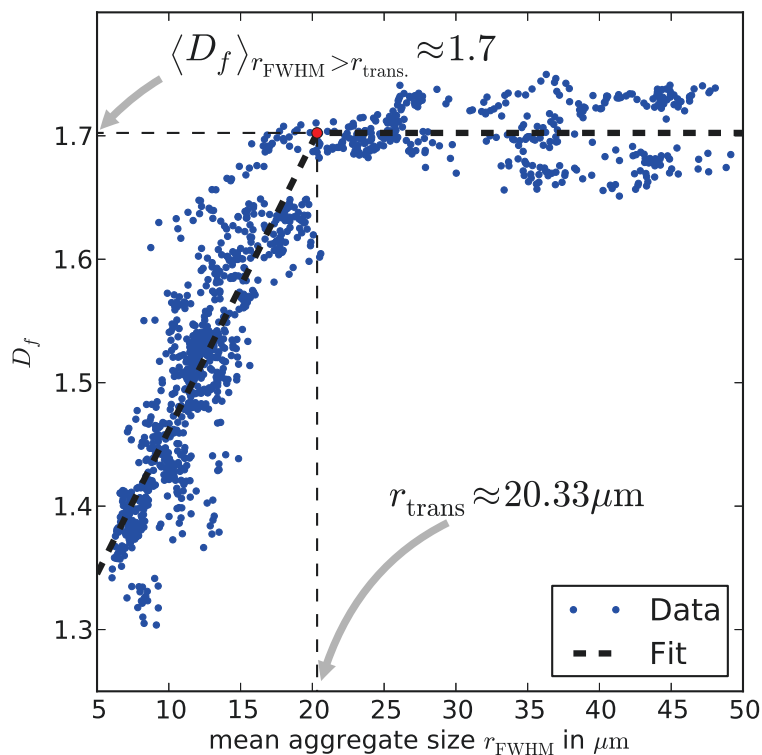
Figure 3.12: Fractal dimension D_f as a function of mean aggregate size r_{FWHM}. The fractal dimension of the aggregates linearly increases until a transition size $r_{\text{trans}} \approx 20.33\,\mu$m. Aggregates larger than this size show an average fractal dimension of $\langle D_f \rangle \approx 1.70$. This is strikingly close to the fractal dimension Witten Jr. and Sander (1981) predicted for the fractal dimension of a diffusion limited aggregation.

All in all, we can monitor the dynamics of the aggregation process using the measures presented in section 3.1. We can divide the aggregation into the two phases spreading – which roughly lasts 20 hours – and contraction as indicated in figure 3.9. We notice a simultaneous increase of fractal dimension and aggregate size over time in figures 3.10 and 3.11. This led to study the relation between aggregate size, measured using the pair correlation function, and fractal dimension. This relation shows a transition from a linear increase of D_f with aggregate size to a saturation at $\langle D_f \rangle \approx 1.70$. The close proximity to the fractal dimension of a two dimensional diffusion limited aggregation indicates that a random motion is the leading process

of this aggregation.

3.3.3 Simulation Results

The simulation results fit the experimental results qualitatively. However, since in this work only one experimental series was analyzed, there are some quantitative differences. The simulations were performed 100 times each and fitted to the fractal dimension resulting from the experimental time-lapse series of 20000 cells per well, 60000 cells per well, and 100000 cells per well using a randomized L-BFGS-B algorithm (Zhu et al., 1997) from the scipy library (Jones et al., 2001)[7]. Figure 3.13 shows a strong decrease in cluster growth after five to ten hours into the simulation.
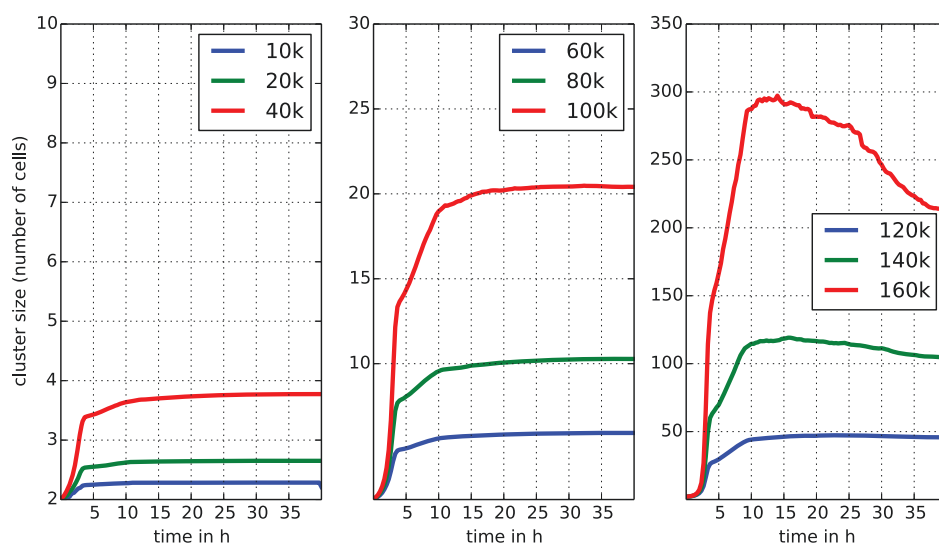
Figure 3.13: Simulated aggregate sizes as a function of time. The size strongly increases with time in the first hours. After about five to ten hours the growth slows down and the aggregate size reaches a saturation in most cases. Exceptions are the results for very high cell densities, where the aggregate size starts decreasing after some time. This is caused by the averaging over a decreasing number of clusters which increases the standard error of the mean accordingly.

---

[7]scipy.optimize.basinhopping, http://docs.scipy.org/doc/scipy-dev/reference/generated/scipy.optimize. basinhopping.html

This decrease finally leads to a saturation at a size correlated to the initial cell concentration where higher concentrations result in bigger aggregates. In contrast to the experimental results, the decrease in growth occurs roughly at the same time for all initial concentrations. This might be a hint, that the cells in the experiment were in different stages of the aggregation process when the experiment was started.

For very high cell densities there is a noticeable decrease in size after some time. Figure 3.14 indicates that this drop in size correlates to a relatively low number of clusters in the simulation.
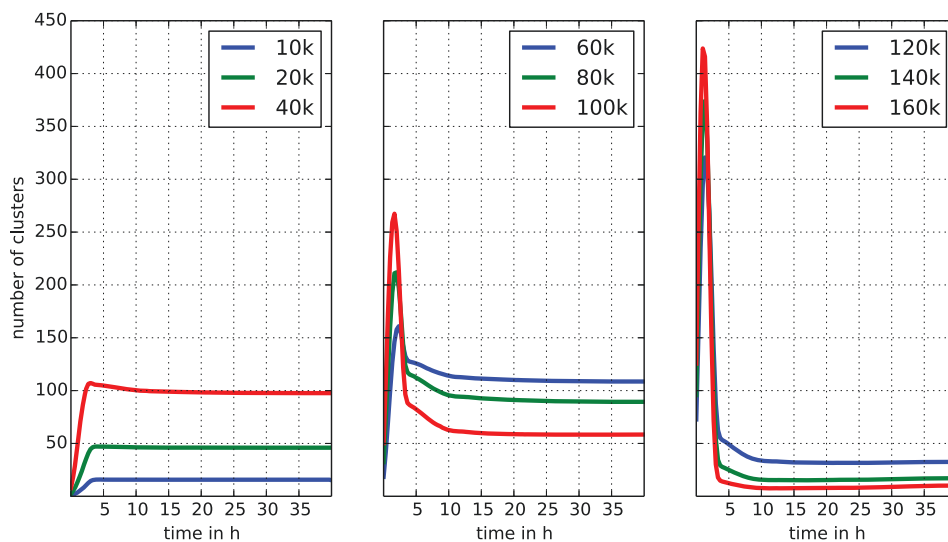
Figure 3.14: Number of clusters as a function of time for the simulations. After a strong increase the number of clusters decreases as the clusters merge to fewer, bigger clusters.

This decreasing number of clusters is caused by merger of clusters in the simulated area. Since no cells are allowed to get in or to leave this area, the standard error of the mean increases with decreasing number of clusters, leading to side effects like one huge cluster surrounded by mini clusters of two to five cells.

The fractal dimensions of the simulations show a very similar behavior as the growth of

the clusters like it does in the experimental series. Figure 3.15 shows that after a strong initial increase the fractal dimension increases slower and slower until it eventually reaches a saturation.
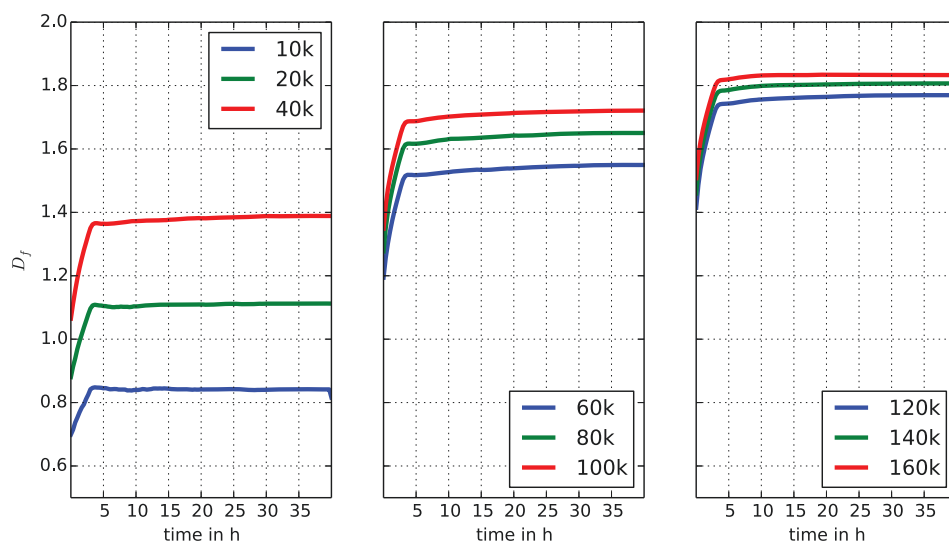
Figure 3.15: Fractal dimension of the simulated aggregates as a function of time. The fractal dimension increases with time and increasing initial concentration until a saturation is reached. If we compare the simulation results to the fractal dimension of the experimental sequence we see a good qualitative agreement. However, the fractal dimension seems to be too small for low cell densities.

Again we see a very nice qualitative agreement. In case of very low initial cell densities, the simulation seems to consequently underestimate the fractal dimension. This can be caused by the simplification to circular cells because the cell shape directly influences the spatial organization of cells. At high densities the fractal dimension fits the experimental data quantitatively, too. However, this simulation result also hints that the cells had to be in a different stage of the aggregation process at the beginning of the experiment.

Because of the noticeably similar course of growth and fractal dimension it makes good scene to look for a correlation of those two like the one shown in figure 3.12. Indeed, we see a similar correlation except that it looks a bit smoother and the saturation occurs at a higher

fractal dimension. Figure 3.16 shows the correlation.
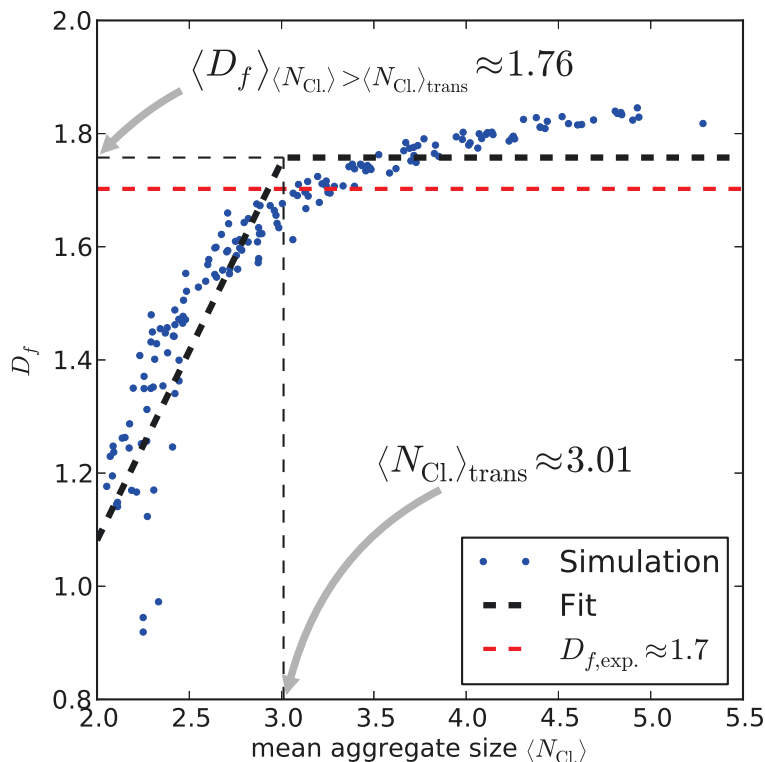
Figure 3.16: Fractal dimension D_f as a function of mean aggregate size $\langle N_{\text{Cl.}} \rangle$ from the simulations. As in the experiments the fractal dimension increases linearly until a transition size is reached ($\langle N_{\text{Cl}} \rangle \approx 3.01$). After this a saturation starts. If the simulation data is fit with the same bilinear curve as the experimental data, it yields a similar saturation value as the experimental data ($D_f \approx 1.76$ compared to $D_f \approx 1.70$ in the experiments).

The saturation value is $D_f = 1.76$ (compared to $D_f = 1.70$ in the experiment) and is reached at a cluster size of about $\langle N_{\text{Cl.}} \rangle = 3.01 \, \text{cells}/\text{cluster}$. This value seems to be too small if compared to the experimental results. If we assume $3.01^{\frac{1}{3}} \approx 1.44$ cells in each direction having a length of about $12 \, \mu\text{m}$ we have an approximate transition cluster size of $r_{\text{trans.}} \approx 17.33 \, \mu\text{m}$ compared to $r_{\text{trans.}} \approx 20.33 \, \mu\text{m}$ in the experiments. However, it is not possible to calculate the cluster size using the pair correlation function in a comparable way to the experiments, because the cells

are represented by circles not by their real shape. This makes it difficult to estimate cluster sizes and to compare them with experimental data.

In addition to the quantities that could be analyzed experimentally, I studied the degree of aggregation, i.e. the fraction of cells that are bound in clusters, and the clustering coefficient, i.e. the average number of bonds per cells divided by the total number of cells.

Figure 3.17 shows the aggregation degree which behaves very similar as does the fractal dimension. It first increases rapidly to start saturation after about five hours into the simulation.
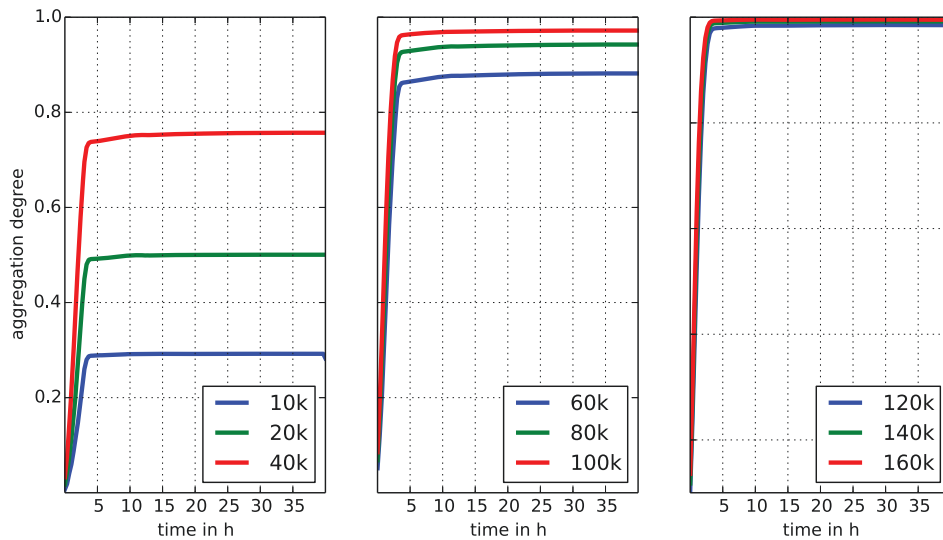
Figure 3.17: Aggregation degree of the simulations as a function of time. The aggregation degree strongly increases with time until it saturates at a value corresponding to the initial cell density. This was expected as the cells at lower cell densities have a much lower probability to encounter other cells and aggregate as cells at high cell densities.

The saturation value reached before the differentiation of the cells, i.e. the transformation into non-motile cells, strongly depends on the initial cell concentration. The higher the initial concentration, the more cells aggregate into clusters.

The cluster coefficient shown in figure 3.18 nicely increases at first and then starts a saturation. For all initial cell concentrations the coefficient increases for about the same time

to about the same value, before the lines split and reach a saturation value depending on the initial cell concentrations.
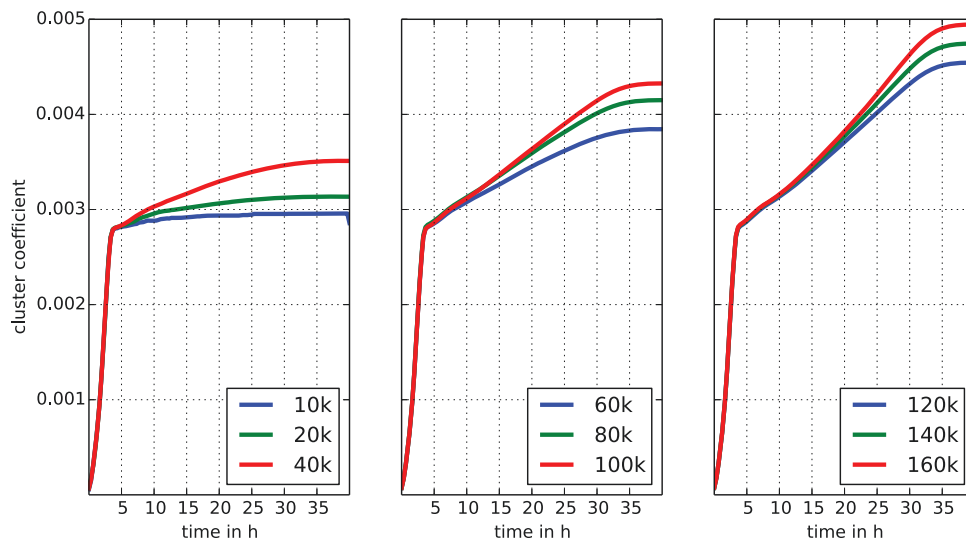
Figure 3.18: Cluster coefficient of the simulations as a function of time. The cluster coefficient sharply increases in the first hours of the simulation. After all initial clusters were formed and the growth of clusters is governed by merging clusters instead of acquiring single cells the cluster coefficient increases less rapidly. It reaches a saturation value that is higher for increasing cell densities.

These splitting points mark the transition from a growth mode that is governed by acquiring single cells to a growth mode that is governed by the merger of whole clusters.

3.3.4 Discussion and Summary

The new experimental approach in combination with the methods to quantify patterns led to new insights into the dynamics and structure of MEF cell aggregation using time-lapse microscopy. The analysis of the microscopy series showed that there are two phases of aggregation. First, the cells spread, i.e. the substrate area occupied by cells increases. After about 20 hours the cells start contracting, i.e. the area covered by cells decreases.

Using the pair-correlation function to extract a measure for the aggregate sizes, it was possible to estimate the growth rate that depends on the initial cell concentration. For very low cell concentrations (10000 - 40000 cells per well) the growth rate is in the order of $0.1\,\mu\mathrm{m/h}$. For the higher concentrations (120000-160000 cells per well) the rate increases more than tenfold to $1.4\,\mu\mathrm{m/h}$. This shows that, as expected, the aggregation is much faster at high cell concentrations.

The fractal dimension is a measure for the regularity and density of the cluster. The more regular the clusters become, the more two dimensional (i.e. area-like) the patterns will be. After a stiff increase, the fractal dimension reaches a saturation value depending on the initial cell concentration (figure 3.11). The reason for this is shown in figure 3.12: the spatial structure of the aggregates (as measured by the fractal dimension) depends on their size. The bigger the aggregates become, the higher gets the fractal dimension meaning that the structure of larger aggregates is more regular than the structure of smaller aggregates. At a size of about $20.33\,\mu\mathrm{m}$ the dependency shifts from linear to a constant fractal dimension of 1.7. This is very similar to the patterns resulting from diffusion limited aggregation as reported by Witten Jr. and Sander (1981). While not proving diffusion limited aggregation as the mechanism at play here, it indicates that the random motion of the cells, i.e. the "diffusion", might be an important factor in this process.

However, the model introduced at the beginning of this section also reproduces the experimental findings quite well without explicitly simulating a random motion. Fitting the behavior of the fractal dimension to the experimental data leads to a good qualitative agreement. In the simulations we could also identify two phases. In the first hours of the simulation the aggregates grow by acquiring single cells. After some time this shifts to growth by merger of larger aggregates.

The aggregate sizes were measured in average number of cells per aggregate. While showing a very similar behavior as the aggregate sizes in the experiments it is difficult to compare

these finding quantitatively. The simulation neglects cell shape that is very important for the calculation of the pair-correlation function because it changes the probability where cells can be found in a certain distance since the real cells cover a larger area compared to the circular ones in the simulation.

Figure 3.16 shows that in the simulation the structure of the aggregates also shifts from an increasing to a constant fractal dimension as a function of aggregate size. The transition is not as sharp as it is in the experiments, but it reaches a similar saturation value (1.76 compared to 1.70).

In addition to the quantities that can be measured experimentally, the simulation allows for the analysis of more quantities. The aggregation degree sharply increases with time and saturates at a level that depends on the initial cell concentrations. In simulations with higher initial concentrations a higher fraction of cells is found in aggregates. Indicating, that the aggregation is not only faster but more effective in terms of cells involved. The clustering coefficient supports the definition of the two aggregation phases in the simulation as growth by acquisition of single cells and growth by merger of clusters. The coefficient strongly increases in a concentration independent way at first. After some hours into the simulation it splits into concentration dependent curves, indicating the shift between the two aggregation phases.

All in all this section showed the importance of quantitative analysis of patterns. Especially the sharp transition in figure 3.16 was found by carefully quantifying the microscopy images. To further quantify the two aggregation phases (spreading and contraction) one has to analyze more experiments for the initial conditions to avoid errors occurring from non isotropic cell concentrations inside one well. For the simulations it would be interesting to incorporate cell shape and random motion. While the latter would be quite easy to implement the first is challenging because the cell shape is highly irregular and it changes randomly with time and through mechanical forces that are applied to the cell.

However, the model so far enabled us to study the basic principle of cellular aggregation

and made it possible to analyze parameters that are not accessible experimentally.

3.4 A New Experimental Approach to Yeast Flocculation

S. cerevisiae, besides its biotechnological relevance, has been used as a model organism to study eukaryotic molecular and cell biology for a long time. Its well known genome and the tremendous amount of experimental tools available for this organism makes it a very good choice as a model.

Flocculation are one of the versatile lifestyles of *S. cerevisiae* in liquid media as reviewed by Brückner and Mösch (2012). In biotechnology this lifestyle is exploited as a cheap and easy way to remove the production cells from the desired medium after the fermentation process.

Smukalla et al. (2008) presented evidence that *S. cerevisiae* uses flocculation as a cooperative mechanism for protection from environmental harms by creating a diffusion barrier that cuts off the inner cells from external chemical compounds. Since this behavior is limited to cells expressing the flocculin, the Flo-gene family contains some candidates that qualify as "green beard genes" as defined by Dawkins (1976).

The work of Smukalla et al. (2008) also showed that, in some cases, cells not expressing the binding protein (cheater cells) were incorporated into the flocs but are enriched at the flocs surface. This makes them the first line of defense that suffers from the harsh surroundings without the benefits of the protective environment inside the flocs. However, the flocs they used to measure the diffusion barrier effect and the enrichment of the cheater cells were rather big and dense.

In this chapter I will present a new experimental approach to quantify the structure of the flocs including the position of the cheater cells. Together with Dr. Stefan Brückner and Prof. Dr. Hans-Ulrich Mösch I developed an experimental setup to generate three dimensional image data that can be analyzed according to the methods shown in section 3.1 and compared to the model presented in section 3.2.

3.4.1 Experimental Setup

In order to analyze the three dimensional structure of the flocs, including the position of the cheater cells, we developed an experimental setup that uses confocal laser scanning microscopy to generate three dimensional image data of them.

We constructed two lab strains, one expressing the Flo5-protein and green fluorescent protein (GFP), the other lacking the *FLO*5 gene and expressing red fluorescent protein (RFP).

In order to ensure a constant copy number of *FLO*5 throughout all cells, we decided to integrate *PGK*1(P)-*FLO*5-*FLO*11(T) into the chromosome of our basic *S. cerevisiae* strain.

3.4.1.1 Plasmids

In order to achieve a chromosomal integration of the *FLO*5 gene we had to construct a plasmid carrying the *PGK*1(P)-*FLO*5-*FLO*11(T) sequence plus an integrative backbone. This was done by cutting out the *PGK*1(P)-*FLO*5-*FLO*11(T) sequence from BHUM1600 using the restriction enzymes EcoRI, EcoRV, and HindIII. The resulting linear DNA strand was ligated to the linearized version of pRS306.

To label the cells according to their cell type we used the BHUM2028 (RFP) plasmid for cheater cells and the BHUM2029 (GFP) plasmid for bearer cells.

All used plasmids (including their references) are listed in table 3.4.

3.4.1.2 *S. cerevisiae* Strains

The strains used in this study are based on the FY1679-14C yeast strain introduced by Winston et al. (1995). To construct the bearer strain we used the BHUM2195 plasmid to integrate a working *FLO*5 gene into FY1679-14C and labeled it with GFP using the BHUM2029 plasmid.

For the cheater strain we used the FY1679-14C strain and transferred the BHUM2028 plasmid carrying RFP into the strain. The Resulting strains are catagolized as YHUM2146

Plasmid	Description	Reference
BHUM1600	*PGK*1(P)-*FLO5-FLO*11(T), *URA*3, AmpR, CEN	Collection AG Mösch
BHUM2028	*TDH*3(P)-tagRFP-*CYC*1(T), *TRP*1, AmpR, CEN	Collection AG Mösch
BHUM2029	*TDH*3(P)-yEGFP-*CYC*1(T), *TRP*1, AmpR, CEN	Collection AG Mösch
BHUM2195	*PGK*1(P)-*FLO5-FLO*11(T), *URA*3, AmpR	This work
pRS306	*URA*3, AmpR	Sikorski and Hieter (1989)

Table 3.4: Plasmids used for the flocculation experiments. BHUM1600 is the source for the *FLO*5 gene, BHUM2028 and BHUM2029 carry the RFP and GFP labels, pRS306 is the integrative shuttle vector, and BHUM2195 is the plasmid carrying *FLO*5 for integration into the chromosome of *S. cerevisiae*.

(bearer) and YHUM2148 (cheater) in the collection of AG Mösch (see table 3.5 for strain descriptions and references).

Strain	Description	Reference
FY1679-14C	S288c MATa *ura*3-52 *trp*1Δ63	Winston et al. (1995)
YHUM2146	S288c MATa *PGK*1(P)-*FLO5-FLO*11(T)::*URA*3 *trp*1Δ63	This work
YHUM2148	S288c MATa *URA*3 *trp*1Δ63	This work

Table 3.5: Used strains of *S. cerevisiae*. FY1679-14C is the basis in which we integrated *FLO*5 and GFP to obtain the bearer strain YHUM2146. We also transferred RFP to an unmodified version of this strain to obtain the cheater strain YHUM2148

3.4.1.3 Flocculation Essays

We used overnight cultures of bearer and cheater cells that are treated with EDTA to dissolve the flocs and diluted the single cells to set a desired OD_{600} value. EDTA binds Ca^{2+} so that the calcium dependent binding of *flo5* is annihilated.

The single cells are diluted to different OD_{600} values. After adding 25 mM Ca_2 to counter the effects of EDTA, the flocs are allowed to settle for one minute. An amount of $100\,\mu$l of the supernatant is mixed with $900\,\mu$l SC-4 medium and the OD_{600} value is measured.

For the mixing experiments the cell cultures were diluted to $OD_{600} = 4$. Then we mixed different fractions of cheater and bearer cells to a total volume of 1 ml including 25 mM $CaCl_2$. After one minute of settling $100\,\mu l$ of the supernatant are diluted in $900\,\mu l$ medium and the OD_{600} is measured.

3.4.1.4 Confocal Laser Scanning Microscopy (CLSM)

We treated overnight cultures of YHUM2146 and YHUM2148 strains with 25 mM ethylenediaminetetraacetic acid (EDTA) to dissolve the flocs into single cells. Then cultures were diluted to set $OD_{600} = 1.5$.

To study the structure of bearer flocs we diluted the overnight culture of bearer cells further to set different OD-values. To study the structure of flocs consisting of bearer and cheater cells we mixed the two cultures in the respective bearer-cheater ratios.

After adding $300\,\mu l$ low melting agarose to $300\,\mu l$ of the prepared culture, we induced the flocculation by adding 30 mM $CaCl_2$ that counters the effects of the EDTA.

We placed $100\,\mu l$ of the culture-agarose mixture on a microscope slide between two spacers consisting of three cover slips (see figure 3.19), covered it, and placed it on a cooled aluminum block to speed up the coagulation of the agarose and protect the flocs.
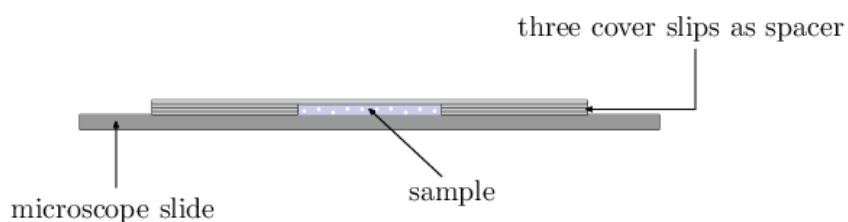
Figure 3.19: Sample preparation for CLSM. on both sides three cover slips act as spacers to prevent the flocs from being squashed. The agarose keeps them at a fixed position so that it is possible to perform three dimensional CLSM evaluations.

Confocal laser scanning microscopy allows analyzing objects in all three spatial dimensions (Roderfeld et al., 2003). With this method one obtains a three dimensional representation of

the cells inside a flock via a stack of optical slices. These slices are put together virtually into a three dimensional matrix and the spatial fluorescence distributions are analyzed using the methods described in section 3.1.

To quantify the density dependence of the structure we varied the OD_{600} value of the cells in the low melting agarose in steps of 0.15 from $OD_{600} = 0.15$ to $OD_{600} = 0.75$ and scanned four flocs for each value.

To quantify the influence and the spatial position of cheater cells, we varied the fraction of cheater cells in steps of 0.1 from $\frac{n_{\text{cheat.}}}{n_{\text{tot}}} = 0.1$ to $\frac{n_{\text{cheat.}}}{n_{\text{tot}}} = 0.9$ and again scanned four flocs for each value.

3.4.2 Experimental Results

3.4.2.1 Flocculation Essays

The flocculation essays show that the flocculation takes place as expected. The efficiency breaks down below a certain concentration of cells (see figure 3.20) because of the decreasing probability of cell-cell encounters with decreasing cell concentration. If we increase the amount of cheater cells, the flocculation efficiency decreases as indicated in figure 3.21.

Figure 3.22 shows the cheater concentration in the supernatant as a function of the cheater concentration in the mixed culture. The fraction of cheater cells in the supernatant is significantly higher than the fraction of cheater cells used in the mix. This means that bearer cells have a significant advantage over the cheater cells to get inside the flocs.

3.4.2.2 Three Dimensional Structure of Flocs

To get a first impression of the three dimensional structure of the flocs we generated a three dimensional projection of a floc using a slice distance of $\Delta z = 0.08\,\mu\text{m}$ and a horizontal res-
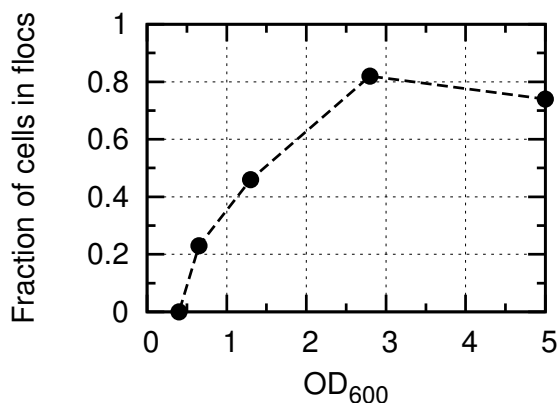
Figure 3.20: Flocculation efficiencies for different OD_{600} for single cells. There seems to be a minimum concentration for flocculation to work and a maximum efficiency at about 80 %.
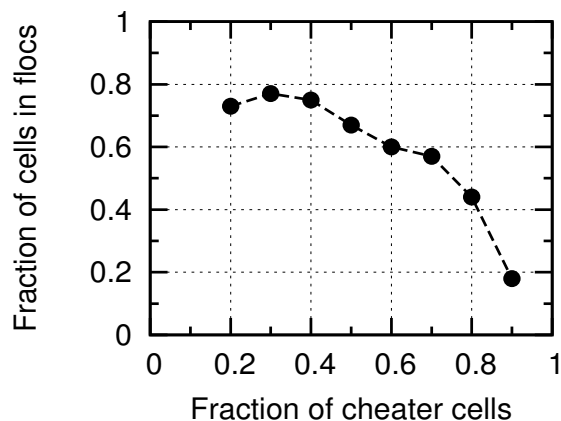
Figure 3.21: Flocculation efficiencies for different fractions of cheater cells in the solution. This data clearly shows that the presence of cheater cells clearly lowers the flocculation efficiency.
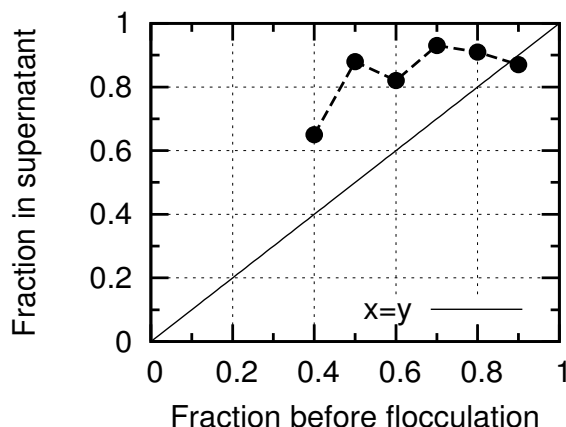
Figure 3.22: The fraction of cheater cells in the supernatant after flocculation is higher than the one used before flocculation. This clearly shows that cells expressing the *Flo*-gene are accumulating inside the flocs.

olution of $0.13\,\mu m/px$. The upper images of figure 3.23 show excerpt of the resulting stack. It nicely shows the positions of bearer cells in green and cheater cells in red.

Using the 3D projection tool of the software ImageJ[8] one can easily generate a rotatable three dimensional projection from this stack of images. The lower images of figure 3.23 show

---

[8]http://imagej.nih.gov/ij/

the resulting projection for a rotation of 150°.
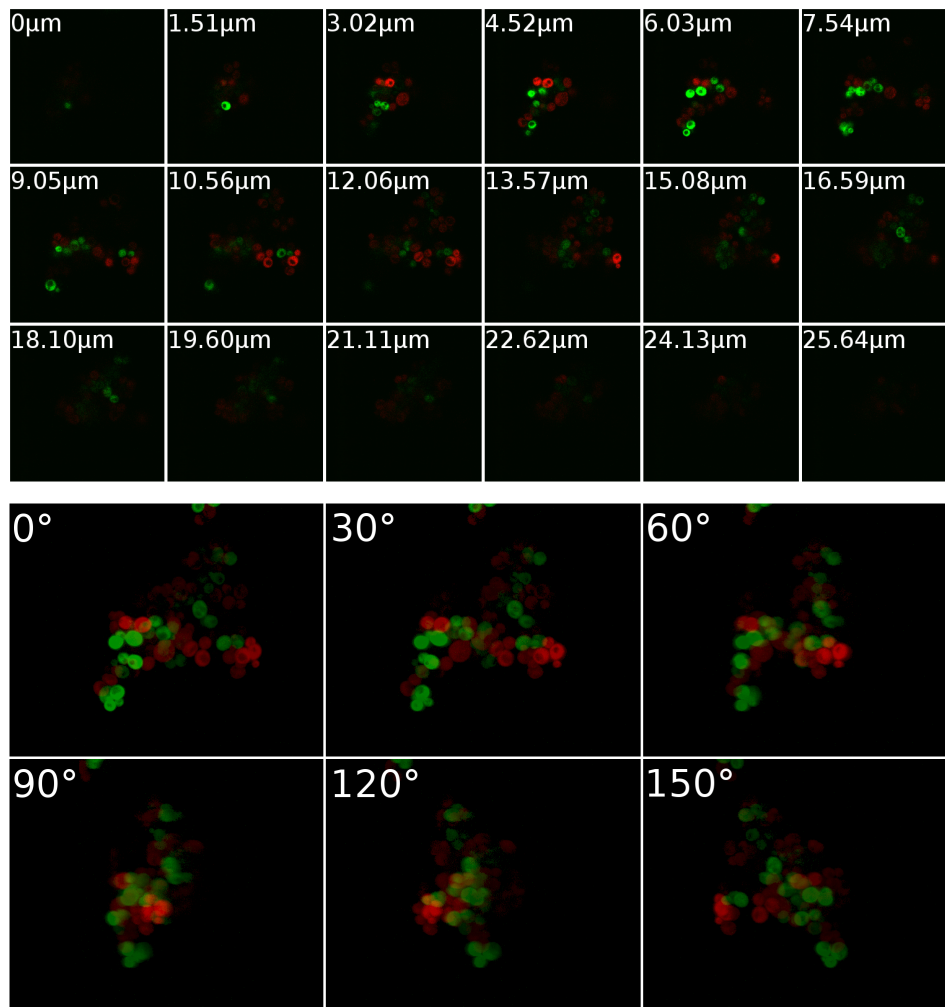
Figure 3.23: Example series of CLSM images for a floc from a mixing experiment with 20% bearer cells and 80% cheater cells (upper series). The slices in this series were taken every $0.08\,\mu$m. This figure shows every 9th slice of 324 total slices. Series like this were converted into binary three dimensional data and evaluated with the methods described in section 3.1. The lower images show the three dimensional projection of the series rotated around the *y*-axis. Using the 3D-Projection tool from ImageJ one can easily generate rotatable three dimensional representations of the flocs. This is very useful to get a feeling for the three dimensional structure of the flocs.

Even though one cannot make any well-founded quantitative statements on the structure of

the flocs from these three dimensional projections it is possible to see some qualitative details.

The most obvious of these details is the irregular structure of the flocs which makes it difficult to define a surface. So instead of discussing the composition of the surface of a floc it is better to look at the composition of the exposed cells.

Another important detail one can directly see from these images is that the bearer cells would often build a flock on their own. This means that cheater cells are not found in positions where the structural integrity of the flocs depends solely on them.

As can be expected from the flocculation assays, the cell concentration has a huge impact on the flocculation rate. However, it also directly influences the floc size. Figure 3.24 shows that the floc size measured through the full width at half maximum (FWHM) of the pair-correlation function $g(r)$) (see section 3.1) is strongly increasing for higher concentrations.
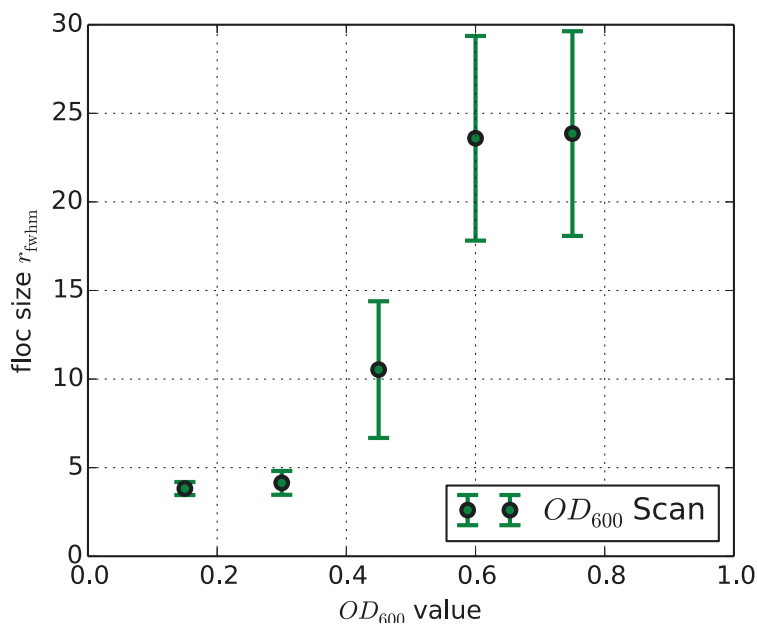
Figure 3.24: The floc size strongly increases with increasing concentration. There is a strong positive correlation of $r_{xy} = 0.73$ between the OD_{600} value (as a measure for concentration) and the FWHM of $g(r)$ (as a measure for the floc size).

A good measure for the impact of a parameter x on another parameter y is the correlation coefficient

$$r_{xy} = \frac{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \langle x_i \rangle_i)(y_i - \langle y_i \rangle_i)}{\sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \langle x_i \rangle_i)^2} \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (y_i - \langle y_i \rangle_i)^2}} \tag{3.38}$$

that is defined for two sample series x_i and y_i, where $\langle \rangle_i$ denotes the mean over all measurements i.

In this case the OD_{600} value and the floc size are strongly, positively correlated with $r_{xy} = 0.73$ indicating the huge impact of concentration not only on the flocculation efficiency but also on the protective potential of each single flock.

The independence of flocsize of the pair correlation function $g(r)$ and boxcount dimension shown in figure 3.25**A** is a first hint that flocs are fractal, self similar objects. The correlation coefficient for flocsize and boxcount dimension is $r_{xy} = 0.31$ indicating that there is little to no correlation.

This is also supported by the fact, that the set cheater concentration, which as I showed above reduces the flocculation efficiency, is also not correlated to the boxcount dimension as shown in figure 3.25**B** ($r_{xy} = 0.40$).

Finally, the set cell concentration (represented by the OD_{600} value) and the boxcount dimension shown in figure 3.25**C** are also uncorrelated with a correlation coefficient of $r_{xy} = 0.43$.

These results justify the assumption of a universal, self similar structure of the flocs. The mean boxcount dimension is $\langle D_f \rangle = 2.64$. Assuming a smallest possible structure that builds the flock, the next logical step is to ask for the size of this "unit floc".

Because the number of cells is an integer and the cell size is roughly the same for every cell, there is only a limited amount of flocs that can have a fractal dimension $D_f = 2.64$. To find those flocs, I wrote a monte-carlo algorithm, that builds flocs on a grid from a central seed and adds single cells if D_f gets closer to the desired value. If D_f gets worse with addition of a single cell this addition will be accepted wit a probability of 5 % to avoid getting caught in
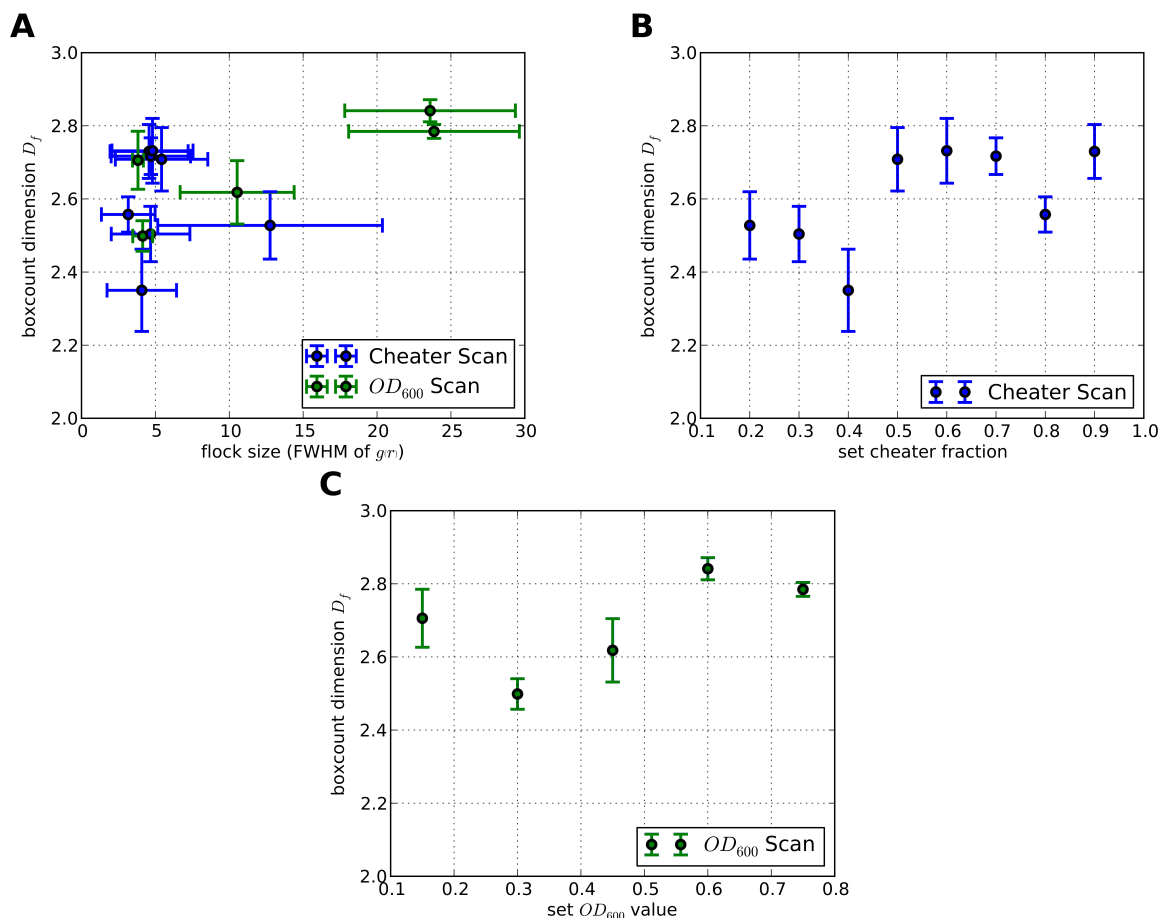
Figure 3.25: **A**: Size of flocks vs. boxdimension. There is little to no correlation. $r_{xy} = 0.31$ $\langle D_f \rangle = 2.64$, **B**: Boxdimension is independent of the set cheater concentration. $r_{xy} = 0.40$ $\langle D_f \rangle = 2.60$, **C**: Boxdimension is independent of the set OD600. This supports the argument of figure 3. $r_{xy} = 0.43$ $\langle D_f \rangle = 2.70$

local minima. After adding 200 cells the algorithm stops and the flocsize where the best fit was encountered is saved and analyzed for a large number of runs.

The distribution of these best fit values as a function of flocsizes is shown in figure 3.26. It has two clear peaks, one at $n_1 = 39$ and one at $n_2 = 78$ cells that persist if we apply the condition that the difference of the fractal dimension to the mean value of the measured fractal dimensions is smaller than twice the standard error of that mean value. From this it can be concluded that the smallest floc having a fractal dimension of 2.64 within a 2σ certainty interval consists of 39 cells while the peak at 78 cells can be interpreted as a floc consisting of
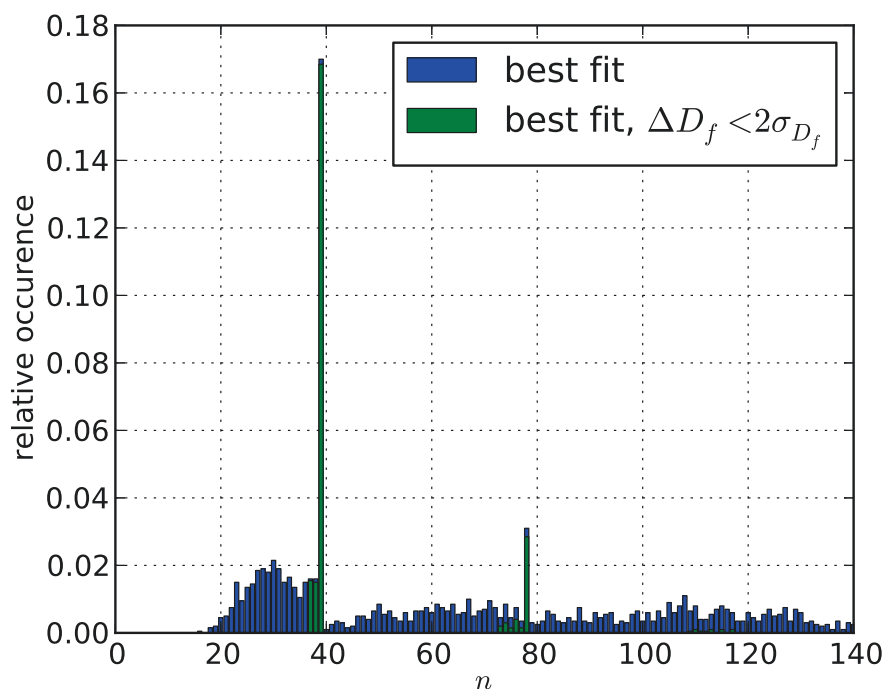
Figure 3.26: Distribution of best fit values encountered by the monte-carlo algorithm. The distribution has two noticeable peaks at $n_1 = 39$ and $n_2 = 78$ that persist if we apply the condition, that the difference in fractal dimension to the mean value of the measured flocks is less than two times the standard error of that mean. Since 78 is twice 39 it is possible to conclude that the unit floc consists of 39 single cells. The second peak can be explained as a combination of two unit flocs.

two 39 cell unit flocs.

To further quantify theses findings, I analyzed the size of the unit flock given by the algorithm as a function of the target fractal dimension (see figure 3.27).

As it turns out, there is an exponential dependency between fractal dimension and unit floc:

$$N_{\text{unit.}}\left(D_f\right) \approx \exp\left(1.39 D_f\right).$$

(3.39)

This is consistent with the assumption that a zero-dimensional object is a single point.
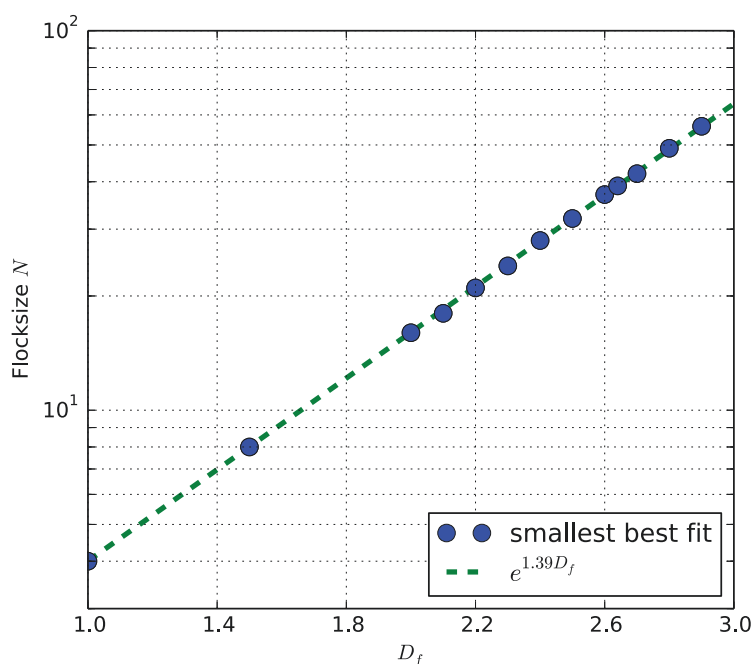
Figure 3.27: Resulting sizes of the unit flocs N for varying fractal dimensions D_f. The resulting sizes are all located on an exponential function $N = \exp\left(1.39D_f\right)$. Though the source of this correlation remains unknown, this indicates, that the minimal unit floc for $D_f = 2.64$ is indeed 39 sharp. Based on the standard error of the mean of the measured D_f of about 10%, the experimental unit flocs contain 39 ± 14.5 cells.

3.4.2.3 Role of the Cheater Cells

The flocculation essays already showed the inhibiting effect of cheater cells in the medium. Figure 3.28 indicates that the presence of cheater cells also decreases the flocsize. The correlation coefficient is $r_{xy} = -0.55$.

There are two possible ways cheater cells could influence the floc size. The first way is an effective decrease in bearer concentration and in this way reducing the collision probability of bearer cells and flocs or other cells. The second way one can think of is the destabilization of flocs with an amount of built in cheaters above a certain threshold. This destabilization is caused by the lack of binding protein in cheater cells resulting in the possibility of "blocking"
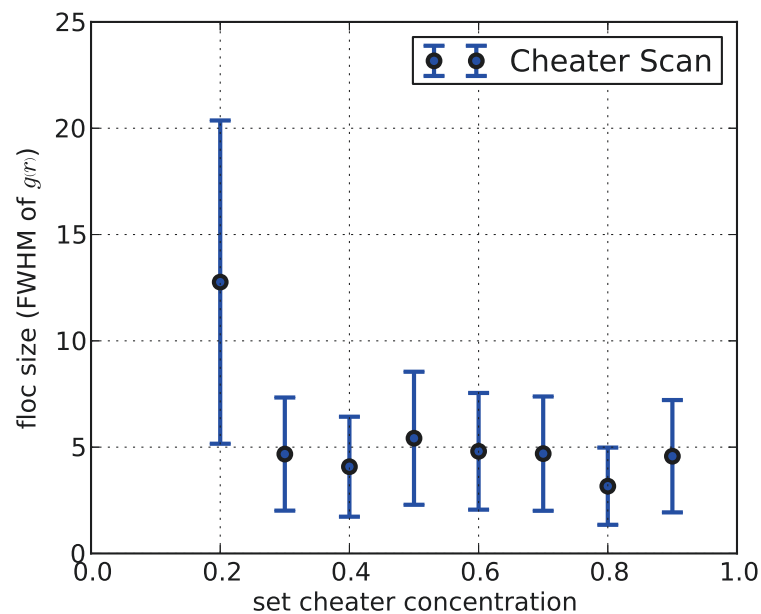
Figure 3.28: The floc size is strongly decreases with increasing initial cheater concentration. The FWHM of $g(r)$ shows a decreasing trend for increasing cheater concentration with a correlation coefficient of $r_{xy} = -0.55$.

bearer cells and weakening the structural integrity of the floc.

Looking at the floc size as a function of the measured cheater concentration reveals that the first is not correlated with a correlation coefficient coefficient of $r_{xy} = -0.14$. This indicates that the cheaters inhibit flocculation mostly by reducing the effective concentration of bearer cells.

The remaining question concerns the position of the cheater cells inside the flock. To answer this question I analyzed the influence of the cheater concentration on the structure of the floc represented by the fractal dimension.

Figure 3.29**A** and figure 3.29**B** show the fractal dimension of the green channel, i.e. the structure of the floc if it is stripped of all cheater cells, as a function of set cheater concentration in the initial mixture and the measured cheater concentration of the floc respectively.

These figures indicate an independence of the floc structure from the amount of cheater

cells over a huge range of cheater concentrations. The correlations coefficients are $r_{xy} = -0.35$ in case of the measured cheater concentration and $r_{xy} = -0.32$ in case of the set cheater concentration.
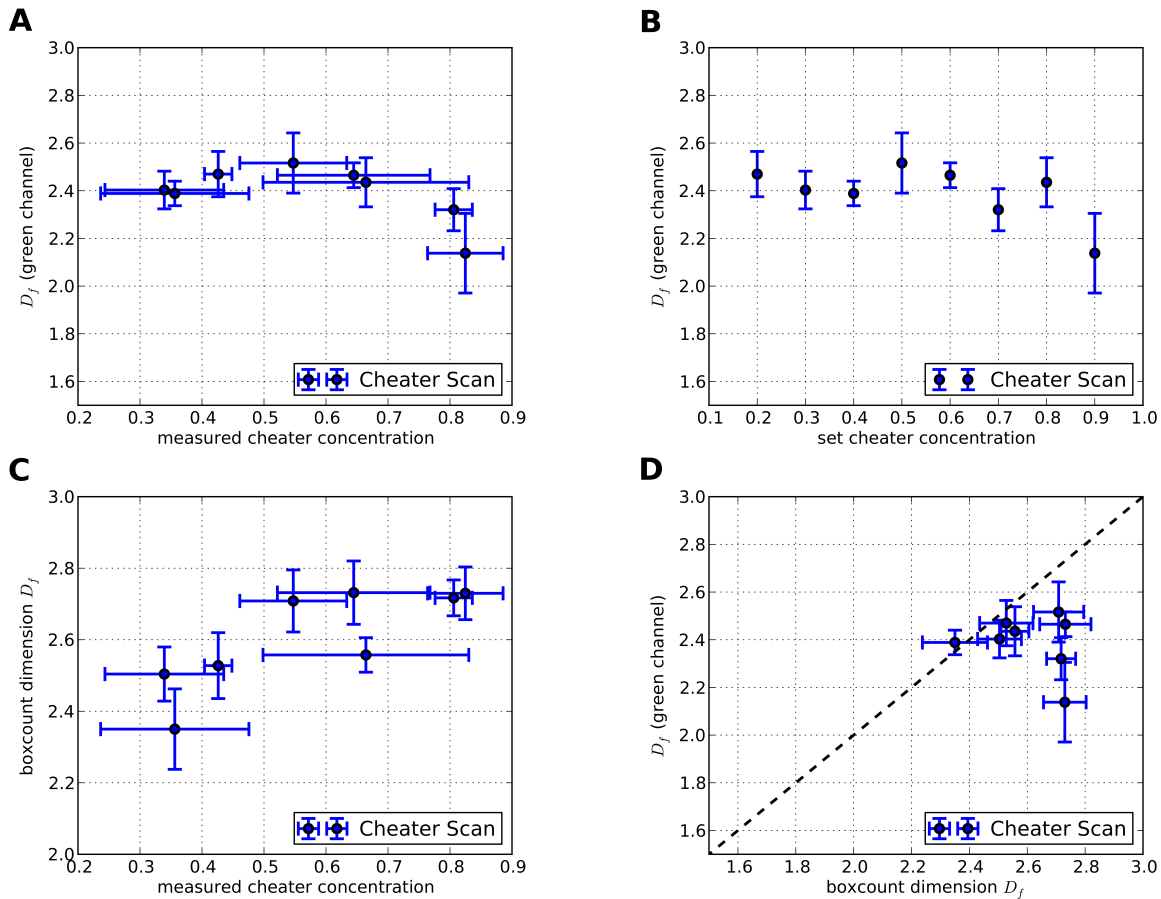
Figure 3.29: Influence of the cheater cells on the structure of the flocs. For this figure I used the fractal dimension of the floc and the fractal dimension of the green channel. In this case "green channel" denotes the structure of the bearer cells that is left if all cheater cells are removed. **A**: The measured cheater concentration has almost no influence on the bearer cell structure ($r_{xy} = -0.35$) **B**: The structure of the bearer cells is not influenced by the initial cheater concentration ($r_{xy} = -0.32$). This indicates that the cheater cells are not part of the integral structure of the floc. **C**: The fractal dimension of a floc clearly increases with increasing cheater concentration ($r_{xy} = 0.53$). So the density and regularity of a floc increase with increasing cheater concentration indicating that the cheater cells fill up gaps in the bearer cell structure. **D**: The fractal dimension of the complete flock is always larger than the fractal dimension of the bearer structure (the two parameters are also uncorrelated, $r_{xy} = -0.22$) supporting the argument of cheater cells filling the gaps of the bearer structure.

Figure 3.29**C** shows that on the other hand the fractal dimension of the complete floc increases with increasing cheater concentration. The correlation coefficient of this increase is $r_{xy} = 0.53$. The fractal dimension of the complete floc is always larger than the fractal dimension of the green channel (see figure 3.29**D**) indicating that the cheater cells make the flocs more regular and more volume-like.

To summarize the results on cheater positions so far four statements can be made: 1) Cheaters inhibit flocculation by decreasing the effective concentration. 2) The structure of bearer cells in a floc is neither influenced by the set cheater concentration in the medium nor by the actual cheater concentration inside the floc. 3) The higher the fraction of cheaters inside a floc, the denser and more volume-like the floc gets. 4) The complete floc is denser and more volume-like than the stripped floc consisting only of bearer cells.

These four results lead to the assumption, that the cheater cells fill up the holes inside the floc that are left by the bearer cells rather than being part of the main structure.

Another interesting fact is revealed by conducting simulated attacks on the flocs, as shown in figure 3.30 (see section 3.1 for a description of the method).

These results show that the cheater cells are more exposed to the environment than bearer cells. In addition to fill the remaining holes inside the floc they are enriched at the floc's surface supporting the idea of cheaters as a first line of defense as claimed by Smukalla et al. (2008). The average shift of about 30% in favor of bearer cells is comparable to their findings concerning the advantage of Flo1 bearer cells under chemical stresses (see (Smukalla et al., 2008) figure 7).

3.4.3 Simulation Results

One of the biggest drawbacks in agent based modeling is the large amount of computing power that is necessary to carry out the simulations. Even with the parallelization mentioned in
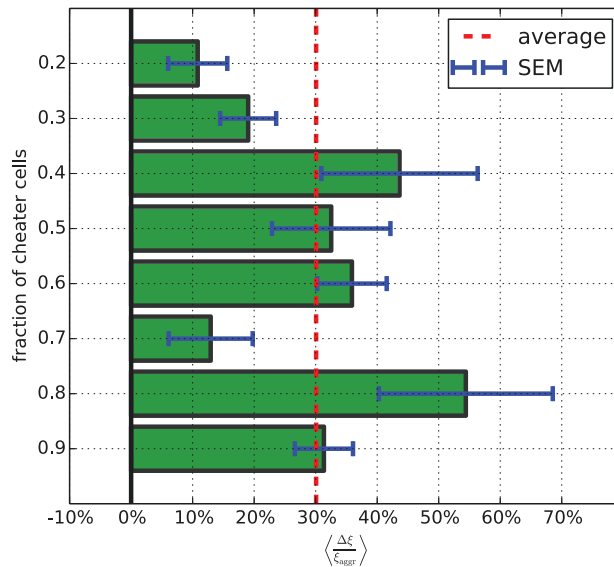
Figure 3.30: The simulated attack reveals that the cheaters are indeed more exposed to the environment. This suggests that they will suffer higher stresses and form a first line of defense as suggested by Smukalla et al. (2008). The average shift of about 30% in favor of bearer cells is comparable to the growth penalty from which they suffer when producing the Flo5 binding protein.

section 3.2.3 it is not possible to do a large parameter scan or a data fit on a single workstation. This restricts this work to a proof of concept.

3.4.3.1 Flocculation Essays

Although I was not able to fit the simulation parameters to the experimental data, the flocculation essays are in strikingly good agreement with the experimental data. Table 3.6 summarizes the used parameters.

Figure 3.31 **A** indicates that – as in the experiments – the flocculation requires a minimal cell density to be efficient.

If the bearer cells are diluted with cheater cells the efficiency of the flocculation decreases rapidly after some time of stagnation as shown in figure 3.31 **B**.

Parameter	Value (arbitrary units)	Description
N	32768	simulated cells
Δt	0.01	division of time-steps
$l_{0,\text{Spring}}$	1.0	equilibrium length of bond (incl. cell radii)
k_{Spring}	4.99787379	spring constant of double bond
$F_{\text{rand.}}$	$7.52538511 \cdot 10^{-2}$	amplitude of single cell random force
F_{box}	$7.88825084 \cdot 10^{-2}$	amplitude of area random force
$F_{\text{roat.}}$	$1.35553341 \cdot 10^{-2}$	amplitude of rotational force
$\Gamma_{\text{fric.}}$	1.09971370	friction parameter
d_{crack}	1.025	distance at which the bonds will break (incl. cell radii)
$p_{\text{react,e}-\text{e}}$	1.0	reaction probability for two bearer cells
$p_{\text{react,e}-\text{c}}$	1.0	reaction probability for bearer and cheater cells
$k_{\text{coll.}}$	2	spring constant for collision
$\gamma_{\text{coll.}}$	2	collision dampening
$\sigma_{\text{coll.}}$	0.5	collision shear
d_{cell}	1	cell diameter

Table 3.6: Parameters used for the simulations. The spatial size of the simulated world and cheater rates were set according to the experiments. It is not possible to give the real parameters at this point, because with the hardware at hand it was not possible to fit the data quantitatively.

Finally figure 3.31 **C** shows the cheater concentration in the supernatant as a function of the cheater concentration is significantly higher than the fraction of cheater cells used in the mix. This behavior is the same in the experiments and in the simulations.

All in all, the simulations qualitatively and almost quantitatively reproduce the experimental findings concerning the flocculation essays without even performing a larger parameter scan.

3.4.3.2 Three Dimensional Structure of Flocks

Using the same parameters that reproduce the flocculation essays so well do not reproduce all structural properties of the flocs especially their size quantitatively. However, since I was not able to perform large parameter scans and data fits, this is likely to be a matter of finding the right combination of parameters.

Figure 3.32 shows a simulated floc for the same initial conditions that produced the floc in
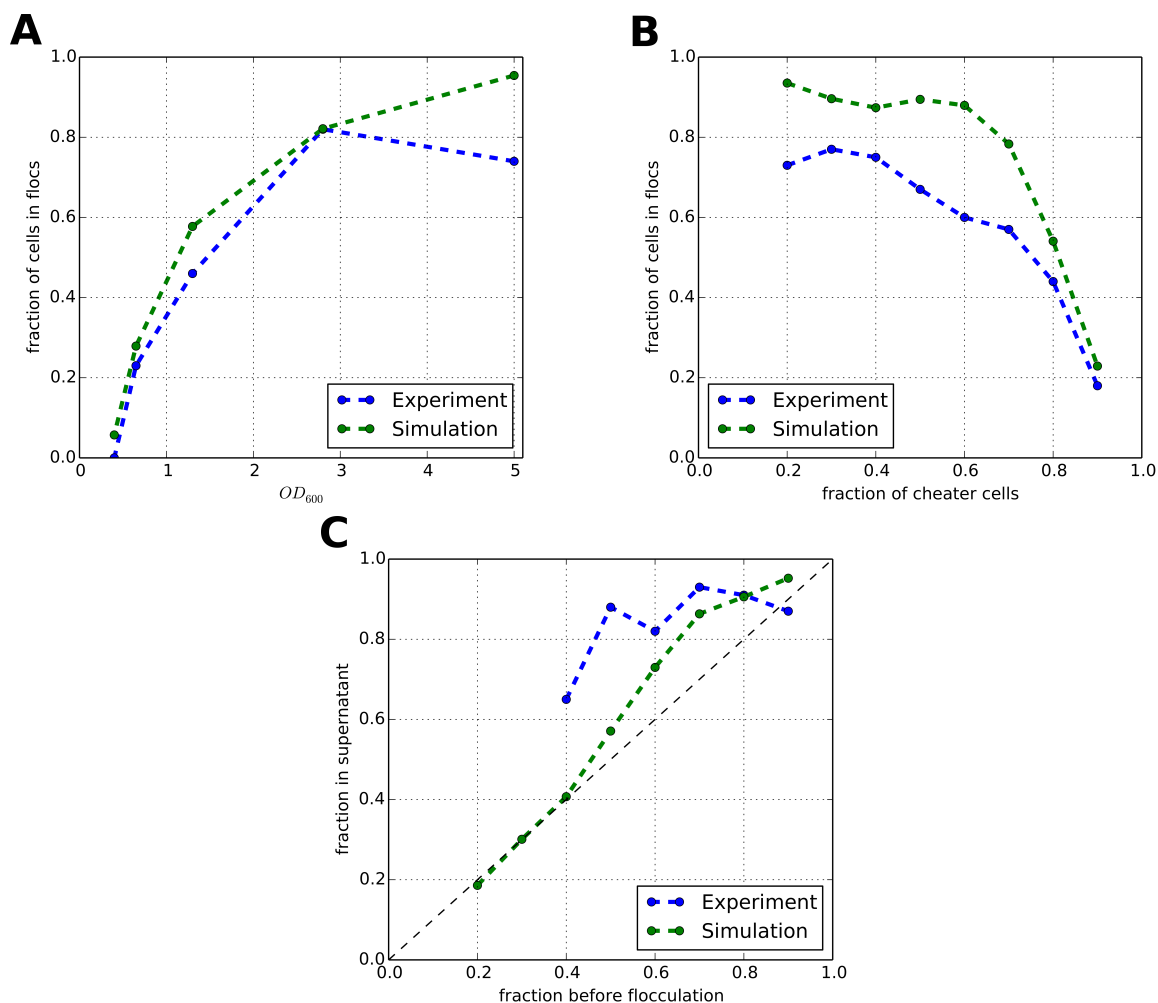
Figure 3.31: Flocculation essays: comparison of the experimental data and the simulations. **A** Flocculation efficiencies for different initial cell concentration from experiment and simulation. The simulation results nicely fit the experimental data leaving aside a little higher efficiencies that might be due to the threshold from which a cluster is considered a floc that settles within one minute. **B** Flocculation efficiencies for different fractions of cheater cells in the solution. The simulations show almost the same decrease in efficiency with an increasing fraction of cheater cells. **C** The fraction of cheater cells in the supernatant after flocculation is higher than the one used before flocculation. This clearly shows that cells expressing the *Flo*-gene are accumulating inside the flocs.

figure 3.23. It clearly shows that the cells do aggregate, but the floc seems too small and too dense compared to the experimental floc.

This can be caused by different effects. The used combination of force parameters might
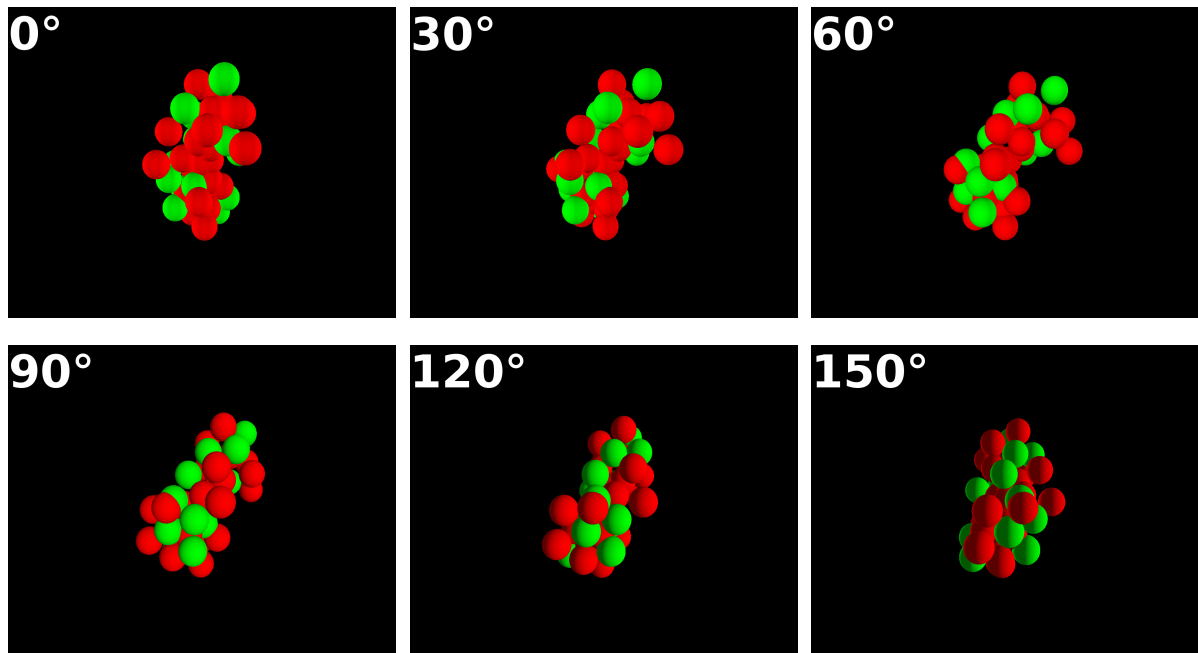
Figure 3.32: Three dimensional view of a simulated floc at the same conditions as the floc shown in figure 3.23. Compared to the experimental images in figure 3.23 the floc looks much smaller (it only contains 42 cells).

result in too slow cells that are not able to aggregate enough through the simulated time. On the other hand the simulated time might be too short or, to be more precise, the combination of force amplitudes and simulated time does not fit the experiment. To simulate larger forces one would have to decrease δt resulting in more iterations per time-step and finally in longer runtimes.

The increased density directly influences the FWHM of $g(r)$ as a measure for the flocsize because if the density strongly increases with the number of cells, the pair-correlation function becomes steeper resulting in a decreasing instead of an increasing FWHM for larger flocs. Figure 3.33 visualizes this effect.

While the flocsize measured by the number of cells inside the floc increases with increasing initial concentration (as did the flocsize measured by the FWHM of $g(r)$ in the experiments), the FWHM of $g(r)$ decreases indicating a strongly increased density of the floc. This density is caused by a non-optimal choice of parameters especially with respect to the relation of
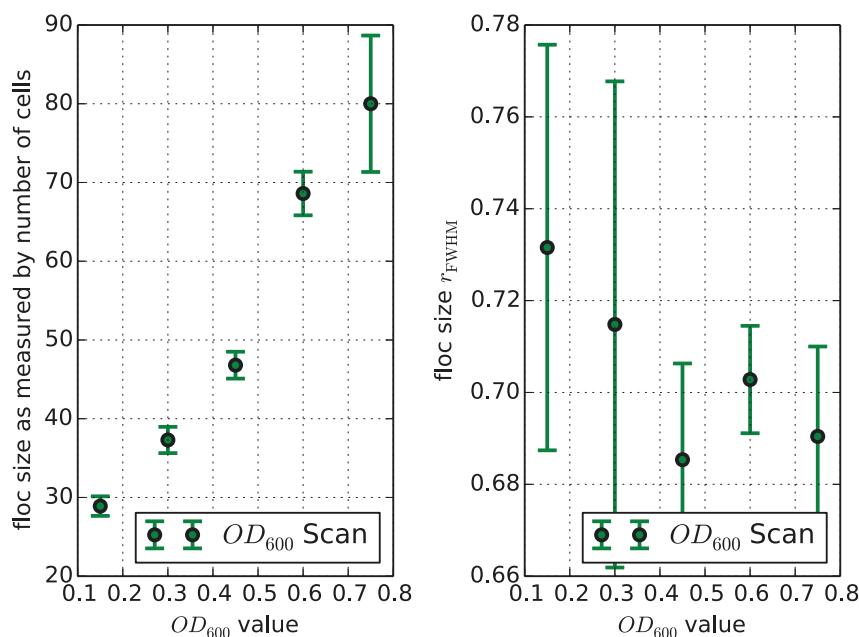
Figure 3.33: Aggregate sizes for the simulated OD_{600} scans. The left image shows the result if we use the number of cells in a cluster as a measure for the flocsize. It shows that the simulation qualitatively reproduces experimental data ($r_{xy} = 0.985$). The image on the right shows the results for the flocsize derived from the pair-correlation function. Compared to the experimental findings this shows the almost opposite behavior ($r_{xy} = -0.795$). This might be caused by the flocs becoming too dense when they grow larger due to a non-optimal relation between collision and attraction force.

attractive and repulsive forces between cells. In general, the repulsive forces should be larger than the attractive forces preventing the cells to occupy the same space. However, it seems that with the chosen parameter set this is not fulfilled and two cells can occupy the same space temporary, resulting in bonds between second next neighbors that do not occur in nature. Through this "forbidden" bonds the floc will become much denser than the natural ones.

A side effect of this is that the FWHM of $g(r)$ can not be trusted as a measure for the flocsize. Instead I will from now on use the average number of cells as a trustworthy parameter for the size of the flocs.

Figures 3.34 **A** - **C** show the results for the structural analysis of the simulated flocs. The first thing that stands out are the obvious correlations between flocsize, set cheater concen-

tration and set OD_{600} value and the fractal dimension that are not present in the experimental data (see figure 3.25).
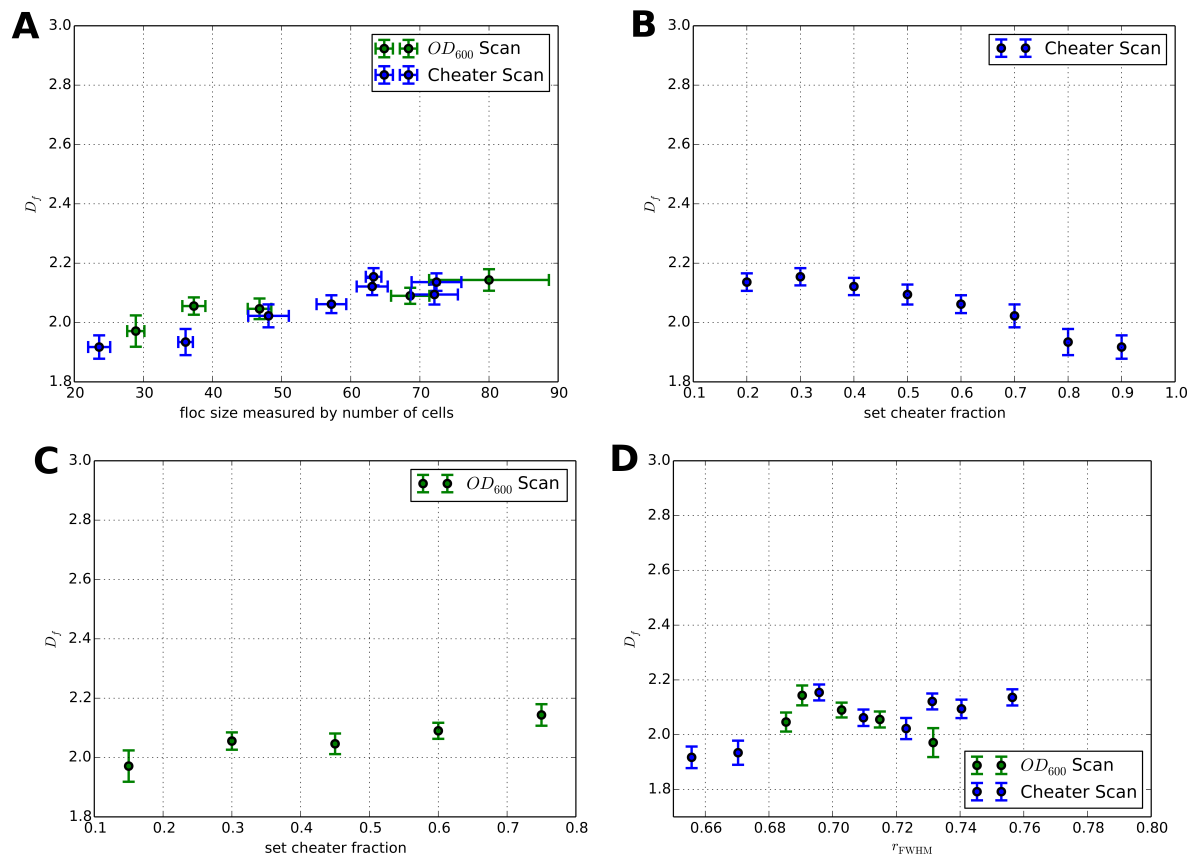
Figure 3.34: **A**: Size of flocks vs. boxdimension. There is a strong positive correlation ($r_{xy} = 0.896$, $\langle D_f \rangle = 2.06$) that result from the combination of two effects. The small size of the flocs in combination with them being much too dense leads to a decrease in fractal dimension. **B** and **C**: In contrast to the experimental findings the simulation results show a rather strong correlations between set cheater fraction and boxcount dimension ($r_{xy} = -0.957$), and set OD_{600} value and boxcount dimension ($r_{xy} = 0.949$) that are caused by the correlation strong correlation between size and density of the flocs. **D**: If we derive the flocsize from the pair-correlation function, the correlations weaken a great deal ($r_{xy} = 0.494$) hinting that the other correlations are indeed caused by the increased density of cells inside a floc.

These correlations are passed on from the negative correlation between the number of cells in a floc and the FWHM of $g(r)$ and as such a result of the too high cell density. This is

indicated in figure 3.34 **D**, where the correlation of flocsize and fractal dimension is strongly decreasing compared to the correlation of flocsize measured by the number of cells of a floc and the fractal dimension.

3.4.3.3 Role of the Cheater Cells

As in the case of the experimental flocs, the flocculation essays already show the inhibitory effect of cheaters in the simulations. In addition there is an even stronger decrease in flocsize with increasing cheater concentration in the simulation (see figure 3.35) than in the experimental data (see figure 3.28). The correlation in the simulations ($r_{xy,\text{sim.}} = -0.90$) is about 1.6 times stronger than in the experiments ($r_{xy,\text{exp.}} = -0.55$).
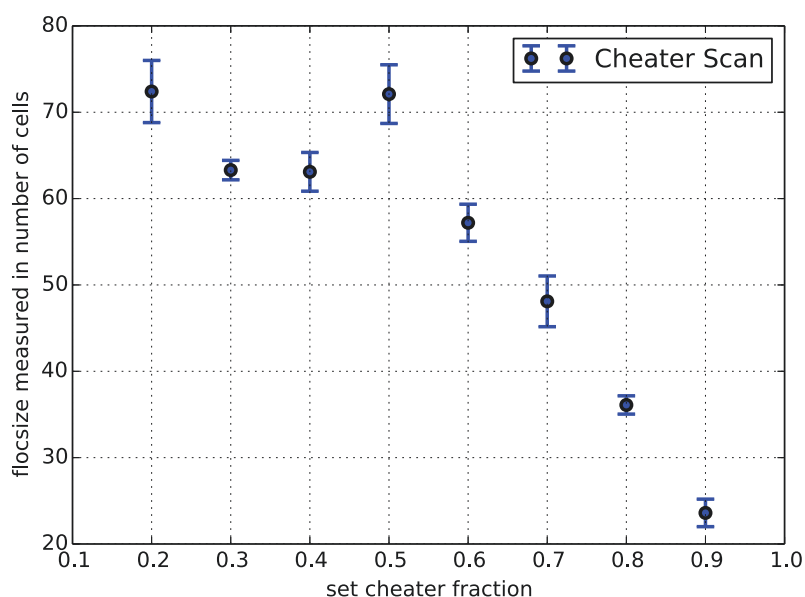
Figure 3.35: The floc size is strongly decreases with increasing initial cheater concentration in the simulations. The floc size shows a strong decreasing trend for increasing cheater concentration with a correlation coefficient of $r_{xy} = -0.900$.

Concerning the position of the cheater cells inside the flocs, the simulation data is not as conclusive as the experimental data because of the increased density of the flocs. This strongly

influences the fractal dimension and therefor passes on the negative correlation between cell number and density (see figures 3.36 **A-C**).
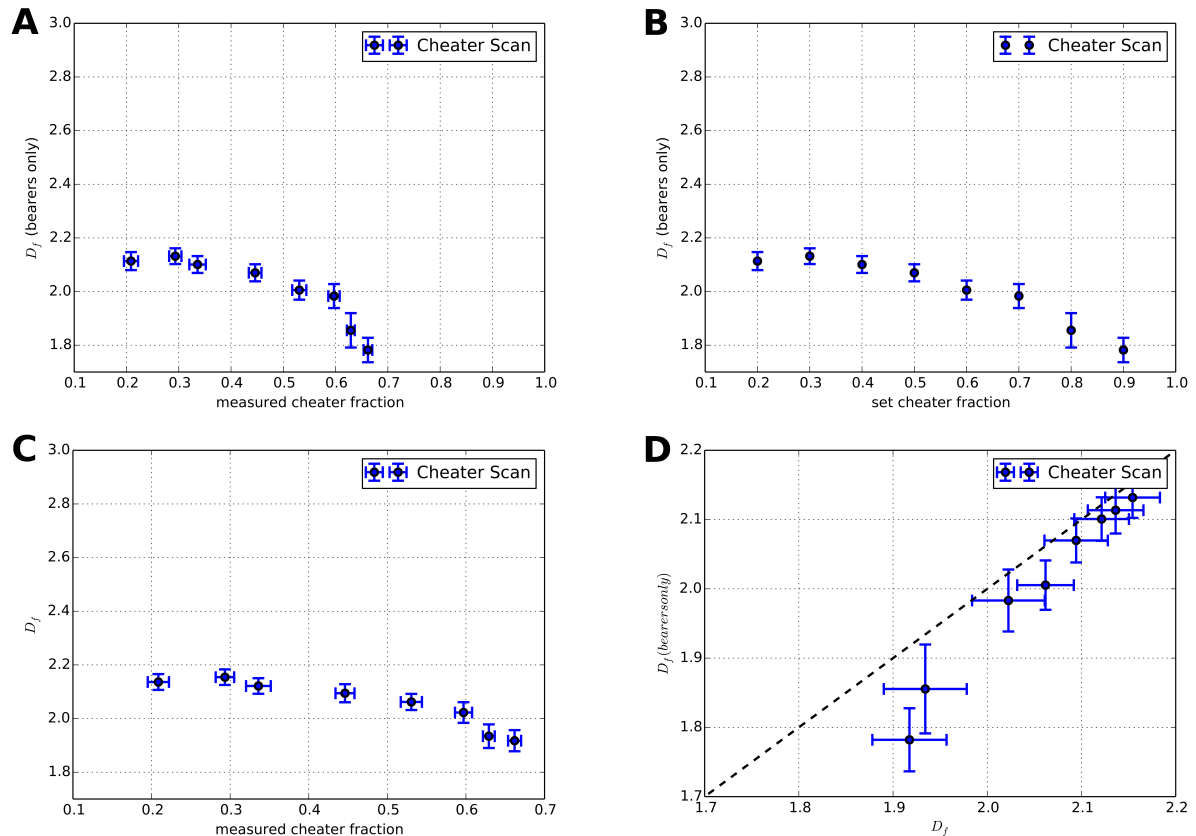
Figure 3.36: Influence of the cheater cells on the structure of the simulated flocs. **A-C**: In contrast to the experimental findings the simulated flocs show strong correlations between measured and set concentration and fractal dimension of the bearer floc. In these cases it is also possible and likely that the correlation is caused by the small size and the more dense structure of the flocs rather than by the fractions of cheater cells. **D**: The fractal dimension of the complete floc is always lower than that of the bearer only structure as it was in the experiments, indicating that the cheaters still fill the holes in the floc even though the floc is much too dense.

However, the fact that the fractal dimension of the complete floc is always larger than the fractal dimension of the floc stripped of all cheater cells is a strong hint that the cheaters also fill the holes in case of the simulations.

Finally, the simulated attack reveals that the cheaters are more exposed to the environment

in the simulations as well. As shown in figure 3.37 the shift is 21.36% in favor of bearer cells.
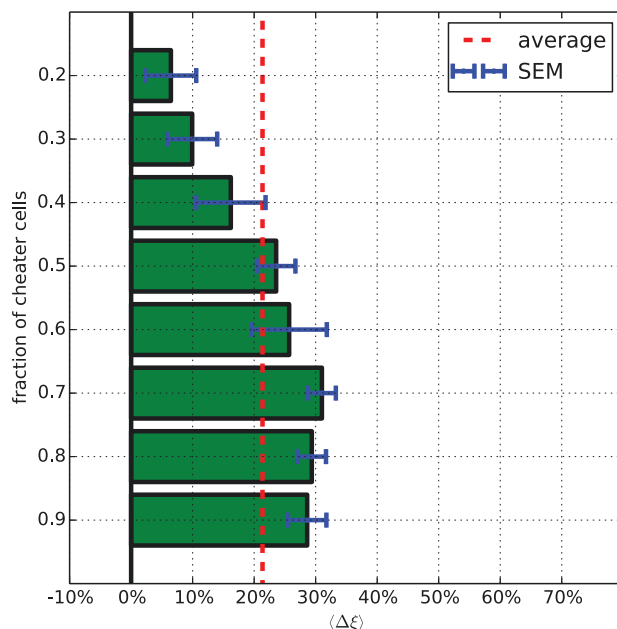
Figure 3.37: The simulated attack reveals that the cheaters in the simulations are indeed more exposed to the environment, too. The average shift is about 10% smaller than in the experiments but it still is 21.36% in favor of the bearer cells.

All in all, the simulation results reproduce the experimental findings qualitatively with the small drawback of producing flocs that are too small and too dense. This shows that the increased exposition of cheater cells is a result of the mechanical properties of the bonds between cheater and bearer cells compared to the properties of the bonds between two bearer cells.

3.4.4 Discussion and Summary

In this section I presented a combination of a novel experimental approach and an individual based model to analyze and understand the principles of the dynamics and structure of the flocculation of *S. cerevisiae* mediated by Flo5. The four most important experimental findings

are that the flocculation efficiency strongly depends on the cell concentration, that flocs are indeed fractal structures (which implies a self similarity), that the basic unit of this self similarity is a floc containing 39 cells and that cheaters inhibit flocculation only by diluting the bearer cells while they fill up holes in the flocculation process and are strongly enriched on the flocs surface.

Especially the role of the cheater cells during and at the end of the flocculation process strongly supports the idea of flocculation as a protective mechanism as reviewed in Soares (2011), Bauer et al. (2010) and Brückner and Mösch (2012). The enrichment on the surface that was shown for the similar Flo1 system is also present in Flo5 flocculation facilitating the cheater cells as a first line of defense as suggested by Smukalla et al. (2008). In addition this work provides analytical tools to analyze the three dimensional structure of flocs (or other cell aggregates) that are non-regular objects, i.e. fractal objects.

As fractal objects the flocs show some kind of self similarity on a certain range of scales as suggested for fractals by Mandelbrot (1977). As mentioned in the first paragraph the smallest of these scales are flocs consisting of 39 cells. While this seems to be a reasonable minimum of the used mechanism (see section 3.4.2.2) it remains to investigate why it is this exact number.

The structure of the flocs that incorporate cheater cells suggests that the flocs consist of a stable bearer backbone with holes stuffed by cheater cells. In this way though inhibiting flocculation by diluting the cell culture the cheater cells can not destabilize the flocs. However through this dilution process they reduce the efficiency of flocculation and the resulting flocs' sizes.

In addition, the mechanical agent based model – even though at a level of a proof of concept – qualitatively reproduces a good part of the experimental findings. To make it fit quantitatively one has to put a lot of computational effort into the simulations. However, this qualitative agreement proves, that there is no active mechanism that rules out the cheater cells but that it is simply the outcome of a two-fold difference in the binding forces that leads to a self

assembly in a way that excludes cheater cells from a big part of the benefits of flocculation and therefore compensates for their advantage in growth.

Another interesting question at hand is if Flo5 could be considered a green beard gene as defined by Dawkins (1976) despite the fact that it does not specifically mediates an interaction between carriers of this gene. The exposition of the cheater cells and their general position inside the flocs are arguments that favor the view, that Flo5 fulfills the conditions to be a green beard gene: while not directly recognizing other carriers of the gene it strongly favors carriers of the gene just by reinforcing the interaction compared to the interaction to other cells.

Chapter 4

Conclusion

This work gives new insights into the dynamics and structure of cellular aggregation starting from cell motility as a precursor. I developed a generic model for mammalian cell migration that is capable of explaining and characterizing several details of the motility of mammalian cells purely based on mechanics.

Furthermore, this study introduces a toolkit to quantify patterns and new experimental procedures to visualize and "measure" patterns. As examples for different types of aggregation I presented the aggregation of mouse embryonic fibroblast cells in two dimensions and the three dimensional flocculation of yeast cells. Hand in hand with the experiments I developed an individual based model capable of reproducing the dynamics and the structure of these aggregation processes.

This chapter summarizes the results for the cell motility model, the MEF cell aggregation and the flocculation. It also presents an outlook on further work that can be done in this fields.

4.1 Mammalian Cell Motility

The mesenchymal mode of migration of epithelial cells is a very important mode of migration in mammalian cells. In chapter 2 I introduced a generic two dimensional model that is capable to simulate the large scale crawling dynamics of an epithelial cell in two dimensions. The experimental observations at steps in substrate rigidity observed by Bordeleau et al. (2013) and cellular morphology observed by Lo et al. (2000) are nicely reproduced by the model. This model captures the very high efficiency in following gradients that has been observed by (Theveneau et al., 2010).

The model can be used to estimate transition probabilities through rigid-soft interfaces which are important in cancer metastasis (Bordeleau et al., 2013). Furthermore, the refractive indices introduced in section 2.4 can be used to characterize substrate rigidity including an appreciation of the biological impact of the material compared to a reference substrate.

All in all, my model produces reasonable results that fit many experimental observations qualitatively. However, further experimental quantifications can be performed to validate the model.

To quantify the chemotactic efficiencies one could rely on an experiment, where a gradient is created using a Dunn-Chamber (Zicha et al., 1991) combined with time-lapse microscopy. This can be done for a large number of cells using automation techniques (Fotos et al., 2006) and cell-tracking methods (Chen et al., 2006).

To validate the results for spatial substrate heterogeneities in substrate adhesiveness, chemoattractant concentration or substrate stiffness one could use the methods presented by Bordeleau et al. (2013) to produce flat substrates of different stiffness and measure the polarization axes of the cells before and after the interface and the transmission coefficients with automatized time lapse microscopy and the before mentioned tracking methods (Chen et al., 2006).

So far I have neglected the mechanical effects that the cells have on the substrate. If a cell attaches to the substrate and contracts, the substrate locally stiffens. This gradual increase in stiffness could mediate a cell-cell interaction since it would attract cells that are in proximity and in this way promote aggregation.

Another possible prosecution of this research is the extension of the model to three dimensions. This requires the incorporation of drags that the cell experiences while migrating through the extracellular matrix (DiMilla et al., 1991).

4.2 MEF Cell Aggregation

Our experimental approach combines time-lapse microscopy with new methods to quantify patterns derived from fractal geometry and graph theory. The analysis of the microscopy series showed that the aggregation can be subdivided into two phases: A spreading phase of about 20 hours and a contraction phase that can be distinguished by monitoring the surface area covered by cells.

It was possible to estimate the aggregate growth rate that depends on the initial cell concentration by using the pair-correlation function as a measure for aggregate size. For very low cell concentrations (10000 - 40000 cells per well) the growth rate is in the order of $0.1\,\mu\mathrm{m/h}$. For the higher concentrations (120000-160000 cells per well) the rate increases more than tenfold to $1.4\,\mu\mathrm{m/h}$. This shows that – as expected – the aggregation is much faster at high cell concentrations.

As a structural measure for the regularity and density of the cluster we used the pair-correlation function. A higher regularity of the cluster, i.e. a fractal dimension closer to two, indicates a more area like structure. For each experiment the fractal dimension finally reaches a saturation value depending on the initial cell concentration. The explanation of this phenomenon is the dependence of the fractal dimension (and therefore the spatial structure of an aggregate) on the aggregate size. Smaller aggregates are less regular. As the aggregates grow in size the fractal dimension increases linearly until at a size of about $20.33\,\mu\mathrm{m}$ the fractal dimension saturates at $D_f = 1.7$. This structural similarity to aggregates formed by diffusion limited aggregation (Witten Jr. and Sander, 1981) hints that the random motion of the cells, i.e. the "diffusion", might be an important factor in the process of aggregation.

However, the introduced model nicely reproduces the experimental findings quite well while not explicitly modeling a random motion. Fitting the temporal dynamics of the fractal dimension to the experimental data leads to a good qualitative agreement. In the simulations

we could also identify two phases: growth by acquisition of single cells in the first few hours and growth by merger of larger clusters.

In the simulations we measured the size of the aggregates as the average number of cells per aggregate. While showing a very similar behavior as the aggregate sizes in the experiments it is difficult to compare these findings quantitatively without fitting the model to a larger amount of experimental data. The simulation completely neglects the cell shape that is of great importance for the final structure of the aggregates because the cells are the building blocks of them.

If we analyze the dependency of the fractal dimension on the aggregate size we also see a transition from an almost linear increase to a saturation. The transition is not as sharp as it is in the experiments, but it reaches a similar saturation value ($D_{f,\text{sim.}} = 1.76$ compared to $D_{f,\text{exp}} \approx 1.70$).

The simulation allows for the analysis of quantities that are unaccessible for experimental analysis. The temporal dynamics of the aggregation degree prove that the aggregation gets faster and more efficient with higher cell concentrations. The finding that the aggregation takes place in two phases in the simulation – growth by acquisition of single cells and growth by merger of clusters – is supported by the temporal behavior of the clustering coefficient that first strongly increases identically for all initial cell concentrations and the splits up to saturate at values depending on the initial cell concentration.

All in all, this study emphasizes the importance of approaches that combine quantitative analysis of patterns with modeling. Especially the sharp transition of the fractal dimension as a function of the aggregate size was found by carefully quantifying the microscopy images using techniques from various mathematical fields.

To further quantify the two aggregation phases (spreading and contraction) one has to increase the number of experiments and analyze the average behavior to avoid errors occurring from non-isotropic cell concentrations inside one well and from other noise effects. For the

simulations it would be interesting to analyze a model extension that incorporates cell shape and random motion. While the latter can be implemented straight forward using the results from chapter 2 the first is challenging because the cell shape is highly irregular and it changes randomly with time and through mechanical forces that are applied to the cell.

Another phenomenon that has been neglected in the model is the stacking of cells, i.e. spreading in the third dimension. Experimentally it becomes evident when the surface area covered by cells decreases below the initial area that the cells either decrease their projected size or stack up into the third dimension. Since this phenomenon is not easily studied experimentally it remains a challenging question.

4.3 Flocculation of *S. cerevisiae*

We used a combination of a novel experimental approach to make a three dimensional visualization of flocs and an individual based model towards the understanding of the principles of *S. cerevisiae* flocculation. The most important results of this study are: (1) the flocculation efficiency strongly depends on the cell concentration, (2) flocs are fractal structures (which implies a self similarity), (3) the basic unit of this self similar structure is a floc that contains 39 cells, (4) cheaters inhibit flocculation only by diluting the bearer cells without destabilizing the flocs and (5) cheaters are strongly enriched on the surface with a shift of about 30% in favor of the bearer cells.

Especially the role of the cheater cells as a filling of a bearer backbone strongly supports the idea of flocculation as a protective mechanism as suggested in Soares (2011), Bauer et al. (2010) and Brückner and Mösch (2012). The enrichment on the surface of a floc that was shown for the similar Flo1 system is also present in Flo5 mediated flocculation facilitating the cheater cells as a first line of defense as suggested by Smukalla et al. (2008). In contrast to this study of Flo1 mediated flocculation we were able to quantify this enrichment. Furthermore, this work provides an analytical toolkit to analyze the three dimensional structure of flocs (or other cell aggregates) that are fractal objects.

Being fractal structures flocs show self similarity. In contrast to mathematical fractals introduced by Mandelbrot (1977) the self similarity of flocs has a finite range of scales in which it really is fractal. The minimum scale of this range is a structure containing 39 cells that I found using a monte-carlo approach to find the smallest structure reproducing the experimental findings.

The fact that the agent-based model that is solely based on mechanical interaction qualitatively reproduces a good part of the experimental findings proves, that there is no active mechanism that rules out the cheater cells. Instead the two-fold difference in the binding forces

lead to a self assembly in a way that excludes cheater cells from a big part of the benefits of flocculation and therefore compensates for their advantage in growth.

However, since a quantitative fit of the model requires a lot of computational effort this results stay on a qualitative level but they provide a proof of concept for the individual based, mechanical modeling of cellular aggregation.

While flo5 does not fulfill the literate definition of a green beard gene as defined by Dawkins (1976) because it does not specifically bind carriers of the same gene. However, the exposition of the cheater cells and their general position inside the flocs are arguments that favor the view, that flo5 does mediate a very similar effect as do green beard genes: while not directly recognizing other carriers of the gene it strongly shifts the outcome of the aggregation in favor of carriers.

Further work could approach this topic from a more topological view. While establishing a 39 cell floc as the unit structure of flocs it remains unclear why it is this exact number. The exponential dependency between the minimal aggregate size and the fractal dimension is a hint that there might be some topological reason for this.

Another interesting question would be the manipulation of the flocculent system. If one could achieve a quantitative fit of the model it is easy to predict the effects of a change in model parameters. This could finally lead to industrial applications in the field of fermentation for the production of chemicals if it is possible to further improve the efficiency of the flocculation.

List of Figures

128

List of Tables

Bibliography

Julius Adler. 'Chemotaxis in Bacteria'. *Science*, 153:708–716 (1966). 17

Julius Adler. 'Chemotaxis in Bacteria'. *Annual Review of Biochemistry*, 44:341–356 (1975). 17

R. Albert and *A.-L. Barabási*. 'Statistical mechanics of complex networks'. *Reviews of Modern Physics*, 74:47–97 (2002). 45, 51, 53

Bruce Alberts, *Alexander Johnson*, *Julian Lewis*, *Martin Raff*, *Keith Roberts* and *Peter Walter*. Garland Science, fourth Auflage (2002). 5, 16

Uri Alon. *An Introduction To Systems Biology*. Chapman & Hall/CRC (2007). 17

Florian F Bauer, *Patrick Govender* and *Michael C Bester*. 'Yeast flocculation and its biotechnological relevance'. *Applied microbiology and biotechnology*, 88(1):31–39 (2010). 1, 11, 114, 123

L. C. Baxter, *V. Frauchiger*, *M. Textor*, *I. ap Gwynn* and *R. G. Richards*. 'Fibroblast and

Osteoblast Adhesion and Morphology on Calcium Phosphate Surfaces'. *European Cells and Materials*, 4:1–17 (2002). 33

Daniel ben-Avraham and *Shlomo Havlin. Diffusion and Reactions in Fractals and Disordered Systems*. Cambridge University Press, 1 Auflage (2004). 48

Eric Bonabeau. 'Agent-based modeling: Methods and techniques for simulating human systems'. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Suppl 3):7280–7287 (2002). 54

F Bordeleau and *Luc Galarneau*. 'Keratin 8/18 modulation of protein kinase C-mediated integrin-dependent adhesion and migration of liver epithelial cells'. *Molecular Biology Of The Cell*, 21:1698–1713 (2010). 5, 14

Francois Bordeleau, *Lauren N Tang* and *Cynthia a Reinhart-King*. 'Topographical guidance of 3D tumor cell migration at an interface of collagen densities.' *Physical biology*, 10(6):065004 (2013). 6, 25, 26, 35, 39, 40, 118

Stefan Brückner and *Hans-Ulrich Mösch*. 'Choosing the right lifestyle: adhesion and development in Saccharomyces cerevisiae'. *FEMS microbiology reviews*, 36(1):25–58 (2012). 3, 11, 12, 43, 90, 114, 123

Mathias Buenemann, *Herbert Levine*, *Wouter-Jan Rappel* and *Leonard M Sander*. 'The role of cell contraction and adhesion in dictyostelium motility.' *Biophysical journal*, 99(1):50–8 (2010). 21, 23, 39

Armin Bunde and *Shlomo Havlin*, editors. *Fractals in Science*. Springer-Verlag (1994). 45, 47

Darci T Butcher, *Tamara Alliston* and *Valerie M Weaver*. 'A tense situation: forcing tumour progression.' *Nature reviews. Cancer*, 9(2):108–22 (2009). 6, 7, 24, 25, 35

Brian A. Camley, *Yanxiang Zhao*, *Bo Li*, *Herbert Levine* and *Wouter-Jan Rappel*. 'Periodic migration in a physical model of cells on micropatterns'. *Phys. Rev. Lett.*, 111:158102 (2013). 23, 39

S. B. Carter. 'Principles of cell motility: the direction of cell movement and cancer metastasis'. *Nature*, 208(5016):1183–1187 (1965). 4, 14

Vikash Chandra, *G Swetha*, *Sudhakar Muthyala*, *Amit K Jaiswal*, *Jayesh R Bellare*, *Prabha D Nair* and *Ramesh R Bhonde*. 'Islet-like cell aggregates generated from human adipose tissue derived stem cells ameliorate experimental diabetes in mice'. *PLoS One*, 6(6):e20615 (2011). 2, 9, 43, 73

Xiaowei Chen, *Xiaobo Zhou* and *Stephen TC Wong*. 'Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy'. *IEEE Transactions on Biomedical Engineering*, 53(4):762–766 (2006). 40, 118

M Chicurel. 'Cell migration research is on the move'. *Science*, 295(January):606–609 (2002). 4, 5, 6, 14, 24

WV Cruess, *CA Weast* and *R Gillilland*. 'Summary of practical investigations on film yeast'. *Fruit Products J*, 67:229–231 (1938). 11

Richard Dawkins. *The selfish gene*. Oxford University Press Oxford (1976). 3, 90, 115, 124

P a DiMilla, *K Barbee* and *D a Lauffenburger*. 'Mathematical model for the effects of adhesion and mechanics on cell migration speed.' *Biophysical journal*, 60(1):15–37 (1991). 6, 7, 24, 25, 119

Anne M Dranginis, *Jason M Rauceo*, *Juan E Coronado* and *Peter N Lipke*. 'A biochemical guide to yeast adhesins: glycoproteins for social and antisocial occasions'. *Microbiology and Molecular Biology Reviews*, 71(2):282–294 (2007). 12, 44, 66

PR Fisher, *Rainer Merkl* and *G Gerisch*. 'Quantitative analysis of cell motility and chemotaxis in Dictyostelium discoideum by using an image processing system and a novel chemotaxis chamber providing stationary chemical gradients.' *The Journal of cell biology*, 108(3):973–984 (1989). 5, 20

G Flierl, *D Grünbaum*, *S Levins* and *D Olson*. 'From individuals to aggregations: the interplay between behavior and physics'. *Journal of Theoretical Biology*, 196(4):397–454 (1999). 2

Kevin R Foster, *Katie Parkinson* and *Christopher RL Thompson*. 'What can microbial genetics teach sociobiology?' *TRENDS in Genetics*, 23(2):74–80 (2007). 3

Joseph S Fotos, *Vivek P Patel*, *Norman J Karin*, *Murali K Temburni*, *John T Koh* and *Deni S Galileo*. 'Automated time-lapse microscopy and high-resolution tracking of cell migration'. *Cytotechnology*, 51(1):7–19 (2006). 40, 118

Kenneth J Freiberg and *WV Cruess*. 'A study of certain factors affecting the growth of flor yeast'. *Applied microbiology*, 3(4):208 (1955). 11

Xiongfei Fu, *Lei-Han Tang*, *Chenli Liu*, *Jian-Dong Huang*, *Terence Hwa* and *Peter Lenz*. 'Stripe formation in bacterial systems with density-suppressed motility'. *Physical review letters*, 108(19):198102 (2012). 8, 42

Danny Fuller, *Wen Chen*, *Micha Adler*, *Alex Groisman*, *Herbert Levine*, *Wouter-Jan Rappel* and *William F Loomis*. 'External And Internal Constrains on Eukaryotic Chemotaxis'. *Proceedings of the National Academy of Sciences of the United States of America*, 107(21) (2010). 5, 19, 20

Juana M Gancedo. 'Control of pseudohyphae formation in Saccharomyces cerevisiae'. *FEMS microbiology reviews*, 25(1):107–123 (2001). 11

A. Gerisch and *M. a. J. Chaplain*. 'Mathematical modelling of cancer cell invasion of tissue: local and non-local models and the effect of adhesion.' *Journal of theoretical biology*, 250(4):684–704 (2008). 5, 7, 14, 25

V Govindasamy, *VS Ronald*, *AN Abdullah*, *KR Ganesan Nathan*, *ZAC Ab Aziz*, *M Abdullah*, *S Musa*, *NH Abu Kasim* and *RR Bhonde*. 'Differentiation of dental pulp stem cells into islet-like aggregates'. *Journal of dental research*, 90(5):646–652 (2011). 2, 9, 43, 73

DS Gray, *Joe Tien* and *CS Chen*. 'Repositioning of cells by mechanotaxis on surfaces with micropatterned Young's modulus'. *J. Biomed. Mater. Res.*, 66:605–614 (2003). 4, 5, 6, 14, 24

M. Griebel, *G. Zumbusch*, *S. Knapek* and *A. Caglar*. *Numerische Simulation in der Molekül-dynamik*. Springer-Verlag (2004). ISBN 3-540-41856-3. 55, 56, 57

Volker Grimm, *Eloy Revilla*, *Uta Berger*, *Florian Jeltsch*, *Wolf M Mooij*, *Steven F Railsback*, *Hans-Hermann Thulke*, *Jacob Weiner*, *Thorsten Wiegand* and *Donald L DeAngelis*. 'Pattern-oriented modeling of agent-based complex systems: lessons from ecology'. *science*, 310(5750):987–991 (2005). 54

Anita Häcker. 'A mathematical model for mesenchymal and chemosensitive cell dynamics'. *J. Math. Biol.*, 64:361–401 (2012). 7, 25

William D Hamilton. 'The genetical evolution of social behaviour. I'. *Journal of theoretical biology*, 7(1):1–16 (1964a). 3

William D Hamilton. 'The genetical evolution of social behaviour. II'. *Journal of theoretical biology*, 7(1):17–51 (1964b). 3

Anandwardhan A Hardikar, *Bernice Marcus-Samuels*, *Elizabeth Geras-Raaka*, *Bruce M Raaka* and *Marvin C Gershengorn*. 'Human pancreatic precursor cells secrete FGF2 to

stimulate clustering into hormone-expressing islet-like cell aggregates'. *Proceedings of the National Academy of Sciences*, 100(12):7117–7122 (2003). 2, 9, 43, 73

Inbal Hecht, *Monica L Skoge*, *Pascale G Charest*, *Eshel Ben-Jacob*, *Richard a Firtel*, *William F Loomis*, *Herbert Levine* and *Wouter-Jan Rappel*. 'Activated membrane patches guide chemotactic cell motility.' *PLoS computational biology*, 7(6):e1002044 (2011). 22, 23, 39

LA Hohl and *WV Cruess*. 'Observations on certain film forming yeasts'. *Zentr. Bakteriol. Parasitenk., Abt. II*, 101:65–78 (1939). 11

William R. Holmes and *Leah Edelstein-Keshet*. 'A comparison of Computational Models for Eukaryotic Cell Shape And Motility'. *PLoS computational biology*, 8(12) (2012). 6, 25

Anna Huttenlocher and *Alan Rick Horwitz*. 'Integrins in cell migration.' *Cold Spring Harbor perspectives in biology*, 3(9):a005074 (2011). 5, 6, 14, 24

Richard O Hynes. 'Integrins: a family of cell surface receptors'. *Cell*, 48(4):549–554 (1987). 10, 43

Richard O Hynes. 'Integrins: versatility, modulation, and signaling in cell adhesion'. *Cell*, 69(1):11–25 (1992). 10, 43

Eric Jones, *Travis E. Oliphant*, *Pearu Peterson et al.* 'SciPy: Open source scientific tools for Python' (2001). 81

Kinneret Keren and *Julie A. Theriot*. 'Biophysical Aspects of Actin-Based Cell Motility in Fish Epithelial Keratocytes'. In *Peter Lenz*, editor, 'Cell Motility', (2008). 5, 6, 24

AJ Koch and *Hans Meinhardt*. 'Biological pattern formation: from basic mechanisms to complex structures'. *Reviews of modern physics*, 66(4):1481 (1994). 9, 42

Shigeru Kondo. 'The reaction-diffusion system: a mechanism for autonomous pattern formation in the animal skin'. *Genes to Cells*, 7(6):535–541 (2002). 9, 42

Shigeru Kondo and *Takashi Miura*. 'Reaction-diffusion model as a framework for understanding biological pattern formation'. *Science*, 329(5999):1616–1620 (2010). 8, 41

D a Lauffenburger and *a F Horwitz*. 'Cell migration: a physically integrated molecular process.' *Cell*, 84(3):359–69 (1996). 6, 7, 24, 25, 35

A. R. Leach. *Molecular Modelling: Principles and Applications*. Pearson Education, second edition Auflage (2001). ISBN 0-582-38210-6. 55

Jean-Luc Legras, *Didier Merdinoglu*, *Jean Cornuet* and *Francis Karst*. 'Bread, beer and wine: Saccharomyces cerevisiae diversity reflects human history'. *Molecular ecology*, 16(10):2091–2102 (2007). 11

Matthias Leinweber, *Patrick Bitter*, *Stefan Brueckner*, *Hans-Ulrich Moesch*, *Peter Lenz* and *Bernd Freisleben*. 'GPU-Based Simulation of Yeast Cell Flocculation'. In 'Parallel, Distributed and Network-Based Processing (PDP), 2014 22nd Euromicro International Conference on', Seiten 601–608. IEEE (2014). 2, 55, 69

Peter Lenz, editor. *Cell motility*. Springer (2008). 2

Herbert Levine and *Wouter-Jan Rappel*. 'Directed Motility and Dictyostelium Aggregation'. In *Peter Lenz*, editor, 'Cell Motility', (2008). 4, 14

Herbert Levine and *Wouter-Jan Rappel*. 'The physics of eukaryotic chemotaxis.' *Physics today*, 66(2) (2013). 20, 22, 23, 125

Tomas Linder and *Claes M Gustafsson*. 'Molecular phylogenetics of ascomycotal adhesins—a novel family of putative cell-surface adhesive proteins in fission yeasts'. *Fungal genetics and biology*, 45(4):485–497 (2008). 12, 44

Chenli Liu, *Xiongfei Fu*, *Lizhong Liu*, *Xiaojing Ren*, *Carlos KL Chau*, *Sihong Li*, *Lu Xiang*, *Hualing Zeng*, *Guanhua Chen*, *Lei-Han Tang et al.* 'Sequential establishment of stripe patterns in an expanding cell population'. *Science*, 334(6053):238–241 (2011). 8, 42

C M Lo, *H B Wang*, *M Dembo* and *Y L Wang*. 'Cell movement is guided by the rigidity of the substrate.' *Biophysical journal*, 79(1):144–52 (2000). 4, 6, 7, 14, 24, 25, 26, 33, 39, 118

Wan-Sheng Lo and *AM Dranginis*. 'FLO11, a yeast gene related to the STA genes, encodes a novel cell surface flocculin.' *Journal of bacteriology*, 178(24):7144–7151 (1996). 3, 12

Joe Lutkenhaus. 'Assembly dynamics of the bacterial MinCDE system and spatial regulation of the Z ring'. *Annu. Rev. Biochem.*, 76:539–562 (2007). 8, 41

RM Macnab and *DE Koshland*. 'The gradient-sensing mechanism in bacterial chemotaxis'. *Proceedings of the National Academy of Sciences of the United States of America*, 69(9):2509–2512 (1972). 4, 14, 17, 18, 39

I V Maly and *G G Borisy*. 'Self-organization of a propulsive actin network as an evolutionary process.' *Proceedings of the National Academy of Sciences of the United States of America*, 98(20):11324–9 (2001). 34

Benoit B. Mandelbrot. Fractals: Form, Chance and Dimension. W. H. Freeman and Company, 1 Auflage (1977). 45, 114, 123

Gail R Martin and *Martin J Evans*. 'Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies in vitro'. *Proceedings of the National Academy of Sciences*, 72(4):1441–1445 (1975). 9, 43, 73

P Martinez, *AC Codón*, *L Pérez* and *T Benitez*. 'Physiological and molecular characterization of flor yeasts: polymorphism of flor yeast populations'. *Yeast*, 11(14):1399–1411 (1995). 11

DA Melton. 'Pattern formation during animal development'. *Science*, 252(5003):234–241 (1991). 2, 8, 42

A Mogilner. 'Mathematics of cell motility: have we got its number?' *Journal of mathematical biology*, 58:105–134 (2009). 5, 6, 7, 24, 25

A Mogilner and *B Rubinstein*. 'The physics of Filopodial Protrusion'. *Biophysical journal*, 89:782–795 (2005). 35

Carey D Nadell, *Joao B Xavier* and *Kevin R Foster*. 'The sociobiology of biofilms'. *FEMS microbiology reviews*, 33(1):206–224 (2009). 1

J Thomas Parsons, *Alan Rick Horwitz* and *Martin a Schwartz*. 'Cell adhesion: integrating cytoskeletal dynamics and cellular tension.' *Nature reviews. Molecular cell biology*, 11(9):633–43 (2010). 4, 5, 6, 14, 24, 35

Amit Pathak and *Sanjay Kumar*. 'From molecular signal activation to locomotion: an integrated, multiscale analysis of cell motility on defined matrices.' *PloS one*, 6(3):e18423 (2011). 6, 24, 25, 73

Matjaž Perc, *Jesús Gómez-Gardeñes*, *Attila Szolnoki*, *Luis M Floría* and *Yamir Moreno*. 'Evolutionary dynamics of group interactions on structured populations: a review'. *Journal of The Royal Society Interface*, 10(80):20120997 (2013). 1

Thomas D Pollard and *Gary G Borisy*. *Cell*, 112(4):453–465 (2003). 5, 20

Jacques Prost, *Jean-Francois Joanny*, *Peter Lenz* and *Cecile Syhes*. 'The physics of Listeria Propulsion'. In *Peter Lenz*, editor, 'Cell Motility', (2008). 4, 14

Paolo P Provenzano, *David R Inman*, *Kevin W Eliceiri*, *Steven M Trier* and *Patricia J Keely*. 'Contact guidance mediated three-dimensional cell migration is regulated by Rho/ROCK-dependent matrix reorganization.' *Biophysical journal*, 95(11):5374–84 (2008). 6, 24, 25

E. M. Purcell. 'Life at low Reynolds number'. *American Journal of Physics*, 45(1):3–11 (1977). 4, 5, 13, 16, 18

David C Queller, *Eleonora Ponte*, *Salvatore Bozzaro* and *Joan E Strassmann*. 'Single-gene greenbeard effects in the social amoeba Dictyostelium discoideum'. *Science*, 299(5603):105–106 (2003). 3

Drazen Raucher and *Michael P. Sheetz*. 'Characteristics of a Membrane Reservoir Buffering Membrane Tension'. *Biophysical journal*, 77:1992–2002 (1999). 33

Todd B Reynolds and *Gerald R Fink*. 'Bakers' yeast, a model for fungal biofilm formation'. *Science*, 291(5505):878–881 (2001). 11

Anne J Ridley, *Martin a Schwartz*, *Keith Burridge*, *Richard a Firtel*, *Mark H Ginsberg*, *Gary Borisy*, *J Thomas Parsons* and *Alan Rick Horwitz*. 'Cell migration: integrating signals from front to back.' *Science (New York, N.Y.)*, 302(5651):1704–9 (2003). 6, 24, 35

Ida Rishal, *Naaman Kam*, *Rotem Ben-Tov Perry*, *Vera Shinder*, *Elizabeth Fisher*, *Giampietro Schiavo* and *Mike Fainzilber*. 'A motor-driven mechanism for cell-length sensing'. *Cell reports*, 1(6):608–616 (2012). 60

M. Roderfeld, *S. Matern* and *E. Roeb*. 'Konfokale Laserscanning-Mikroskopie: Der Blick in die Zelle'. *Dtsch Med Wochenschr*, 128:2539–2542 (2003). ISSN 0012-0472. 93

B Rubinstein, *K Jacobson* and *A Mogilner*. 'Multiscale two-dimensional modeling of a motile simple-shaped cell'. *Multiscale Modeling & Simulation*, 3(2):413–439 (2005). 7, 25

Joseph Schlessinger, *Yoram Shechter*, *Mark C Willingham* and *Ira Pastan*. 'Direct visualization of binding, aggregation, and internalization of insulin and epidermal growth factor on living fibroblastic cells'. *Proceedings of the National Academy of Sciences*, 75(6):2659–2663 (1978). 9, 43

Robert S Sikorski and *Philip Hieter*. 'A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae.' *Genetics*, 122(1):19–27 (1989). 92

J Victor Small and *Guenter P Resch*. 'The comings and goings of actin: coupling protrusion and retraction in cell motility'. *Current opinion in cell biology*, 17(5):517–523 (2005). 5, 20

Alexander E Smith, *Zhibing Zhang*, *Colin R Thomas*, *Kennith E Moxham* and *Anton PJ Middelberg*. 'The mechanical properties of Saccharomyces cerevisiae'. *Proceedings of the National Academy of Sciences*, 97(18):9871–9874 (2000). 66

Scott Smukalla, *Marina Caldara*, *Nathalie Pochet*, *Anne Beauvais*, *Stephanie Guadagnini*, *Chen Yan*, *Marcelo D Vinces*, *An Jansen*, *Marie Christine Prevost*, *Jean-Paul Latgé et al.* 'FLO1 Is a Variable Green Beard Gene that Drives Biofilm-like Cooperation in Budding Yeast'. *Cell*, 135(4):726–737 (2008). 90, 104, 105, 114, 123

Eduardo V Soares. 'Flocculation in Saccharomyces cerevisiae: a review'. *Journal of applied microbiology*, 110(1):1–18 (2011). 12, 43, 114, 123

Joan E Strassmann, *Yong Zhu* and *David C Queller*. 'Altruism and social cheating in the social amoeba Dictyostelium discoideum'. *Nature*, 408(6815):965–967 (2000). 3

Tatyana M. Svitkina, *Elena A. Bulanova*, *Oleg Y. Chaga*, *Danijela M. Vignjevic*, *Shin-ichiro Kojima*, *Jury M. Vasiliev* and *Gary G. Borisy*. 'Mechanism of filopodia initiation by reorganization of a dendritic network'. *The Journal of cell biology*, 160(3):409–421 (2003). 34

Robert Tarjan. 'Depth-first search and linear graph algorithms'. *SIAM journal on computing*, 1(2):146–160 (1972). 52

Martin Thanbichler and *Lucy Shapiro*. 'MipZ, a Spatial Regulator Coordinating Chromosome Segregation with Cell Division in< i> Caulobacter'. *Cell*, 126(1):147–162 (2006). 8, 41

Eric Theveneau, *Lorena Marchant*, *Sei Kuriyama*, *Mazhar Gull*, *Barbara Moepps*, *Maddy Parsons* and *Roberto Mayor*. 'Collective Chemotaxis Requires Contact-Dependent Cell Polarity'. *Developmental Cell*, 19:39–53 (2010). 6, 24, 26, 118

Bryan C Thorne, *Alexander M Bailey* and *Shayn M Peirce*. 'Combining experiments with multi-cell agent-based modeling to study biological tissue patterning'. *Briefings in bioinformatics*, 8(4):245–257 (2007). 54

M J Tindall, *P K Maini*, *S L Porter* and *J P Armitage*. 'Overview of mathematical approaches used to model bacterial chemotaxis II: bacterial populations.' *Bulletin of mathematical biology*, 70(6):1570–607 (2008a). 4, 14, 18

M J Tindall, *S L Porter*, *P K Maini*, *G Gaglia* and *J P Armitage*. 'Overview of mathematical approaches used to model bacterial chemotaxis I: the single cell.' *Bulletin of mathematical biology*, 70(6):1525–69 (2008b). 18

George J Todaro and *Howard Green*. 'Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines'. *J. Cell Biol.*, 17(2):299–313 (1963). 9, 43

Christopher P Toret, *Michael V D'Ambrosio*, *Ronald D Vale*, *Michael A Simon* and *W James Nelson*. 'A genome-wide screen identifies conserved protein hubs required for cadherin-mediated cell–cell adhesion'. *The Journal of cell biology*, 204(2):265–279 (2014). 10, 43

Léa Trichet, *Jimmy Le Digabel*, *Rhoda J Hawkins*, *Sri Ram Krishna Vedula*, *Mukund Gupta*, *Claire Ribrault*, *Pascal Hersen*, *Raphaël Voituriez* and *Benoît Ladoux*. 'Evidence of a large-

scale mechanosensing mechanism for cellular adaptation to substrate stiffness.' *Proceedings of the National Academy of Sciences of the United States of America*, 109(18):6933–8 (2012). 6, 25

Yuhai Tu, *Thomas S Shimizu* and *Howard C. Berg*. 'Modeling the chemotactic response of Escherichia coli to time-varying stimuli'. *Proceedings of the National Academy of Sciences of the United States of America*, 105(39):14855–14860 (2008). 18

Alan Mathison Turing. 'The chemical basis of morphogenesis'. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 237(641):37–72 (1952). 8, 41

Sarah L. Veatch, *Benjamin B. Machta*, *Sarah A. Shelby*, *Ethan N. Chiang*, *David A. Holowka* and *Barbara A. Baird*. 'Correlation Functions Quantify Super-Resolution Images and Estimate Apparent Clustering Due to Over-Counting'. *PLoS ONE*, 7(2):e31457 (2012). 49

Maik Veelders, *Stefan Brückner*, *Dimitri Ott*, *Carlo Unverzagt*, *Hans-Ulrich Mösch* and *Lars-Oliver Essen*. 'Structural basis of flocculin-mediated social behavior in yeast'. *Proceedings of the National Academy of Sciences*, 107(52):22511–22516 (2010). 12

Kevin J Verstrepen and *Frans M Klis*. 'Flocculation, adhesion and biofilm formation in yeasts'. *Molecular microbiology*, 60(1):5–15 (2006). 12, 44

Kevin J Verstrepen, *Todd B Reynolds* and *Gerald R Fink*. 'Origins of variation in the fungal cell surface'. *Nature Reviews Microbiology*, 2(7):533–540 (2004). 12, 44

Tamas Vicsek. 'Pattern formation in diffusion-limited aggregation'. *Phys. Rev. Lett.*, 53(24):2281 (1984). 9, 42

David Sloan Wilson and *Edward O Wilson*. 'Rethinking the theoretical foundation of sociobiology'. *The Quarterly review of biology*, 82(4):327–348 (2007). 1

Fred Winston, *Catherine Dollard* and *Stephanie L Ricupero-Hovasse*. *Yeast*, 11(1):53–55 (1995). 91, 92

T A Witten and *L M Sander*. *Phys. Rev. B*, 27(9):5686 (1983). 9, 42

TA Witten Jr. and *Leonard M Sander*. 'Diffusion-limited aggregation, a kinetic critical phenomenon'. *Physical review letters*, 47(19):1400 (1981). 9, 42, 79, 80, 87, 120

Lewis Wolpert. 'Positional information and the spatial pattern of cellular differentiation'. *Journal of theoretical biology*, 25(1):1–47 (1969). 8, 41, 73

Joyce Y. Wong, *Alan Velasco*, *Padmavathy Rajagopalan* and *Quynh Pham*. 'Directed Movement of Vascular Smooth Muscle Cells on Gradient-Compliant Hydrogels †'. *Langmuir*, 19(5):1908–1913 (2003). 5, 14

Tony Yeung, *Penelope C Georges*, *Lisa a Flanagan*, *Beatrice Marg*, *Miguelina Ortiz*, *Makoto Funaki*, *Nastaran Zahir*, *Wenyu Ming*, *Valerie Weaver* and *Paul a Janmey*. 'Effects of substrate stiffness on cell morphology, cytoskeletal structure, and adhesion.' *Cell motility and the cytoskeleton*, 60(1):24–34 (2005). 6, 24

John C Zak, *Michael R Willig*, *Daryl L Moorhead* and *Howard G Wildman*. 'Functional diversity of microbial communities: a quantitative approach'. *Soil Biology and Biochemistry*, 26(9):1101–1108 (1994). 3

Muhammad H Zaman, *Roger D Kamm*, *Paul Matsudaira* and *Douglas a Lauffenburger*. 'Computational model for cell migration in three-dimensional matrices.' *Biophysical journal*, 89(2):1389–97 (2005). 6, 7, 24, 25

Severino Zara, *Alan T Bakalinsky*, *Giacomo Zara*, *Giorgia Pirino*, *Maria Antonietta Demontis* and *Marilena Budroni*. 'FLO11-based model for air-liquid interfacial biofilm formation by

Saccharomyces cerevisiae'. *Applied and Environmental microbiology*, 71(6):2934–2939 (2005). 11

Ciyou Zhu, *Richard H Byrd*, *Peihuang Lu* and *Jorge Nocedal*. 'Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization'. *ACM Transactions on Mathematical Software (TOMS)*, 23(4):550–560 (1997). 81

Daniel Zicha, *Graham A. Dunn* and *Alastair F. Brown*. 'A new direct-viewing chemotaxis chamber.' *Journal of cell science*, 99:769–75 (1991). 40, 118

Danksagung

Zunächst möchte ich mich bei meinem Betreuer Prof. Dr. Peter Lenz bedanken, der immer ein offenes Ohr und einen guten Rat für mich hatte. Außerdem bei Prof. Dr. Hans-Ullrich Mösch, meinem Zweitgutachter, bei dem ich einige Zeit im Labor verbringen durfte.

Auch bei meinen Kooperationspartnern aus Marburg, Stefan Brückner und Matthias Leinweber möchte ich mich herzlich bedanken. Stefan hat mich bei meinen ersten Schritten im Labor angeleitet und mich bei den Experimenten unterstützt. Matthias hat meinen Algorithmus sehr erfolgreich auf die Grafikkarte portiert und sehr zu seiner Verbesserung beigetragen.

Einen herzlichen Dank auch an meine Kollegen aus Hongkong, Huang Wei, Huang Jiandong und Hu Jianjiang, die mich bei meinen Arbeiten an der University of Hong Kong angeleitet und unterstützt haben.

Ich bedanke mich bei den Korrekturlesern Kristof Beck und Alexander Orlov, die mich vor einigen Peinlichkeiten bewahrt haben.

Des Weiteren möchte ich mich bei den Kollegen und Freunden bedanken, die mir durch Gespräche, Diskussionen und Hilfestellungen bei der Erstellung meiner Dissertation sehr geholfen haben: Kristof Beck, Alexander Orlov, Konstanze Bandmann Tobias Kreilos, Hannes

Brauckmann, Kevin Yam, Marcus Assmann, Shuxin Jia, Lisa Beran, Michael Grau, Stefan Zammert und Michael Rath.

Außerdem möchte ich meinen Freunden und Kollegen aus den Arbeitsgruppen Biologische Physik, Komplexe Systeme, Verteilte Systeme, Molekulare Genetik und aus SYNMIKRO für die angenehme Arbeitsatmosphäre und die schöne Zeit Danken.

Insbesondere möchte ich mich noch bei meinen Eltern bedanken, ohne die mir ein Studium niemals möglich gewesen wäre. Ihre Unterstützung, gerade in den schwierigen Phasen meines Studiums und meiner Promotion, hat mich immer wieder weiter gebracht.

Zu guter Letzt danke ich meiner Frau für ihre Geduld, insbesondere in den letzten Tagen und Wochen vor der Abgabe, und ihre uneingeschränkte moralische Unterstützung.

Danke!

Erklärungen

Versicherung

Die vorliegende Dissertation wurde von mir selbst und ohne fremde Hilfe verfasst. Es wurden keine anderen als die in ihr angegebenen Quellen oder Hilfsmittel benutzt. Alle vollständigen oder sinngemäß übernommenen Zitate wurden als solche gekennzeichnet und die Dissertation wurde in der vorliegenden oder einer ähnlichen Form noch bei keiner anderen in- oder ausländischen Hochschule anlässlich eines Promotionsgesuchs oder zu anderen Prüfungszwecken eingereicht.

Marburg, 17.12.2014

Patrick Bitter

Erklärung

Ich habe vor diesem Promotionsversuch keinen weiteren Versuch unternommen.

Marburg, 17.12.2014

Patrick Bitter

Lebenslauf

Patrick Bitter
Geboren am 19.03.1986 in Arolsen (jetzt Bad Arolsen)

Akademischer Werdegang:

1996 – 2005   Christian-Rauch-Schule Bad Arolsen
Abschluss: Abitur

2005-2011   Studium der Physik an der Philipps-Universität Marburg
Abschluss: Diplom-Physiker

2011-2015   Promotion in theoretischer Biophysik an der Philipps-Universität Marburg