Universiteit Leiden
The Netherlands

Insights from modeling metabolism and amoeboid cell motility in the immune system

Steijn, L. van

Citation

Steijn, L. van. (2021, July 15). *Insights from modeling metabolism and amoeboid cell motility in the immune system*. Retrieved from https://hdl.handle.net/1887/3195085

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3195085

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

Leiden University Repository

The handle https://hdl.handle.net/1887/3195085 holds various files of this Leiden University dissertation.

Author: Steijn, L. van
Title: Insights from modeling metabolism and amoeboid cell motility in the immune system
Issue Date: 2021-07-15

Insights from modeling metabolism and amoeboid cell motility in the immune system

Leonie van Steijn

Copyright: Leonie van Steijn, 2021
ISBN: 978-94-6332-766-4
Print: GVO drukkers & vormgevers B.V.
Omslagontwerp: Leonie van Steijn, Vera Lippai, Remie Janssen

Insights from modeling metabolism and amoeboid cell motility in the immune system

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 15 juli 2021
klokke 13.45 uur door

Leonie van Steijn

geboren te Noordwijkerhout
in 1991

Promotores

Prof. dr. R.M.H. Merks
Prof. dr. H.P. Spaink
Prof. dr. ir. F.J. Verbeek

Promotiecommissie

Prof. dr. F.A. van der Duijn Schouten	Voorzitter
Prof. dr. A. Doelman	Secretaris
Prof. dr. A.H. Meijer	
Prof. dr. E.C.M. de Lange	
Dr. J.C. Textor	Radboud Universiteit, Nijmegen
Dr. L. Dupré	Toulouse Institute for Infectious and Inflammatory Diseases, Toulouse, Frankrijk

Contents

1 Introduction **1**
1.1 Cell migration and metabolism: linked by infection 1
1.2 Metabolism 2
1.3 Cell migration 7
1.4 Thesis overview 14

2 Predicting metabolism in a metabolic model of *Danio rerio* **17**
2.1 Introduction 18
2.2 Methods 21
2.3 Results 27
2.4 Discussion 44
2.5 Supplementary material 48

3 Cell-matrix adhesion affects cell motility mode **49**
3.1 Introduction 50
3.2 Results 53
3.3 Discussion 68
3.4 Methods 73
3.5 Supporting information 79

4 Topotaxis on silicon and *in silico* **81**
4.1 Introduction 83
4.2 Results 85
4.3 Discussion 104
4.4 Materials, Methods, and Model 109

5 Function of TLR2 and MyD88 in leukocyte migration **121**
5.1 Introduction 122
5.2 Materials and methods 124
5.3 Results 128
5.4 Discussion 150

6 Discussion **155**
6.1 Summarizing discussion 155
6.2 Future work 157

Bibliography **163**

Samenvatting **183**

Summary **187**

Dankwoord **189**

Curriculum Vitae **191**

Chapter 1

Introduction

Multicellular organisms, including plants and animals, continuously struggle with infections by bacteria, fungi and viruses. In animals, infections are cleared by the innate and acquired immune system. Despite the successes of biomedical research in fighting infectious diseases, many disease processes are still little understood. In this thesis, we will focus on two processes involved in fighting infections: metabolism, and immune cell motility and navigation. Specifically, we use mathematical and computational models to address questions surrounding these processes. In this chapter, we first explain the link between metabolism and immune cell motility, before we introduce both the main questions and modeling methods for these subjects.

1.1 Cell migration and metabolism: linked by infection

During infection, metabolism shifts from maintenance of the body to fighting the infection. This happens at a single cell scale as well as on tissue or organism scale. Examples of common global changes are the production of heat in inflamed tissues or fever on the scale of the entire body [1]. Some infectious diseases, for example tuberculosis, are associated with wasting, a rapid and involuntary loss of muscle and fat tissue and a still little understood metabolic change. As these tissues can serve as energy storage, the breakdown of these tissues is paired with the release of a large amount of energy. It is yet unclear which processes would require such an amount of energy and which changes in metabolism allow for this. A potential process that is under debate is whether immune cells in tuberculosis lesion exhibit the Warburg effect, which is the rapid conversion of glucose into lactate, despite the availability of oxygen to fully oxidize the glucose [2]. Furthermore, it is unclear how wasting is induced: by the pathogen/infection or by the host [3]?

Local changes in metabolism can be seen in many types of immune cells and occur at different stages of infection clearance; immune cells, such as leukocytes or cytotoxic T lymphocytes, first have to locate pathogens and then neutralize them. A number of different metabolic demands are the production of ATP to fuel cell migration, the production of reactive oxygen species and other compounds that destroy the pathogen, and cytokine production to recruit more cells to infected areas [4, 5].

However, some metabolic changes in immune cells will not occur unless the immune system finds the pathogen within the body. Detection of pathogens is thought to happen by random encounters of immune cell and pathogen. Navigation through a tissue is greatly influenced by the type of motility immune cells display and determines the occurrence of such random encounters [6, 7]. How immune cells such as lymphocytes and leukocytes navigate through a tissue is a topic of ongoing research. The extracellular matrix (ECM), the network surrounding cells within tissues, is thought to play an important role [8]. As a framework around cells in tissues, it can function as a scaffold for cells to push off from, as well as a physical barrier blocking cell movement. Besides that, the ECM also plays a role in signaling, both by binding signaling molecules and by its composition and stiffness. By modeling immune cell motility, we can untangle the roles of the ECM and study the effect of the ECM as a substrate, an obstacle or its signaling function separately.

1.2 Metabolism

In studying metabolism, there are two main questions: which metabolites are present and at what rate are they being converted into one another. Considering the changes in metabolism during infection, and the wide range of involved processes, we should not be limited to a small set of metabolites. Metabolomics can study a wide array of molecules using techniques such as NMR and mass-spectrometry, and thus answer the first question. Combining metabolomics data with other data and methods can be used to study reactions rates, so called systems metabolomics [10]. A rough indication of the present enzymes and, by proxy, fluxes, can be obtained through gene expression and proteomic data [11]. Detailed flux information can be derived from metabolomics combined with isotope labeling. However, this requires good prior knowledge of the metabolic network [12]; by extensive calculation of fluxes and the corresponding isotope distributions along the network, one can match fluxes to the measured

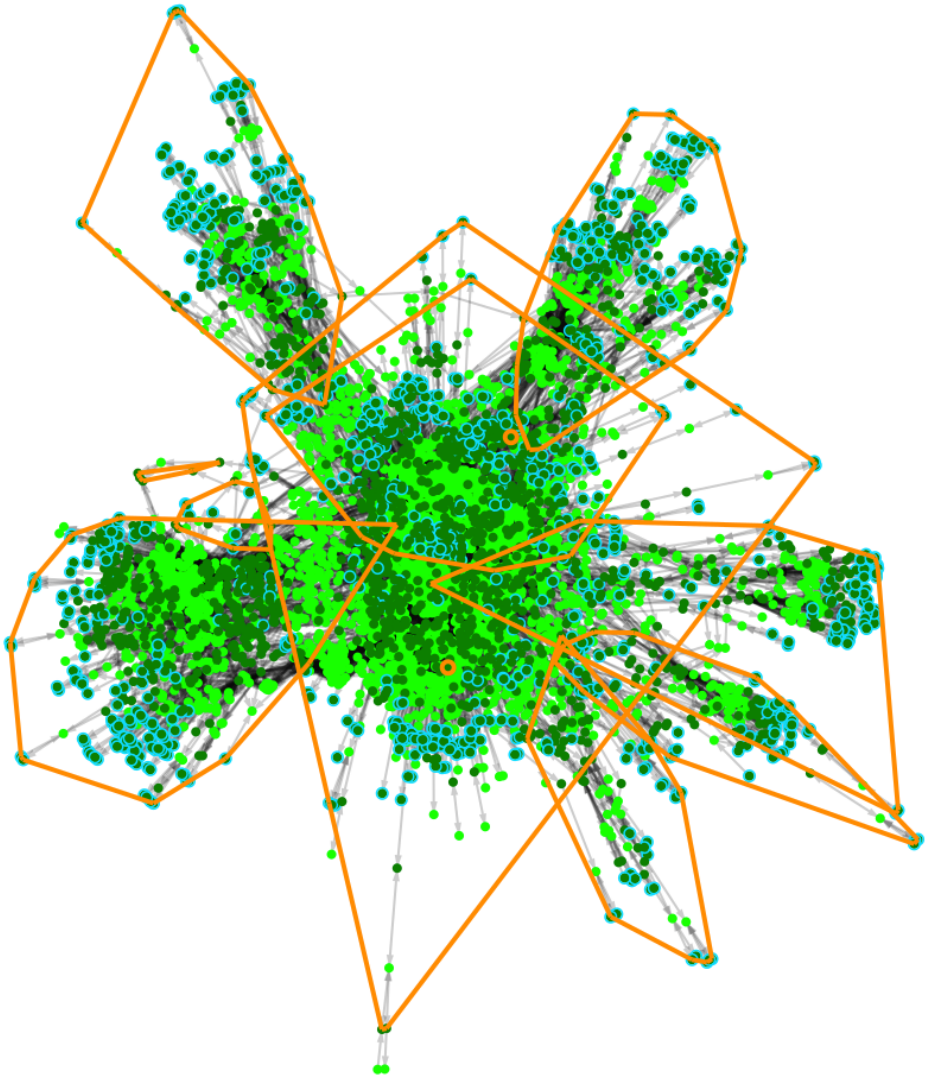


FIGURE 1.1: **Visualization of the metabolic network from ZebraGEM 2.0.** Light green nodes: reactions. Dark green nodes: metabolites (with cyan lining: end point metabolites). Orange boxes: cellular compartments. Visualization made using ModelExplorer [9].

isotope labeled metabolomics data [13]. This prior knowledge, collected into a metabolic model, can be used for other modeling efforts as well.

Aside from being a computational tool in isotope labeling, metabolic models can be used as a predictive tool of how internal and external changes alter an organism's metabolism. Current models of metabolism exist varying in the level of details. Kinetic models are systems of coupled ordinary differential equations that describe the change of metabolites over time. Often, they are limited to single pathways for which they can give detailed time-dependent predictions of metabolite concentrations and metabolite flows [14]. They come, however, at the cost of large numbers of parameters. These parameters can be obtained through experiments, and a kinetic *S. cerevisiae* metabolism model [15] and, more recently, a kinetic model of *E. coli* core metabolism [16] have been established. Nonetheless, this requires extensive studies, as kinetics are influenced by substrate concentration, temperature, pH, and many other factors. Another angle is taken in constraint-based metabolic models, or stoichiometric models. As the stoichiometry of many metabolic pathways is known, the metabolic network can be represented by a stoichiometric matrix. Constraint-based metabolic models explore the properties of the network, and how the network itself and physical and chemical properties constrain fluxes through that network [17]. They can predict the most efficient metabolic pathway to create certain metabolites or alternative pathways in case of disruption by blocking or limiting certain reactions.

The most commonly used method with constraint-based modeling is Flux Balance Analysis (FBA) [18]. The main assumption of this method is a steady state of internal metabolites; hence, flux balance. As the concentration of metabolites, c, is governed by $\frac{dc}{dt} = S \cdot f$, with S the stoichiometric matrix corresponding to the metabolic network and f the fluxes through that network, this assumption can be written as $\frac{dc}{dt} = S \cdot f = 0$. This gives a system of linear equations. However, usually this system of equations is underdetermined, resulting in a multidimensional solution space with an infinite number of solutions (Figure 1.2). This brings us to another important assumption: the organism optimizes metabolism for a certain metabolic objective. This metabolic objective depends on the research question, but frequent objectives are maximization of biomass production (i.e. growth), maximization of energy production (i.e ATP production), or maximal production of a specific desirable compound like a drug compound. In addition, metabolites flowing in and out of the system are limited, resulting in additional constraints. With these two assumptions and the constraints, the model then gives the following linear programming problem (LP) that can be solved computationally:

Optimize

$$f_{obj}$$

such that:

$$S \cdot f = 0,$$
$$a_i \leq f_i \leq b_i$$

The solution of that LP problem is a set of reaction fluxes that ascertain steady state and optimize the objective.

The main limitation of FBA is that it predicts fluxes in equilibrium with available nutrients. However, in most natural systems, nutrient availability varies over time. Extending the FBA with a simple system of ordinary differential equations, one can do dynamic FBA [19], and account for the fluctuations of nutrients and growing organisms, described by the following equations:

$$\frac{dc}{dt} = S \cdot fB, \quad \frac{dB}{dt} = f_B B, \tag{1.1}$$

where B denotes the biomass of the organism and f_B is the growth rate per unit biomass obtained from computing the fluxes. One example of where dynamic FBA is applied, is the modeling of small populations of bacteria. Although dynamic FBA is a dynamic model, the time scale is coarser than for kinetic metabolic models, hence allowing for the use of FBA as an intermediary step. Another limitation of FBA is that it gives only a single solution from the usually large solution space. The outlines of this solution space can be explored by using flux variability analysis (FVA, Figure 1.2). Here, aside from optimizing the objective, each single reaction in the network is minimized and maximized [20], which draws the contours of the solution space.

Lastly, for many organisms, and especially multicellular ones, data on nutrient availability is hard to obtain or even define, and finding the right constraints is impossible. However, other data can be collected much easier, such as gene expression data, or proteomics and metabolomics data. These data can be integrated into the constraints-based modeling framework as additional constraints. Multiple methods on how to incorporate experimental data in constraints-based modeling have been developed, depending on the specific type of data; gene expression profiles [21, 22], proteomics data [23, 24] and metabolomics data [23, 25, 22, 26].

Essential to the integration of gene expression data is that the genes are represented in the model. Most of these constraint-based models, especially

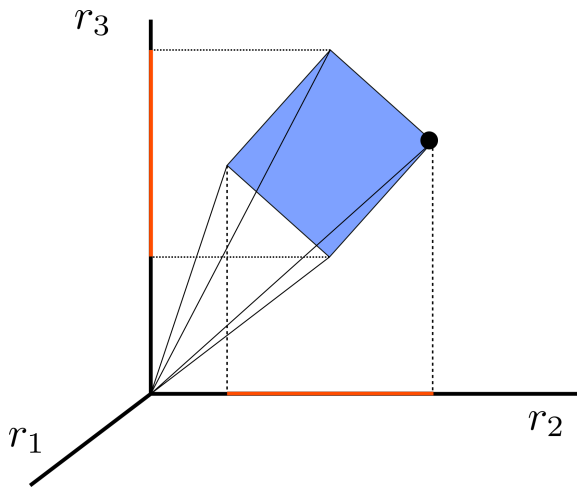


FIGURE 1.2: Solution space of constraints based models. Constraints on the fluxes ($r_{1,2,3}$) result in a convex subspace that represents feasible flux distributions. Optimizing a certain objective, as in FBA, gives an optimal solution space (blue face) of which LP solvers give a single point (black point). By individually minimizing and maximizing each flux within the optimal solution space (FVA), one can draw the contours (red projections on r_2 and r_3 axis) of the optimal solution space.

in recent years, are so called whole-genome scale metabolic reconstructions and, hence, contain all metabolic genes, or at least all suspected metabolic genes, of an organism. The corresponding metabolic reactions then form the metabolic network. By including all metabolic reactions, these model encompass an organism's metabolic network. The relation between reactions and genes is described by gene-protein-reaction associations (GPR) [27]. Many organisms are already captured in such whole-genome scale metabolic reconstructions. Most of those organisms are unicellular organisms, such as a multitude of *E. coli* strains [28] and *S. cerevisiae* [29]. Multicellular organisms, and of special interest for studying immune response, vertebrates, are less well represented. To current date, whole-genome scale metabolic reconstruction of human [30], mouse [31], Chinese hamster [32], rat [33], zebrafish [34, 35], and cod [36] have been assembled. Although the metabolic diversity among vertebrates is limited in comparison to the zoo of reconstructions of unicellular organisms, organism specific models are still preferable. Essential nutrients differ among organisms, think of vitamin C that is essential for humans but not for most other vertebrates, and inositol, which zebrafish are unable to produce themselves [37]. Furthermore, genetic structure between vertebrates is sufficiently different to

justify separate reconstructions, as GPRs become dramatically different.

1.3 Cell migration

The metabolic demands of immune cells are diverse and change during the infection. Powering cell migration is an important demand from the beginning of infection. Immune cells have to be able to find the pathogen. Both how they manage to propel themselves and how they navigate through tissues play a role in their ability of finding the pathogen. Here, we shortly discuss how immune cells move, how to analyze cell motility and distinguish different types of motilities, and finally, how cell motility is modeled.

1.3.1 Amoeboid motion

Most eukaryotic cells move by reorganizing their internal structure, the cytoskeleton, and hence their shape. This type of motion is called amoeboid motion. Amoeboid cell motility is driven by actin polymerization (Figure 1.3). Actin filaments are polarized and polymerization occurs mainly at one end. Due to thermal fluctuations of the cell membrane, actin subunits can be added in the space between an existing filament and the membrane. When an actin subunit is added to the actin filament, its attached ATP is hydrolyzed. This releases energy and combined with the elongation of the actin filament, this results in a force that can be used to push the cell membrane forward [38].

The polymerizing actin is organized in a network, an actin front. The size and shape of this network can be seen in the deformation of the cell. Actin fronts pushing out the cell membrane form a flat structure called a lamellipodium. Lamellipodia vary in size and stability. Very broad and stable lamellipodia are seen in keratocytes and these cells often cover very straight trajectories [39]. Other cells, for example the cellular slime mold *Dictoystelium*, have multiple, forking lamellipodia which they sometimes retract, and hence, change direction more often than keratocytes [40].

With so many regulating compounds, cells can also swiftly react to external signaling. The best known example of this is chemotaxis, where cells migrate towards or away from a compound such as food, oxygen or signaling molecules. Membrane located receptors influence in migration direction by relaying their signals to the actin polymerization regulating factors [41] such as the Arp2/3 complex which influences branching of actin filaments [42], and the antagonistic small GTPases Rac1 and RhoA which determine cell polarization [43]. Another cue that relays to the actin machinery is durotaxis, where the cell responds to the stiffness of the

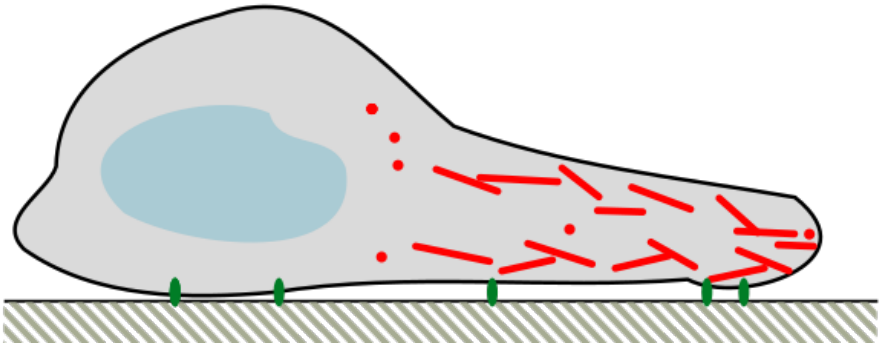


FIGURE 1.3: **Schematic view of actin polymerization within a lamellipodium** Red: actin filaments and actin monomers. Blue: cell nucleus. Green: cell-matrix adhesion complexes. Shaded gray: substrate with extracellular matrix. Actin filaments are polymerized at the front (right) of the cell. This generates a force that can push the cell membrane outwards. Actin filaments attached to the ECM can transfer the force more efficiently to outwards motion.

extracellular matrix through focal adhesions, protein complexes connecting the cytoskeleton to the ECM [44]. In this thesis, we argue that topotaxis, guidance by topographical cues in the environment, could also work by changing the actin polymerization, simply by putting obstacles in the way of cells.

1.3.2 Analysing immune cell motility

There are many determining characteristics in cell motility. Hence, describing cell motility by just a single characteristic is insufficient to fully capture the behavior of a motile cell. The most simple characteristic of cell motion is speed. Defining an instantaneous speed requires some considerations. Using a short time frame to define instantaneous speed can measure fluctuations of the cell membrane which do not contribute to actual cell displacement, and overestimates cell speed. On the other hand, measuring instantaneous speed with too large time steps disregards actual short excursions of the cell, and underestimates cell speed. Hence, choosing a time frame for computing the instantaneous velocity requires some consideration. For experimental data, there are additional constraints in the feasible spatial and time resolution.

The trajectory of cells can be described as a random walk: they regularly stop and move into a new direction. Random walk theory has brought forth

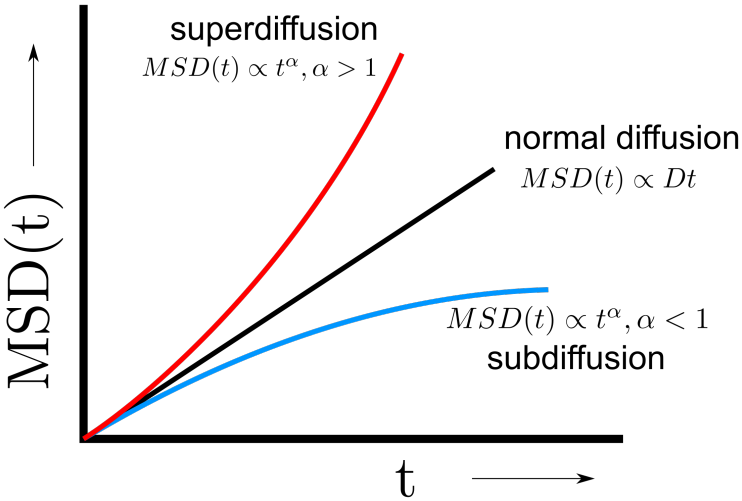


FIGURE 1.4: **Mean squared displacement curve of diffusive, superdiffusive and subdiffusive processes**

a number of characteristics and methods that we can apply to study cell motility. The first measure we discuss is the mean squared displacement (MSD). It is a measure of the deviation of the position of a random walker, particle or cell from a reference position over time, defined by:

$$MSD(t) = \langle |x(t) - x_0|^2 \rangle, \tag{1.2}$$

where the mean is taken either over a number of cells or over multiple time points. The MSD can be interpreted as the area explored by a population of cells in a given time and is closely related to the concept of diffusivity.

By computing the MSD of a large number of cell trajectories, information can be extracted from the curve of the MSD. If the MSD increases linearly over time, the observed cells are performing a Brownian random walk. Einstein and Sutherland derived independently that the MSD of Brownian walkers is given by

$$MSD(t) = 2nDt, \tag{1.3}$$

with n the dimension in which the walk takes place, and D the diffusion constant of the walker [45, 46]. Hence, from a linear MSD, we can straightforwardly compute the diffusivity of the cells.

Brownian walkers do not show any correlation in their step directions. However, some cells are able to maintain their direction for some amount

of time. This is called persistence. One can recognize persistent random walkers when the MSD curve shows a more than linear increase with respect to time, i.e. $MSD(t) \propto t^{\alpha}$ with $\alpha > 1$ (figure 1.4). This is called superdiffusion. The time scale at which $\alpha > 1$ is an interesting matter. For persistent random walkers, $\alpha > 1$ on short time scales, but on longer timescales the walker still perform a Brownian walk and $\alpha = 1$. The diffusivity is then no longer related to the slope of the MSD curve only. Instead, for persistent random walkers a different expression gives us the MSD, the so called Fürth's equation [47]:

$$MSD(t) = 2v_0^2\tau t - 2(v_0\tau)^2(1 - e^{-t/\tau}), \tag{1.4}$$

with v_0 an intrinsic cell velocity and τ the cell persistence time, i.e. the time frame in which the cell keeps moving in the same direction. The effective diffusivity of these walkers in two dimensions is given by $\frac{1}{2}v_0^2\tau$. So, the cell can explore more area by increasing its intrinsic velocity, or by increasing its persistence time. Random walks with $\alpha > 1$ on all time scales are called superdiffusive. A good example are Lévy walkers [48], which have fractal-like property.

A less than linear increase in MSD with time ($\alpha < 1$) is also possible and is called subdiffusion (Figure 1.4). Usually, subdiffusion is caused by some physical constraint on the cell, like limited space in which the cells can move around or forces keeping the cell in place.

Cell persistence can also be measured by the meandering index. This index is defined as the net displacement divided by the total displacement of a cell:

$$\frac{\|x_T - x_0\|}{\sum_{i=1}^{T}\|x_i - x_{i-1}\|},$$

with T the duration in time steps of the cell trajectory. A high index corresponds to straighter trajectories, and a low index to more convoluted trajectories. An important factor is the time resolution, similar to the computation of cell speed. Too low resolution will overestimate the index, and higher resolutions will most likely always result in a lower index.

Finally, we can define the 'directional bias' as the population mean velocity in a direction of interest. If the mean is equal to zero, there is no bias in that direction, and if the mean is significantly different from zero, there is. Directional bias can arise by factors such as chemotaxis and other forms of taxis.

1.3.3 Modeling amoeboid motion

In analysis of cell tracking data, a common question is the characterization of the type of cell motility observed in the data: which random walk recapitulates the same statistical properties as the cell tracking data [49, 50, 51, 52]? These types of analyses show that T cells display a Brownian walk, persistent random walk, Lévy walk or subdiffusive walk, depending on the tissue they reside in [53, 54, 55, 56, 57], but they give no insight in what mechanisms underlie those different types of motility.

Mechanistic models have focused on how the cellular components such as actin, myosin, Arp2/3 and other actin network regulating factors, focal adhesion complexes, and feedback from Rac and Rho signaling and phosphoinositide contribute to cell polarization and cell locomotion [58, 59, 60, 61, 62]. We cannot possibly describe all the work, and hence direct the reader to some reviews on this topic [63, 64].

In this thesis, we are interested in studying the effect of cell-ECM interactions on the type of cell motility using mathematical models. We have a number of requirements for our model: 1) we want a model that includes cell-ECM interactions on a subcellular scale, 2) the effects of cell-ECM interaction remain localized within the cell and hence we want our model to integrate intracellular processes and cell and cytoskeleton plasticity, and 3) the model should be coarse-grained enough that we can simulate sufficiently many cell tracks for statistical analysis. The combination of these three requirements further requires us to restrict to cell motility models that take the shape of the cell into account. These types of models can be found in the modeling framework of phase field models [65, 66, 67] and the Cellular Potts model [68]. Phase field models are a good tool in understanding the role of different components in cell motility such as adhesion, contractility and actin polymerization, yet, they are computationally quite expensive which makes them less suitable for the generation of cell trajectories. In this thesis, we use the Cellular Potts modeling framework, described in detail in Chapters 3 and 4, and shortly introduce it here.

1.3.4 Cellular Potts model

The Cellular Potts model, also known as the Glazier-Graner-Hogeweg model, is a lattice based model of deformable cells [68]. A single cell consists of a connected set of lattice sites having spin, or cell ID, denoted by σ. Cells move by both copying their σ into neighboring lattice sites or other σ's being copied into their lattice sites (Figure 1.5). Copy attempts are done iteratively by a Metropolis algorithm. The algorithm selects a random lattice site u and one of its neighboring lattice sites v. If $\sigma(u) \neq \sigma(v)$, $\sigma(v)$

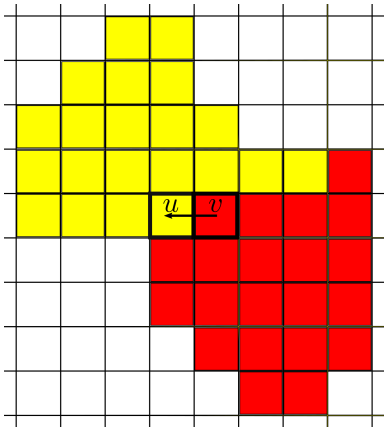


FIGURE 1.5: **Copy attempt in the Cellular Potts model** A configuration of the CPM with a yellow and red cell. A random lattice site u and one of its neighbors v are chosen. As $\sigma(u) \neq \sigma(v)$, lattice site u can obtain $\sigma(v)$ causing the red cell to expand and the yellow cell to retract. Whether this copy attempt is actually successful depends on the Hamiltonian $\mathcal{H}$.

can be copied into u. Whether this copy attempt is accepted depends on the balance of forces described in the Hamiltonian $\mathcal{H}$. If a copy attempt decreases $\mathcal{H}$, it is always accepted. For copy attempts which increase $\mathcal{H}$, the Boltzmann probability is used, which allows for stochasticity in cell movement. In total, the probability of a successful copy attempt is given by:

$$P(\Delta\mathcal{H}_{v\to u}) = \left\{ \begin{array}{ll} 1 & \text{if } \Delta\mathcal{H} < 0 \\ e^{-(\Delta\mathcal{H}/T)} & \text{if } \Delta\mathcal{H} \geq 0 \end{array} \right\}, \tag{1.5}$$

where T denotes a temperature of the system. For the Metropolis algorithm, a number of consecutive copy attempts equal to the total number of lattice sites is called a Monte Carlo step (MCS) and is used as a time measure. Within one MCS, each lattice site is expected to have been updated exactly once. Rejection-free versions of the Metropolis algorithm, which exclude sampling over neighbors with identical σ, have modified definitions of the Monte Carlo step [69], but are more efficient in computation time.

The Hamiltonian describes the balance of forces in the model. Terms included in the Hamiltonian can be varied to include cellular mechanisms of interest, such as adhesion energies between cells and between cell and medium, area constraints to ensure cell area, or perimeter constraints to influence cell shape. An example Hamiltonian including these terms could

look like this [70]:

$$\mathcal{H} = \sum_{x,y} J_{\sigma_x,\sigma_y}(1 - \delta_{\sigma_x,\sigma_y}) + \lambda_{\text{area}} \sum_{\sigma} (a_\sigma - A_\sigma)^2 + \lambda_{\text{perimeter}} \sum_{\sigma} (p_\sigma - P_\sigma)^2. \tag{1.6}$$

Some dissipative processes are not represented in the Hamiltonian directly, but are accounted for in $\Delta\mathcal{H}$ as additional work terms. One example is chemotaxis. A very simple chemotaxis algorithm for the CPM is to compare concentrations C at u and v such that $\Delta\mathcal{H}_{\text{chem}} = \chi(C(v) - C(u))$, where χ denotes the strength of chemotaxis [71]. This influences the probability of a successful copy attempt by replacing $\Delta\mathcal{H}$ in Eq. 1.5 by the sum $\Delta\mathcal{H}_{\text{total}} = \Delta\mathcal{H} + \Delta\mathcal{H}_{\text{chem}}$.

Cell motility in the CPM with the Hamiltonian described by Eq. 1.6 is limited to random passive fluctuations of the membrane and has low diffusivity. Multiple extensions of the CPM have been developed to simulate actively moving cells [58, 72, 59, 73]. In this thesis, we mainly use the Act model to model actively moving immune cells [73]. This extension models actin dynamics in a phenomenological way: the branching and polymerizing actin network pushes the membrane outward. This is represented in the Act model through so called Act values: aside from keeping track of the σ of each lattice site, each lattice site u now also has an Act value $Act(u)$. These values indicate how recently that lattice site was added to the cell and can be viewed as a measure of actin activity. More recently active sites can be viewed as containing polymerizing actin network pushing against the membrane, and cell extension from lattice sites with a local neighborhood of high Act values is favored. Similarly, retraction at such lattice sites is suppressed, according to:

$$\Delta\mathcal{H}_{Act}(u \to v) = \frac{\lambda_{Act}}{\max_{Act}} \left(\sqrt[|\text{NB}(u)|]{\prod_{y \in \text{NB}(u)} Act(y)} - \sqrt[|\text{NB}(v)|]{\prod_{y \in \text{NB}(v)} Act(y)} \right). \tag{1.7}$$

Here, λ_{Act} is a parameter determining the strength of this process in relation to the other terms in the Hamiltonian. Once a lattice site is added to a cell, that site will obtain the maximum Act value $\max_{Act}$. This initiates a feedback loop where further extension at the same edge of cell remains favored, resulting in cell polarization. The use of the geometric means in Eq. 1.7 ensures that neighborhoods where all Act values are high are favored and neighborhoods with 'holes' are nullified. A biological interpretation for this is that actin subunits contribute to growth of the actin network by attaching to the already existing network. By using the geometric mean, attachment of actin subunits to the existing network is enforced.

By tuning the parameters $\max_{Act}$ and λ_{Act}, the stability and strength of the cell front can be tuned. This results in persistent random walks with a variety in persistence times, and different motility behaviors can be observed, such as *Dictoystelium*-like motion and keratocyte-like motion.

1.4 Thesis overview

This thesis is organized as follows. In Chapter 2, we present ZebraGEM 2.0, an improved whole-genome scale metabolic reconstruction for zebrafish. The improvements include the addition of GPRs and the oxidative phosphorylation pathway, and make it possible to use the model for knock-out studies, simulating respiration experiments and predicting changes in metabolism based on gene expression data. We specifically study zebrafish metabolism upon infection with *Mycobacterium marinum* integrating gene expression data from control and infected zebrafish larvae.

The following chapters will focus on cell motility in response to the environment and revolves around the question how environmental input shapes and guides cell motility. Chapter 3 explores the different types of lymphocyte motility that can arise by the interactions between cell and extracellular matrix. We introduce an extension of the Act model with adhesion dynamics that can both show cell motion on a short time scale as well as derive statistical properties of cell motility. We observe a range of motility modes: 1) Brownian walks, 2) stick-and-slip walks where cells alternate between sufficient adhesion and too little adhesion to gain traction, 3) highly persistent walks, and 4) short-term persistent long-term subdiffusive walks.

In Chapter 4, we focus on a different role of the environment, namely how obstacles in the environment can guide cell movement. Immune cells have to navigate within tissues around the other cells in the tissue as well as the extracellular matrix. A model system for the role of obstacles in cell motility has previously been set up to study *Dictyostelium discoideum* cells in a gradient in density of cell sized pillars [74]. The *Dictyostelium discoideum* cells display actin-driven amoeboid motility, very similar to immune cells, and are shown to be guided by this gradient in pillar density from areas with high pillar density to area with low pillar densities. This process is called topotaxis: cell movement guided by topographical cues. Previous work on Active Brownian particles, which perform a persistent random walk, similar to the *Dictyostelium* cells, has shown that part of the topotaxis effect can be explained by altered cell persistence in the pillar grid [75]. However, the extent to which Active Brownian particles perform topotaxis

is lower than measured in *Dictyostelium* cells. We hypothesize that the amoeboid motility of *Dictyostelium discoideum* is better captured by the Act model and show that both deformable cell shape and active reorientation upon collision make cells more efficient at topotaxis.

The final chapter on cell motility is Chapter 5, but this also marks our return to zebrafish. Here, leukocyte trajectories from a tail-wounding assay in zebrafish larvae are analyzed. Two mutants in the TLR-signalling pathway are studied and we show that leukocyte migration towards the wounded area is significantly lowered in the mutant compared to wild type.

Finally, we conclude this thesis in Chapter 6, where we discuss the results of this thesis and propose ideas for future work.

Chapter 2

Predicting metabolism from gene expression in an improved whole-genome metabolic network model of *Danio rerio*

Leonie van Steijn, Fons J. Verbeek, Herman P. Spaink, Roeland M.H. Merks*

Abstract

Zebrafish is a useful modeling organism for the study of vertebrate development, immune response, and metabolism. Metabolic studies can be aided by mathematical reconstructions of the metabolic network of zebrafish. These list the substrates and products of all biochemical reactions that occur in the zebrafish. Mathematical techniques such as flux-balance analysis then make it possible to predict the possible metabolic flux distributions that optimize, for example, the turnover of food into biomass. The only available genome-scale reconstruction of zebrafish metabolism is ZebraGEM. In this study, we present ZebraGEM 2.0, an updated and validated version of ZebraGEM. ZebraGEM 2.0 is extended with gene-protein-reaction associations (GPRs) that are required to integrate genetic data with the metabolic model. To demonstrate the use of these GPRs, we performed an in silico genetic screening for knockouts of metabolic genes and validated the results against published in vivo genetic knockout and knockdown screenings. Among the single knockout simulations, we identified 74 essential genes, whose knockout stopped growth completely. Among these, 11 genes are known have an abnormal knockout or knockdown phenotype in vivo (partial), and 41 have human homologs associated with metabolic diseases. We also added the oxidative phosphorylation pathway, which was unavailable in the published version of ZebraGEM. The updated model performs better than the original model on a predetermined list of metabolic functions. We also determined a minimal feed

*Published as Leonie van Steijn et al. "Predicting Metabolism from Gene Expression in an Improved Whole-Genome Metabolic Network Model of Danio rerio". *Zebrafish* 16.4 (2019), 348–362

composition. The oxidative phosphorylation pathways were validated by comparing with published experiments in which key components of the oxidative phosphorylation pathway were pharmacologically inhibited. To test the utility of ZebraGEM2.0 for obtaining new results, we integrated gene expression data from control and Mycobacterium marinum-infected zebrafish larvae. The resulting model predicts impeded growth and altered histidine metabolism in the infected larvae.

2.1 Introduction

The zebrafish (*Danio rerio*) has become a widely used model organism for the study of vertebrate metabolism [76, 77]. Its genome has been sequenced and annotated [78] and the CRIPSR-Cas technique has made it easier than ever to study the role of specific metabolic genes [79]. For example, zebrafish have been used to test the toxicity of drugs on liver metabolism and the effect of liver metabolism on internal drug concentration [80]. Zebrafish have also been used in studies of metabolic diseases such as diabetes, obesity, and fatty liver disease, often combining sequencing with visualization of gene expression [76].

Mathematical and computational techniques make it possible to use such metabolic gene expression data to predict the flux of metabolites through single cells or even whole organisms. Genome-scale metabolic reconstructions, or metabolic maps for short, are models that consist of two parts: a metabolic network of the organism and the genes underlying this network. This network reconstruction is based on the genes coding for metabolic proteins present in the genome and sometimes requires manual curation to fills in gaps in the network [81].

Metabolic maps make it possible to predict how metabolites flow through a network of biochemical reactions, finally resulting in resources for growth or the availability of energy. Because in one network, an infinite number of alternative flow distributions are equally likely, a sensible prediction can only be made under the assumption of an objective, for example, optimal biomass production or optimal production of ATP, and a number of constraints on the possible fluxes. Most techniques assume flux balance, meaning that all biochemical concentrations are in equilibrium. Additional constraints can be given by known or assumed concentrations of enzymes, leading to a maximum flux through the reaction.

Mathematical techniques to make these predictions include Flux-Balance Analysis (FBA) [18] and derivate methods as Flux Variance Analysis, [20] Minimization of Metabolic Adjustment, [82] and Expression flux [83].These

predict the production rate of biomass or of a certain metabolite, for a given substrate, and sometimes supplemented with expression data. These predictions are valuable for finding suitable substrates for microorganism-based production in bioreactors. Another feature of these methods used to predict the flux through genome-scale metabolic models is the ability to study the effects of gene knockouts or gene expression on metabolism by constraining or removing reactions in the reaction network [28, 84]. This gives insight into the metabolic routing or rerouting of an organism and can be helpful in acquiring the aspired phenotype of an organism, but it can also give insight into the metabolic fluxes of different cell types.

With the increasing presence of metabolic data of healthy and diseased zebrafish, and the availability of genetic data, a genome-scale metabolic model of the zebrafish is tremendously useful. So far, genome-scale metabolic models have been proposed mainly for single-cell model organisms, such as *Escherichia coli* and *Saccharomyces cervesiae*, [85, 29, 86] as well as pathogens such as *Salmonella typhimurium* [87] and *Mycobacterium tuberculosis* [88]. For these unicellular organisms, very accurate growth predictions have been made. Multicellular organisms, particularly vertebrates, are less well represented in the list of genome-scale metabolic models. So far, reconstructions have been made for human, [30] mouse, [31] Chinese hamster, [32] fish, [89, 34] and recently, rat [33]. Whole-organism modeling is less common for these multicellular organisms, as metabolic functions are distributed over different tissues. However, modeling specific cell types has been done, such as erythrocytes [90] and cancer cell lines, [23] as well as integrating different cell types into a larger model, such as a combined model, including adipocytes, myocytes, and hepatocytes [91].

Why do we require a specific zebrafish genome-scale metabolic reconstruction when other vertebrate models exist? Despite the high metabolic similarity to human and mouse, there are subtle differences between zebrafish metabolism and the metabolism of these mammals that affect their required nutrients. For example, inositol-3-phosphate synthase is an enzyme present in humans and mice, but it is absent in zebrafish, preventing it from converting glucose-6-phosphate into inositol 3-phosphate [37]. This makes inositol an essential nutrient for zebrafish.

The difference in metabolism aside, the main reason to make a specific zebrafish genome-scale metabolic model is the genomic structure. The teleost lineage underwent a whole-genome duplication event after the radiation from their common ancestor with mammals, which resulted in numerous genes still having duplicate copies compared to mammals [92]. As a result, there are more paralogous genes in the zebrafish genome than in mammals. Hence, if one wants to study the effects of genes on metabolism,

translating a human or mouse genome-scale metabolic reconstruction into a zebrafish specific model by orthologous genes is not sufficient. Foremost, this translation is hampered by these paralogs as it does not make the translation one-to-one, and furthermore, many paralogs have evolved different subfunctions, increasing the functional difference between the zebrafish paralogs and the human or mouse orthologs. So to model the effects of genes on zebrafish metabolism, a zebrafish-specific genome-scale model is necessary.

Existing genome-scale models for zebrafish are MetaFishNet [89] and ZebraGEM [34]. MetaFishNet is a metabolic model derived from the genome of multiple fish species, including zebrafish, and focuses on individual pathways. As these pathways are not interconnected or divided into cell compartments, MetaFishNet is not suitable for whole-cell or whole-organism modeling using Flux Balance Analysis (FBA) methods, and therefore functions mainly as a reference tool, instead of a simulation tool. The fact that it combines multiple fish genomes also makes it harder to compare insights gained from this model to in vivo experimental results, as some pathways are solely based on the genome of one of those five fish species and do not occur in the other four fish species.

The other model, ZebraGEM, is based on the zebrafish genome and is a whole-cell and compartmentalized reconstruction. It contains 2911 reactions, of which 2446 are gene-associated reactions based upon 1498 genes and can be used for whole-cell metabolism modeling. It was reported to fulfill a list of 160 metabolic functions, such as the production of amino acids and biosynthesis and degradation of secondary metabolites. The model also predicted that the synthesis of taurine is through a metabolic pathway dependent on cysteine sulfinic acid decarboxylase, which is in line with experimental findings [93].

Currently, ZebraGEM cannot be used for modeling large screens of single gene knockouts or for the integration of gene expression data, as it lacks GPR. GPRs describe how gene products associated to a reaction work together, that is, whether they form a complex enzyme, are isoenzymes, or a combination of these. They provide a logical framework to decide whether a reaction can take place when one or more of its underlying genes are knocked out, and hence, they are of great importance when it comes to modeling gene knockouts.

In this article, we describe the modifications applied to ZebraGEM to fit our modeling needs and to fit standards of genome-scale metabolic reconstructions, as well as demonstrate a number of ways in which the updated model can be used. Briefly, the modifications fall into three categories. First, we added the GPRs, to facilitate gene knockout and gene expression

modeling. Second, we renamed components of the model according to BiGG Models standard names, [94] to ease comparison with genome-scale metabolic reconstructions of other organisms. Finally, we extended the model with essential reactions for pathways already present, or changed the reversibility of reactions already present in the model.

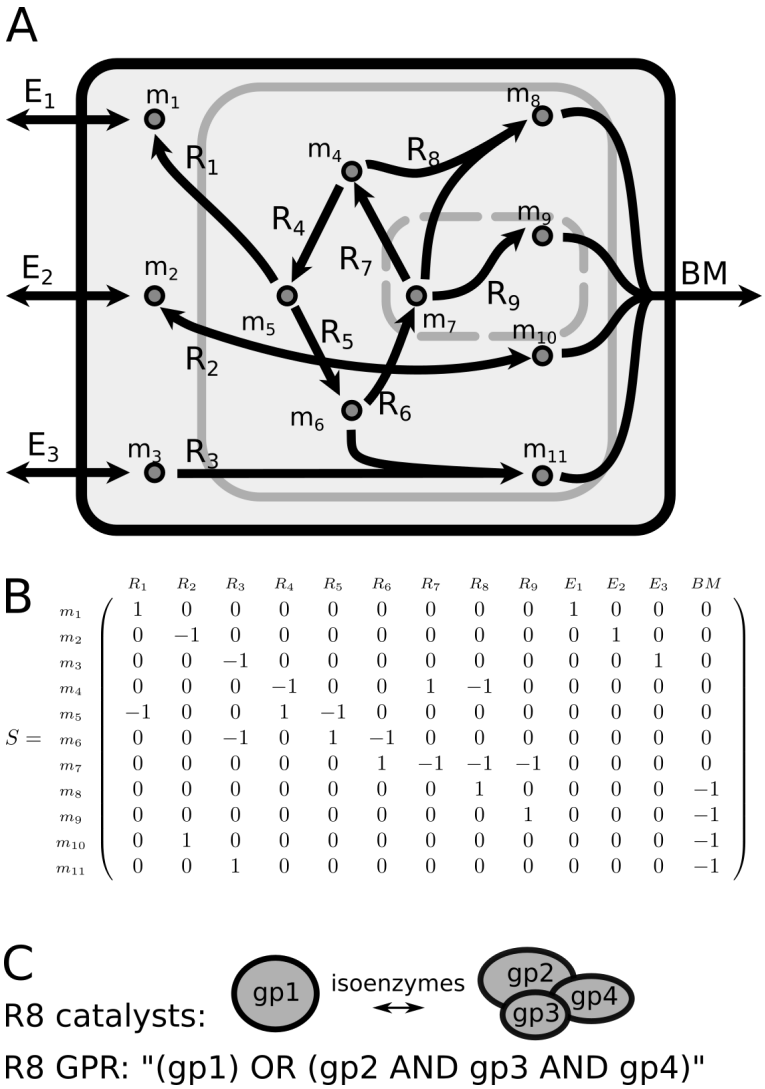
We have validated the renewed model against the metabolic functions the original model was reported to fulfill. Using the updated model, we predicted a minimal feed composition and were able to make predictions of mitochondrial function with respiration simulations. Finally, we also proved the usefulness of the newly added GPRs: we performed a large single-knockout and double-knockout screening and predicted lethal knockouts, and we also integrated gene expression data with the model to predict metabolic differences between control zebrafish larvae and larvae infected with *Mycobacterium marinum*.

2.2 Methods

The genome-scale metabolic reconstruction ("metabolic map") of zebrafish consists of the following: (1) a metabolic network describing the reactions that can occur in the organism and (2) the genes that are associated with those reactions (Fig 2.1). The network on its own can be used for modeling metabolism, and the associated genes give extra handles to this modeling. In this section, we give a general overview of the metabolic network component and gene component of a genome-scale metabolic reconstruction, as well as describe the modeling method called FBA. We also briefly address the representation of this model in a computer file.

2.2.1 Metabolic network

The metabolic network part of a metabolic map can be represented by a matrix S (2.1 A B). This matrix contains the ratio between reactants and products, or stoichiometry, for each reaction within the network, and is called a stoichiometric matrix. The rows represent the metabolites and the columns represent the reactions. The coefficient at the intersection of a specific row and column indicates the contribution of that metabolite to that reaction. Some of the reactions are of a special type, the so-called exchange reactions. These exchange reactions either have only a reactant or only a product, and hence do not preserve mass. They represent the influx and efflux of metabolites in and out of the system.



$$
S = \begin{array}{c} \\ m_1 \\ m_2 \\ m_3 \\ m_4 \\ m_5 \\ m_6 \\ m_7 \\ m_8 \\ m_9 \\ m_{10} \\ m_{11} \end{array}
\begin{array}{c}
\begin{array}{ccccccccccccc} R_1 & R_2 & R_3 & R_4 & R_5 & R_6 & R_7 & R_8 & R_9 & E_1 & E_2 & E_3 & BM \end{array} \\
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1
\end{pmatrix}
\end{array}
$$

FIGURE 2.1: Important components of a genome-scale metabolic reconstruction are the metabolic network (A, B) and the GPR (C). (A) Graphical overview of a simplified metabolic network. Reactions within the black border are part of the system and hence have mass balance. The solid gray border indicated the cell membrane and the dashed gray border indicates cell organelle membranes. Reactions E1–E3 are exchange reactions and are not mass balanced, allowing for import and export of metabolites. Reaction BM is a biomass reaction, taking biomass precursor metabolites and exporting them to biomass; (B) stoichiometric matrix representation of the network shown in (A); (C) example of how isoenzymes and protein complexes are translated into a GPR. GPR, gene-protein-reaction associations; gpx, gene product x.

2.2.2 Flux Balance Analysis

The standard method for constraint-based metabolic modeling is FBA [18]. For a given metabolic network and a given objective function, FBA computes the optimal flux through the metabolic network that minimizes or maximizes the objective function. The first assumption upon which FBA is based, is that an organism will adjust its fluxes such that the internal metabolites, indicated with c, are in equilibrium, that is

$$\frac{dc}{dt} = S \cdot \vec{f} = 0, \tag{2.1}$$

with $\vec{f}$ the vector representing the fluxes of the reactions in the metabolic network. Some of these fluxes can be constrained. For example, exchange reactions can be constrained due to limited availability of the exchanged metabolite in the environment. Also, irreversible reactions can be constrained, as they cannot have a negative flux. This can be formulated as follows:

$$a_i \leq f_i \leq b_i, \tag{2.2}$$

with a_i and b_i indicating the lower bound and upper bound of the flux of reaction i. Sometimes an exchange reaction has a strictly positive lower bound, indicating that the system should at least produce that amount of the exchanged metabolite. These reactions are called demand reactions.

Solving equation 2.1 and 2.2 together can lead to an infinite number of solutions. Within this solution space, FBA selects for a smaller solution space based on a predefined objective, for example, that the organism optimizes its metabolic fluxes for a specific reaction or for biomass production. This optimized reaction, or objective function f_{obj}, can be any reaction in the metabolic network, but most often, it is a biomass function. The biomass function lists all the precursor metabolites and energy-carrying metabolites required for the accumulation of biomass. Unless stated otherwise, we will use the biomass function as the objective function. The full formulation of the FBA problem then becomes as follows:

Optimize

$$f_{obj} \tag{2.3}$$

such that:

$$S \cdot \vec{f} = 0,$$
$$a_i \leq f_i \leq b_i$$

This forms a linear programming problem and can easily be solved using

linear programming solver software, for example, GNU linear programming kit (GLPK) or Gurobi. In this work, we have used CPLEX IBM ILOG CPLEX.

Once the linear programming problem is solved, the solution $\vec{f}$ gives a flux distribution of the metabolic network for the given constraints. This gives insight into which pathways are used and their relative contribution can be computed. By changing the upper and lower bounds in 2.2, one can test the flux distribution in different scenarios, such as comparing the growth rate under different sets of substrates.

Some common variations on FBA are parsimonious FBA [95] (pFBA) and Flux Variability Analysis (FVA) [20], which are multiobjective linear programming problems. After solving the original FBA problem, they then optimize a second objective. For pFBA, the secondary objective is to minimize the total sum of fluxes, that is, $\min \sum |f_i|$, while maintaining the same constraints as in the FBA problem, together with keeping the previous objective f_{obj} at its optimum. FVA is a method that explores more of the solution space, by searching for the minimum and maximum flux of each reaction. So after doing FBA, a new linear programming problem first minimizes and then maximizes each f_i, while also maintaining f_{obj} at its optimum and regarding all the previous constraints.

Multiple software packages for FBA exist. These function as an interface between the user and the linear programming solver. They allow for easy manipulation of bounds, easy addition and removal of reactions in the metabolic network, and modification of the GPRs, without having to keep track of the linear programming problem manually. The software used in this study is CobraPy [96], combined with the CPLEX solver.

2.2.3 Genes and constraint-based modeling

The second part of the metabolic map is the associated genes. These genes, responsible for the enzymatic reactions in the metabolic network, are represented using GPR. In its simplest form, the GPR links each enzyme with a biochemical reaction. If two enzymes catalyze the same reaction, the GPR becomes a logical expression. If they are isoenzymes, for example, they can both independently catalyze the reaction, an "OR" function is used. If the two enzymes form a complex such that both must be present to catalyze the reaction, an "AND" function is used. More complex GPRs can be described by nested logical expressions (2.1 Fig. 1C). In case multiple, equivalent logical expressions are possible, the disjunctive normal form is used, that is, a summation of all possible isoenzymes. Using the GPRs, gene knockouts or gene expression data can be integrated into constraint-based models. A

standard way of integrating gene knockouts is to set each occurrence of the knocked-out gene in a GPR to False and evaluate the GPRs. If any of these GPRs also evaluates to false, then constrain the corresponding reaction to 0 flux by setting its upper and lower bound to 0. Gene expression data can be integrated into constraint-based modeling in alternative ways [97, 25, 98, 99]. Although details vary, these methods either penalize fluxes over reactions with no or low expression and minimize the penalty or they set the lower and upper bound of fluxes depending on the expression level. The gene expression data integration method used in this study is Gene-centric flux (GC-flux) [100]. In this study, the linear programming problem is slightly altered from the original stoichiometric matrix-based linear programming problem. Using the GPRTransform package [101], we split up each reaction into multiple versions of the same reaction, one for every possible isoenzyme. The sum of the fluxes of all the reactions containing a certain gene in their GPR is then constrained by the expression level of that gene. Although many choices exists for how the expression level gives an upper bound, the simplest one is to take the expression level itself. So if we rephrase Equation 2.3 with the altered stoichiometric matrix S', the new programming problem becomes as follows:

First optimize

$$f_{obj} \tag{2.4}$$

such that,

$$\begin{aligned} S \cdot \vec{f} &= 0, \\ a_i \leq f_i &\leq b_i \\ \sum_{r \in R_g} |\vec{f}_r| &\leq E_g \forall g \in G \end{aligned}$$

Here R_g denotes the reactions belonging to gene g, E_g the expression of that gene, and G the total gene set. Basically, this algorithm distributes the gene expression among the different enzyme complexes, and hence the related reactions, of that gene, assuming that each molecule of a gene product can only take part in one complex at a time. The GC-flux algorithm originally also minimized the length of the flux vector, to obtain the most parsimonious flux distribution that optimizes the objective. We did not minimize the flux vector length, but applied FVA together with computing the relative flux range change (RFRC) to compare between the different gene expression data sets. With FVA, we determine for each f_i its minimum and maximum value that still allow for the objective to be optimized. To compare the flux ranges between different conditions, we compute the

RFRC of reaction i as follows [102]:

$$RFRC_i = \frac{c_{2,i} - c_{1,i}}{\frac{1}{2}(r_{2,i} + r_{1,i})},$$

with $c_{n,i}$ the center, $\frac{1}{2}(f_{i,max} + f_{i,min})$ of the flux range of reaction i in condition n, and $r_{n,i}$ the range width $f_{i,max} - f_{i,min}$.

2.2.4 Data standards for representation of metabolic maps

To facilitate exchange of computational models, such as metabolic models, in systems biology, the Systems Biology Markup Language (SBML) has been developed [103]. Different elements of a metabolic map, such as metabolites, reactions, genes, and GPRs, are represented by their own class in SBML. For this, we use the fbc package, the Flux Balance Constraints extension of SBML. This package is especially designed to describe these genome-scale metabolic reconstruction elements, and has specified guidelines on how an entity should be represented in an SBML file [104]. 41 The original model was already an SBML file, but predates the fbc package's release. Therefore, we adapted the model to fit with the fbc package guidelines.

Metabolite, reaction, and gene nomenclature

Aside from the file structure, there are also standards for the names of metabolites and reactions. This facilitates comparison and interfacing with metabolic maps of other organisms. We renamed the metabolites, reactions, and genes. Genes were renamed with their Entrez id [105]. The metabolites and reactions were renamed using, if possible, the data standard from BiGG Models, a knowledgebase of genome-scale metabolic network reconstructions [94]. Metabolites without BiGG name were renamed to their corresponding identifier in the Kyoto Encyclopedia of Genes and Genomes (KEGG) to facilitate easy lookup [106, 107, 108]. Reactions without BiGG name were not renamed, as no standardized names exist for these reactions yet, making up 689 of not-renamed reactions. The reactions that did not need renaming can be categorized into three groups. The first group includes transport reactions of metabolites without BiGG name. These reactions can be identified by the description of the reaction. The second group consists of reactions involved in the exchange of fatty acids between metabolites. The third group contains reactions involved in oxidation and reduction of metabolites using NADH/NAD+ or NADPH/NADP+. The

second and third group kept their original annotation, linking the reaction to a KEGG entry.

2.3 Results

In this section, we first describe the alterations in the model. These include alterations to the metabolic network, as well as the part of the model describing the relationships between genes and reactions. After that, we present the results validating our updated model. We first tested the metabolic expansion of the model by checking it for a list of metabolic functions, determining a minimal feed, and predicting mitochondrial function in respiration simulations. Next, we tested the GPRs in the model by doing knockout simulations. Finally, we apply the model to predict metabolic changes due to infection with *M. marinum*.

2.3.1 Reaction network

The alterations to the metabolic network encompassed the following five issues: (1) improvement of the biomass function and addition of reactions to enable synthesis of biomass precursor metabolites; (2) addition of oxidative phosphorylation; (3) correction of starch metabolism; (4) correction of the reversibility of reactions and their catalyzed or spontaneous nature; and (5) validation of the list of metabolic functions ZebraGEM was reported to be able to fulfill. Figure 2.2 summarizes the update in ZebraGEM, categorized into subsystems following the subsystem reaction associations from Virtual Metabolic Human (VMH), a human- and microbe-specific database on metabolism and metabolism modeling [109, 110]. The subsystems are sorted according to the number of reactions changed in each subsystem. Changes are of three types: "reaction added," "reaction deleted," and "reversibility changed."

Biomass function and biomass precursors

FBA and related modeling approaches [18, 25, 111, 112] assume that an organism or cell channels the metabolic fluxes to optimize a metabolic function, called the objective function. This objective function is often a biomass function, describing the relative amounts of precursor metabolites required for biomass production. Realistic biomass functions improve the realism of model predictions [113]. In the absence of exact data for zebrafish, we based the updated biomass function upon data from other vertebrates.

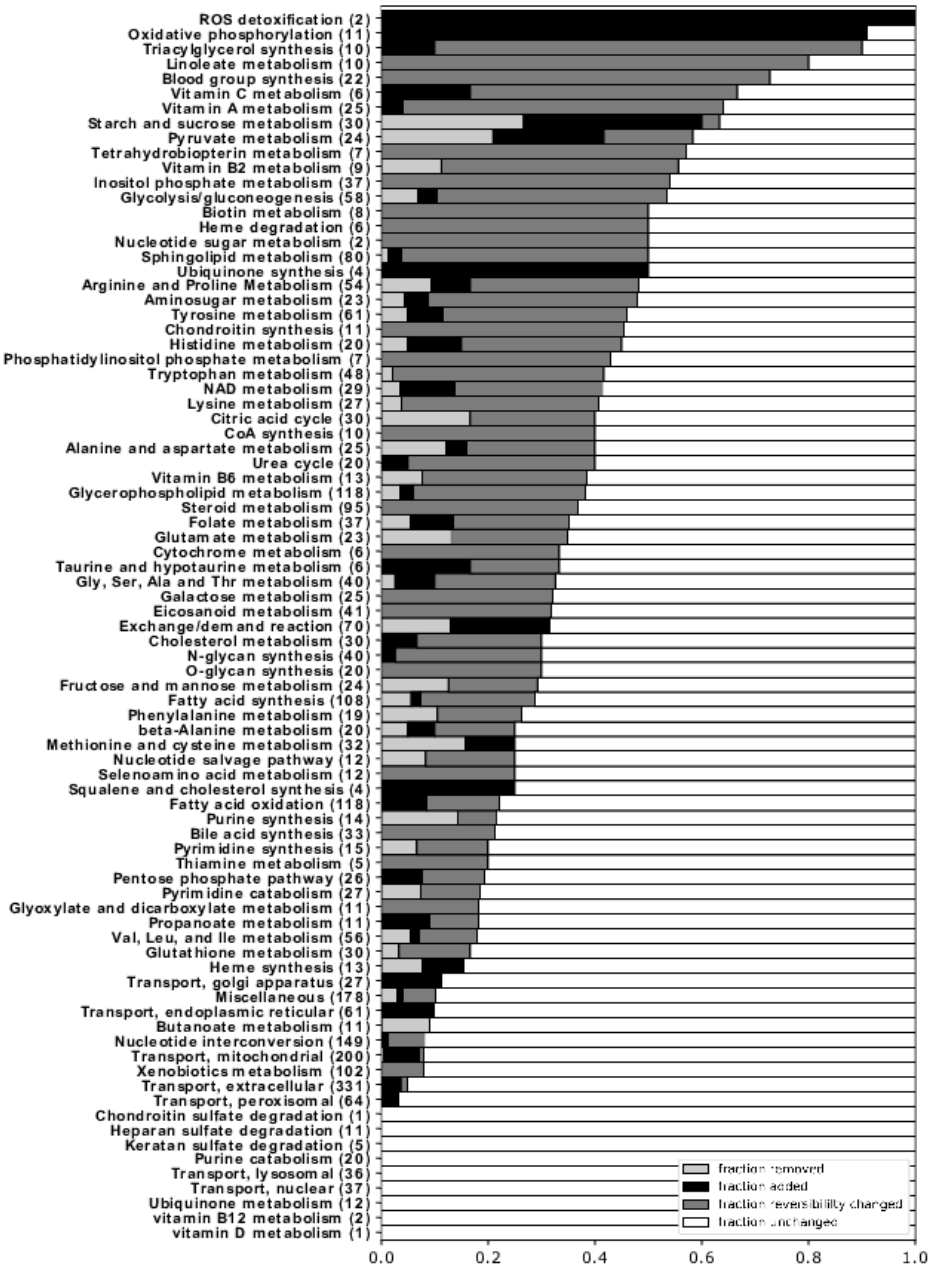


FIGURE 2.2: Subsystem overview of the adaptations made to ZebraGEM. For each subsystem, the total number of reactions, including the removed and added reactions, is noted in between brackets.

The biomass function coefficients were taken to be the average of the coefficients of biomass function of a human genome-scale reconstruction (Recon 2 [30]) and a mouse genome-scale reconstruction (iMM1415 [31]), so far the only other vertebrates with genome-scale reconstructions, together with Chinese hamster [32] and rat [33]. If a metabolite was a precursor in only one of Recon 2 and iMM1415, the coefficient was taken directly from the model in which the metabolite was present. If a metabolite was not present in both models, the coefficient was the average of a third, human three-tissue model, which had a biomass function for each tissue type [91].

Of the biomass precursors, 14 reactants and 2 products originally had stoichiometry coefficient 0 and were put in the biomass reaction for future work. Three of the reactants were cysteine, proline, and tyrosine, and with addition of reactions to their synthesis pathways, they could be produced. Nine of the reactants were membrane lipids, like cholesterol, sphingomyelin, and phosphatidylinositol, which also could be produced after the addition of reactions involved in their synthesis. We updated their coefficients in the same way as the other metabolites taking part in the biomass function. The remaining four metabolites were NAD, NADP, NADH, and NADPH. These were omitted from the biomass function, following Recon 2, iMM1415 and the human three-tissue model. iMM1415 nor the three-tissue model contained these metabolites in their biomass function. The resulting coefficients and their origin can be found in Supplementary Table S2.1.

Oxidative phosphorylation and starch metabolism

Oxidative phosphorylation in the model is an essential pathway for respiration. The corresponding reactions and genes were added to the model, using the human metabolic model Recon 2 as a template. Along with oxidative phosphorylation, it was also necessary to update "Ubiquinone synthesis," as well as to add the reactions CATm and SPODMm, represented in "reactive oxygen species (ROS) detoxification," to have a functional oxidative phosphorylation pathway.

We have also revised glycogen metabolism, using Recon 2 as a template, as the stoichiometry in the original model led to mass imbalance. The original reactions were replaced with those from Recon 2, replacing the genes within the GPRs for zebrafish orthologs. Changes in glycogen metabolism are shown in Figure 2.2 under subsystem "Starch and sucrose metabolism" according to VHM.

Reaction reversibility and reaction nature

All reactions in the model were checked for reaction reversibility. This corrected two types of unrealistic behavior. First, ZebraGEM produced essential nutrients through backward reactions (Supplementary Table S2.2). This was solved by correcting nonbiological reversible reactions in the corresponding pathways. Second, several metabolites were tunneled over membranes, as the same reaction occurred on both sides of a membrane that involved a membrane metabolite. If at least one of these reactions was reversible, this could result in spurious transport of the nonmembrane metabolites, often NAD or NADP. By checking the reversibility of the reactions with the reaction databases BiGG, VMH, and KEGG combined, this free transport cycle could be broken. The fraction of reactions with reversibility changed per subsystem is shown in Figure 2.2. In total, the reversibility of 543 out of 3023 reactions was changed.

A final check was done to ensure that all reactions in the updated model do occur in zebrafish metabolism. Reactions without gene regulation were checked using the KEGG database, a database containing information on genes and reactions. Their KEGG entries were tested for two conditions: (1) whether the reaction could occur nonenzymatically, and if not, then (2) it was checked whether the reaction has an enzyme associated to vertebrates, thus excluding reactions that occur in bacteria only. If any of these two conditions was met, the reaction was kept; otherwise, we deleted the reaction. The subsystems with deleted reactions are also shown in Figure 2.2.

Metabolic functions

The original model was reported to fulfill 160 metabolic functions, ranging from amino acid metabolism to pyrimidine and purine metabolism. In our hands, using the downloadable SBML file of the original model in the supplements, only 92 of these functions were fulfilled (Supplementary Table S2.4). Twenty-seven of the failed functions required metabolites in compartments that were absent in those compartments in the model. The other failed functions were checked manually using From Metabolite to Metabolite (FMM [114]) and KEGG for missing reactions, or for missing transport reactions that should be present in zebrafish. The missing reactions and their corresponding genes were added to the model. An overview of the subsystems with reactions added is shown in Figure 2.2.

TABLE 2.1: Comparison of the original ZebraGEM model with the updated version.

Number of	ZebraGEM	ZebraGEM 2.0
reactions	2911	3023
metabolites	2742	2810
unique metabolites	1554	1557
genes	1498	1636
gene regulated reactions	2446	2523
blocked reactions	1572	1678
successful metabolic functions	92	123
failed metabolic functions	41	12
metabolic functions missing metabolites	27	25

Genes and gene-protein-reaction associations

The original model already had 2446 gene-associated reactions coded for by 4988 genes (1498 unique genes). We extended the model by putting these gene products into a GPR, and added this to the model according to the SBML guidelines. As a result, the full model can now be read and run using constraint-based modeling software, and is now suitable for gene knockout simulations and simulations with gene expression data integration.

In summary, 95 reactions were removed and 140 were added to the model, and 543 reactions had changed reaction reversibility. The updated model now contains 3023 reactions with 2810 metabolites, of which 1557 were unique, and 1636 genes. Two thousand five hundred and twenty-three reactions are gene regulated and 1678 reactions are blocked, that is, are unable to carry any flux due to dead-end metabolites. A comparison between the original ZebraGEM model and the updated model is shown in Table 2.1.

2.3.2 Model validation

To check whether the changes in the model network improved the performance of the model, we tested the model predictions as follows: (1) we checked whether the model performed the metabolic functions reported in Bekaert [34]; (2) we checked for biological validity of the minimal set of metabolites required for model growth; (3) we checked whether the model could reproduce pharmacological interference with respiration. We utilized the addition of the GPR by doing single- and double-knockout experiments, and ultimately by gene expression data integration.

Model metabolic functions

ZebraGEM was published with a list of 160 metabolic functions it was reported to fulfill (Supplementary table 3 of Bekaert [34]). A metabolic function on this list consists of one or multiple starting metabolites and one or more end metabolites, indicating that a metabolic route between these metabolites fulfills this function. We tested these functions by setting an import reaction for the starting metabolites and an export reaction for the end metabolites. The export reaction for the end metabolites was chosen as the objective function, and a function was deemed successful if the model imported the starting metabolites and exported the end metabolites. Some of these metabolic functions could not be tested, as the starting or end metabolite was not present in the model. Metabolic functions that did not result in a success immediately were checked by hand to see whether the model has an alternative path to fulfill the demand for the end metabolite.

Out of the 160 metabolic functions, after the corrections, ZebraGEM 2.0 was able to perform 123 functions successfully and still failed to perform 12 functions. Of the remaining 25 metabolic functions, the starting or end metabolite was absent in the model and the corresponding function could not be tested (Table 2.1).

Minimal feed composition

To validate the new biomass function and the changes to the reaction reversibility, which corrected spurious production of essential amino acids, we determined a minimal feed composition that would allow for growth. The model was set to produce 1 arbitrary unit of biomass flux. As the model objective, we minimized the uptake of metabolites from the environment. The source metabolites include amino acids, the fatty acids linoleic acid and linolenic acid, minerals, oxygen, and inositol (Fig. 2.3). We chose glucose as the sole carbohydrate source.

The updated model predicts that the amino acids arginine, histidine, and threonine are essential for biomass production, whereas they were nonessential in the original model (Fig. 2.3). The updated model also predicts additional uptake of glucose. In the original model, spurious glucose was produced from imbalanced glycogen reactions, leading to increased glucose uptake in the updated model. The updated model now also predicts uptake of oxygen, due to the updated model for oxidative phosphorylation (data not shown). The ratio between the metabolite species taken up from the environment has also changed in the updated model, due to the updated stoichiometry of the biomass function. This is most

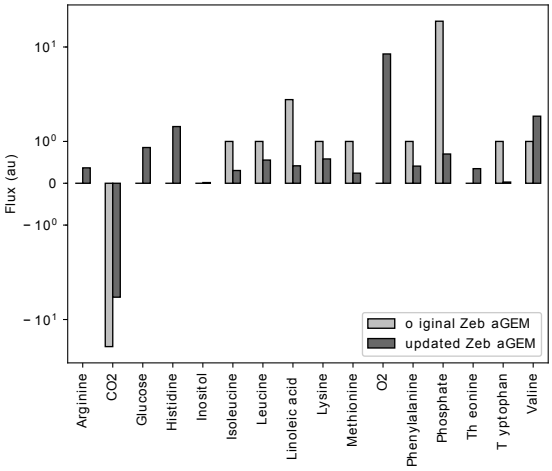


FIGURE 2.3: Minimal required metabolite uptake fluxes for the production of 1 arbitrary unit of biomass flux for both the original model and the updated model. Metabolite excretion fluxes are also shown, but were not constraining the minimization.

clearly the case for phosphate uptake (Fig. 2.3), which dropped from 71% of total metabolite uptake to 3%.

Thanks to the updated biomass function, inositol is now also an essential metabolite for growth in the model. Inositol is thought to be essential for zebrafish as no gene for inositol-3-phosphate synthase has been found. Inositol essentiality has been experimentally confirmed in other fish species, even in fish species with *de novo* synthesis of inositol [115, 116, 117]. The model currently does not require the essential fatty acid linolenic acid to grow, as the lipid metabolism in the model uses a generic fatty acid and the correct conversion of linolenic acid into this generic fatty acid is not present in the model. Further improvements connecting and specifying the used fatty acid in the lipid metabolism subsystem are required; see also in the Discussion.

Respiration

We next tested if ZebraGEM 2.0 correctly predicts oxidative phosphorylation. The mitochondrial oxidative function of zebrafish can be tested in vivo by measuring the oxygen consumption rate, which has been done in zebrafish embryos [118]. In Gibert et al. [118], the consumption rate of oxygen has been measured under the addition of three different compounds disrupting oxidative phosphorylation. We have simulated the effects of

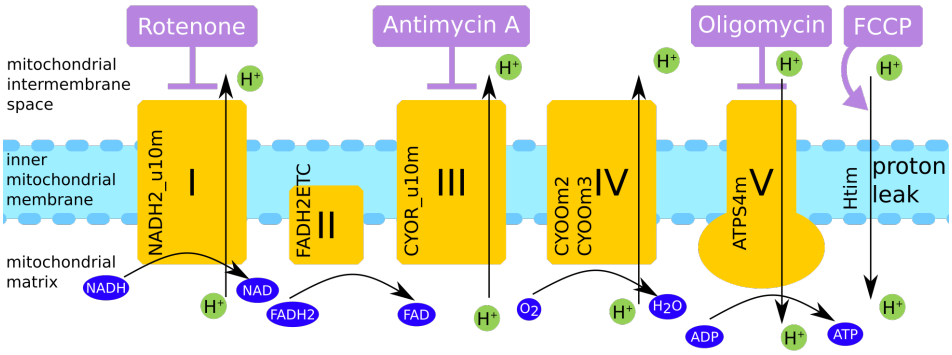


FIGURE 2.4: Overview of oxidative phosphorylation, with the site of action of the disrupting compounds rotenone, Antimycin A, oligomycin, and FCCP. The model reaction names are next to the corresponding enzyme, except for Htim, which represent, the proton leak and hence has no corresponding enzyme.

these compounds using the updated ZebraGEM model with pFBA. The site of action of these compounds and the model reactions active in oxidative phosphorylation are shown in Figure 2.4.

The mitochondrial oxidative function can be tested *in vivo* by measuring the oxygen consumption rate, which has been done on zebrafish embryos [118]. In this study, the consumption rate of oxygen has been measured under the addition of four different compounds disrupting oxidative phosphorylation. Without any compound the basal respiration is measured. Adding oligomycin, an ATPase inhibitor, the respiration related to ATP production can be derived. Under the addition of FCCP, a proton uncoupler, the maximal respiration rate is measured. Finally, rotenone, a complex I inhibitor, and Antimycin A, a complex III inhibitor, are added such that the non-mitochondrial respiration can be measured. However, only rotenone was used in the last step in the zebrafish embryo study [118].

First, the basal respiration rate is determined. In the experimental setup, this was done by measuring the oxygen consumption flux of embryos in the absence of disrupting chemicals. In our simulations, we optimize the model for biomass production with pFBA. Because the cellular environment within zebrafish is unknown, we used 1000 randomly created environments. For each of these environments, we sampled the upper bounds of metabolite uptake from selected ranges, such that the uptake was the constraining factor in biomass optimization. We used the same random environments for simulations of disruptive compounds.

Second, in Gibert et al. [118] the maximal respiration rate was measured

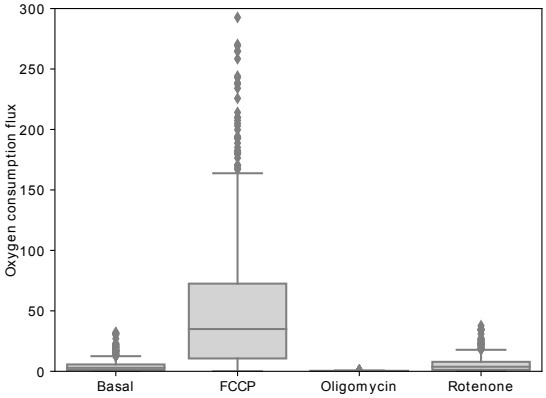


FIGURE 2.5: Oxygen exchange for the four modeling conditions shown in box plots.

after exposure to the proton uncoupler FCCP. This uncoupler allows for proton flux over the inner mitochondrial membrane, bypassing ATPase. We simulated this by blocking the model reaction ATPS4m (Fig. 2.4), the model equivalent of ATPase, and again optimizing for biomass production with pFBA. The experimental results show a 29% increase in respiration compared to basal respiration. Our FCCP simulations, Figure 2.5, second column, show a 10-fold increase in mean value compared to our basal respiration simulations mean value.

After that, a new assay was performed in Gibert et al. [118] exposing the embryos alternatively to oligomycin, an ATPase inhibitor, and rotenone, a complex I inhibitor. By comparing the respiration rate after oligomycin addition, the respiration related to ATP production can be derived. We simulated the effect of oligomycin by again blocking ATPS4m, together with limiting the flux through the uncoupling reaction that transports protons over the inner membrane (Htim, Fig. 2.4). The latter constraint is necessary as proton gradients cannot develop in FBA. The Htim flux upper bound was set equal to the Htim flux from the basal respiration simulations to reflect the maximal buildup of proton gradient. The experimental results show that ATP turnover-related respiration contributes about 60% to basal mitochondrial respiration; in our simulations, this would be about 90%. This is due to a side effect of blocking ATPS4m together with the limit on Htim. As the proton back flow is limited, ubiquinone cycling is also limited. Ubiquinone is required for the reaction catalyzed by dihydroorotate dehydrogenase, an essential part of pyrimidine synthesis. With limited pyrimidine synthesis, the biomass production is also limited. As the upper bound for Htim is often 0, the model does not grow at all, and hence

requires no oxygen.

The final compound rotenone can be used to measure the nonmitochondrial respiration, as the electron transport chain is blocked and no oxygen is consumed by complex IV. We modeled the effect of rotenone by blocking the reaction associated to complex I: NADH2_u10m (Fig. 2.4). The experimental results show that nonmitochondrial respiration contributes to about 40% of basal respiration. Our simulations show a different picture, as the oxygen consumption flux is larger in the rotenone simulation than in the basal simulation. (Fig. 2.5, column 4). The rotenone simulation should represent respiration where the entire electron transport chain has been blocked, resulting in nonmitochondrial respiration. However, by only restricting the flux of NADH2_u10m, the electron transport chain is not entirely blocked in the model, allowing for respiration similar to the basal case. An extra compound that can be used to study nonmitochondrial respiration is Antimycin A, which inhibits complex III. Although not used in Gibert et al. [118] we tried simulating the effects by blocking the complex III corresponding reaction CYOR_u10m. However, in this case, the model fails to grow at all.

Overall, the model is able to simulate the qualitative behavior of basal, FCCP-influenced, and oligomycin-influenced respiration. It is impossible to use FBA to describe the proton gradient. Our choice to describe the proton gradient with Htim flux from the basal simulation proved too strict, and choosing a higher Htim upper bound could improve the model outcome. The rotenone/Antimycin A simulations also exposed some problems with the model that are still open, such as alternative electron transport routing and total biomass dependency on the reaction CYOR_u10m.

Gene knock-out simulations

Next, we validated the utility of the GPRS by performing an *in silico* screen for gene knockouts. To simulate a gene knockout, we set gene activity to "false" in each GPR that contains the gene. The other genes in the GPRs were set to "true," and the logical expression of the GPR was evaluated. If the GPR evaluated as "false," the flux through the associated reaction was blocked. Using FBA, we optimized biomass production in the presence of the additional constraint. The procedure was repeated for each gene. We also screened for double gene knockouts. In this case, each pair of genes in the network was set to "false" and the same procedure was applied for double knockouts. The resulting knockout biomass production rate was expressed as a fraction of the wild-type biomass production rate, that is,

we divide to optimal biomass production rate in the knockout case over the optimal biomass production rate in the "wild-type" control.

Out of the 1636 genes in the model, 74 single knockouts completely blocked biomass production. For further 30 genes, single knockout reduced biomass production rates. Out of these 30 single knockouts, 13 single knockouts resulted in a biomass production rate ranging from 0.4038 to 0.8 of the optimal biomass production rate and 17 have a slightly reduced biomass production rate ranging from 0.8 to 0.95 of the optimal rate. A further 42 single knockouts resulted in a very minor reduction in biomass production, ranging from 0.95 to 0.9998 of that of the wild type. All these genes are listed in Supplementary Table S2.3A. The model was robust to single knockout of the 1490 other genes in the model, yielding a biomass production rate identical to that of the wild type. The genes resulting in a nonoptimal phenotype were mostly involved in oxidative phosphorylation (37 of 146), followed by cholesterol metabolism (14), nucleotide interconversion (8), and synthesis (11). We see a good correlation of the essential and partial-essential genes and the pathways for biomass precursors that we added to the biomass function as well as oxidative phosphorylation.

To validate our single-gene knockout simulation results, we searched the literature for mutagenesis screens in zebrafish screening for visible defects (Fig. 2.6) [119, 120, 121, 122, 123, 124, 125, 126]. Thirty-six of all our model genes had at least one record in these screens. Out of these 36 genes, 6 knockouts were among the 74 knockouts with fully blocked biomass production (*paics*, *tyms*, *cdipt*, *rrm1*, and *cad*). One knockout (*atp5po*) resulted in a reduced biomass production rate of 0.509 of the wild-type rate. For the remaining 29 knockouts from these *in vivo* screens, ZebraGEM 2.0 did not predict a reduced biomass production. These genes without model phenotype are also included in Supplementary Table S2.3A.

We next used ZebrafishMine to extract single-gene knockdown non-normal phenotypes from the Zebrafish Information Network (ZFIN) [127]. Around 232 genes present in ZebraGEM 2.0 had a knockdown phenotype in ZFIN. Of those 232 genes, 18 genes also had reduced biomass production in the single knockout simulations (Supplementary Table S2.3A and Fig. 2.6), 8 had no growth, 1 had rate 0.647 of wild-type rate, 5 had a rate in the range 0.8–0.95 of wild-type rate, and 4 had a rate ranging from 0.95 to 0.9998 of wild-type rate. The low number in overlap between model knockout phenotypes and *in vivo* phenotypes can be caused by open problems within the model.

On the other hand, not every gene has been extensively studied in zebrafish, which might also explain part of the model knockouts with reduced biomass production rate, but no record in the zebrafish literature.

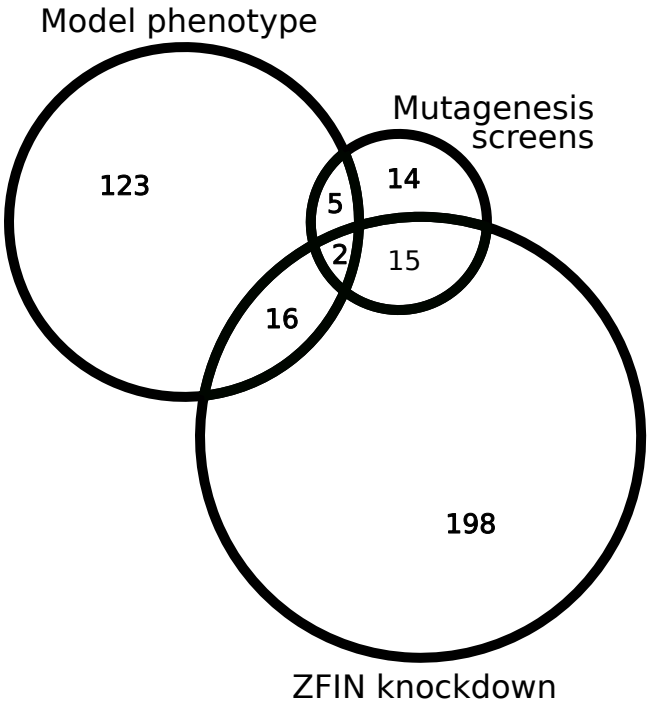


FIGURE 2.6: Venn diagram of genes present in the model that result in a phenotype in the single knockout simulation (model phenotype), are present in the genetic screen studies (screens) [119, 120, 121, 122, 123, 124, 125, 126], and have a knockdown abnormal phenotype registered in the Zebrafish Information Network (ZFIN) (knockdown).

For this reason, we also used ZebrafishMine to check the remaining 123 genes that have a phenotype in the model for diseases associated with their human orthologs. Of these 123 genes, 69 have a metabolic disease associated to their human ortholog, with the exception of *sod2* and *got1* that are associated with microvascular complications of diabetes and low serum levels of aspartate aminotransferase, respectively (Supplementary Table S2.3A). Of the remaining 54 genes without associated disease, there is still the possibility that they point to problems in the model, or that they are associated with rare mutations that have not been studied yet. Twenty-five of these genes were related to oxidative phosphorylation, which might indicate the latter.

In total, 228 genes appeared in Refs. [119, 120, 121, 122, 123, 124, 125] and ZFIN with a non-normal phenotype, but showed no phenotype in the single-gene knockout simulation. We categorized the effects of the knockout of these genes. One hundred and seven genes were involved in blocked reactions only, so knocking those out results in no change in the model. For 59 genes, the corresponding reactions of the genes would divert flux from the biomass production; thus, if wild-type model is optimized for biomass production, those reactions are already minimized to 0 flux. Next, there were also 42 genes that are redundant in our model: knocking those out does not block any reaction. It could be that subfunctionalization on the level of enzyme kinetics causes the in vivo phenotype, which cannot be represented with FBA modeling. Finally, there are 20 remaining genes that do not fit any of the three categories mentioned. Their associated reactions might be redundant within the network or do not contribute to biomass production.

For the double knockouts, we looked at two sets of genes pairs. First, we looked for pairs of genes with lower growth rates, which do not involve genes with phenotype in the single knockout simulation. The gene pairs with lowered growth rate (44 in total, 22 of which show no growth at all) are shown in Supplementary Table S2.3B, and are often paralogous genes. We also checked gene pairs involving at least one gene with a lowered growth rate in the single knockout experiment, which resulted in no growth, and found 36 pairs, also shown in Supplementary Table S2.3B. Lethal double knockouts are mainly involved in lipid metabolism, amino acid metabolism, and the citric acid cycle. In contrast to the single knockout simulation, the gene pairs that are lethal only in double knockouts do not account for much of the newly added reactions, with the exception of gene pairs involved in oxidative phosphorylation.

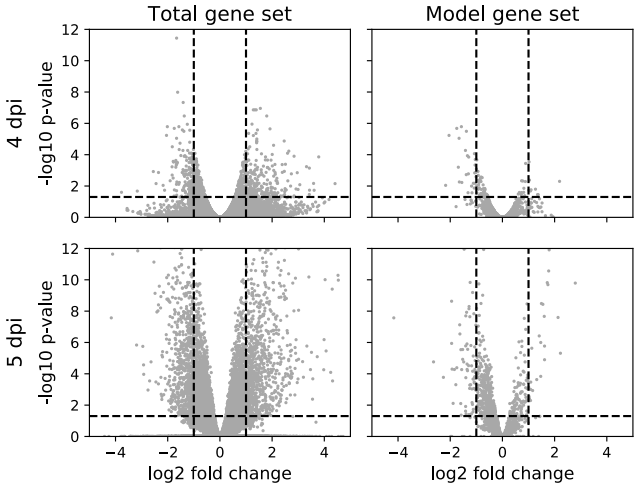


FIGURE 2.7: Volcano plots of the gene expression data set for both 4 and 5 dpi. Total data set on the left, the model subset on the right. Dashed lines indicate cut-off values: $-\log_{10}(p) > 1.301$, $|\log_2(foldchange)| > 1$. dpi, days post infection.

2.3.3 Integration of expression data

Thanks to the GPRs, ZebraGEM 2.0 can predict metabolic changes driven by changes in gene expression. We demonstrate this application of ZebraGEM 2.0 with a published dataset of infection with the fish tuberculosis bacterium *M. marinum* [128]. Briefly, zebrafish larvae were injected in the yolk with *M. marinum* at 2 h postfertilization [128]. Gene expression in infected and control larvae was measured at 4 and 5 days postfertilization using RNA deep sequencing. This yielded a data set containing the expression of 31,388 genes.

Of these 31,388 genes, 1608 genes are present in ZebraGEM 2.0. Although this is a small fraction of the total gene set, it covers 98% of the model genes. From these 1608 genes present in ZebraGEM 2.0, we selected genes with differential expression in the infected and control groups at 4 and 5 days postinfection (dpi). Genes were considered "differentially expressed" if they had a fold change or a fold change , together with an adjusted p-value threshold (Fig. 2.7). We thus identified 24 metabolic genes in ZebraGEM 2.0 that were differentially expressed both at 4 dpi and 5 dpi (Tables 2.2, and 2.3).

We next predicted the metabolic changes caused by differential expression of these 24 expressed genes. We made use of GC-flux [100]. The GC-flux algorithm constrains the rate of the metabolic reaction in the model

TABLE 2.2: Number of differentially expressed genes in the total gene expression dataset and the subset of genes present in the model.

	Total gene set	Model gene set
4 dpi	408	35
5 dpi	1714	106
both dpi	226	24

TABLE 2.3: List of genes differentially expressed at both 4 and 5 dpi that are present in the model.

Gene symbol	Gene name
acsl5	acyl-CoA synthetase long-chain family member 5
ampd3b	adenosine monophosphate deaminase 3b
anpepb	alanyl (membrane) aminopeptidase b
asah2	N-acylsphingosine amidohydrolase 2
dpys	dihydropyrimidinase
elovl8b	ELOVL fatty acid elongase 8b
enpp7.1	ectonucleotide pyrophosphatase/phosphodiesterase 7, tandem duplicate 1
ftcd	formimidoyltransferase cyclodeaminase
gch2	GTP cyclohydrolase 2
ggt1b	gamma-glutamyltransferase 1b
mboat2a	membrane bound O-acyltransferase domain containing 2a
neu3.3	sialidase 3 (membrane sialidase), tandem duplicate 3
neu3.4	sialidase 3 (membrane sialidase), tandem duplicate 4
pfkfb3	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3
ptgs2a	prostaglandin-endoperoxide synthase 2a
sat1a.2	spermidine/spermine N1-acetyltransferase 1a, duplicate 2
slc13a3	solute carrier family 13 (sodium-dependent dicarboxylate transporter), member 3
slc26a3.2	solute carrier family 26 (anion exchanger), member 3, tandem duplicate 2
slc7a7	solute carrier family 7 (amino acid transporter light chain, y+L system), member 7
tdo2a	tryptophan 2,3-dioxygenase a
tyms	thymidylate synthetase
ugt1ab	UDP glucuronosyltransferase 1 family a, b
uroc1	urocanate hydratase 1
zgc:92040	zgc:92040

based on the expression levels of the genes coding for the corresponding enzymes. GC flux distributes the gene expression of a single gene over all reactions associated with that gene, such that the total sum of those reaction fluxes cannot exceed maximum flux associated with the gene expression value. We performed this analysis for control and infected larvae at 4 and 5 days dpi.

After the model was constrained with the gene expression data, a method called FVA was applied [20]. FVA predicts the minimum and maximum possible flux ranges for each reaction, given an objective function; in this study, we used biomass production rate. To compare the flux ranges between control infected at 4 and 5 dpi, we used the RFRC [102]. The RFRC is a measure that indicates how much the flux ranges differ between the control and infected simulations. When the RFRC is greater than 1 or smaller than -1, the centers of the compared flux ranges are separated by more than the averaged width of those flux ranges, with negative values indicating that the infected case has a range lower than the control case.

An important reaction with an absolute RFRC greater than 1 is the biomass function BIO_L_2 and it appears in the list for both 4 and 5 dpi. The RFRC of BIO_L_2 is negative in both cases, -18.371 for 4 dpi and -17.421 for 5 dpi, suggesting that infection reduces biomass production rate. When comparing the maximal growth rates, the growth rate of the infected simulation was 83% of the control growth rate at 4 dpi, and at 5 dpi, the infected group reached 84% of the growth rate of the control. Further examination of the list with reactions with absolute RFRC greater than one (Supplementary Table S2.5) shows that affected reactions (with) at 5 dpi (46 reactions in total) are also affected at 4 dpi (56 reactions in total). Most of these 46 reactions were essential reactions involved in biomass precursor production and their knockouts are lethal (Supplementary Table S2.3A). The fluxes of the biomass precursor reactions co-vary, because they contribute, often in parallel, to the biomass reaction. If one of the fluxes is reduced, biomass production rate is also reduced. Due to flux balance, all the other biomass precursor fluxes must be reduced as well.

To gain insight in which genes give rise to such restricting reactions, and hence are limiting growth in our simulations, we identified the genes that restricted biomass production by comparing the flux corresponding to each gene with the expression level of each gene (Table 2.4). In total, 17 genes restricted biomass production in at least one of the four cases (condition x dpi). Aside from essential biomass precursor reaction-associated genes (essential genes for the model), 9 genes out of 17 are not essential to the model. Among these are *si:ch1073-100f3.2*, *slc5a9*, and *tha1*, all associated to monosaccharide transporters. The differential expression of

TABLE 2.4: Genes with gene expression restricting biomass production in the model with their fold change and their essentiality within the model, according to lethal phenotypes (essential) and reduced growth phenotypes (semiessential) in Supplementary Table S2.3A

Gene	FC 4dpi	FC 5dpi	Essentiality
acacb	0.522	0.036	essential
arg1	-0.402	-0.837	semi-essential
atp5s	0.358	0.088	semi-essential
bdh2*	-0.403	-0.810	semi-essential
cox6a2	-0.437	-0.633	essential
ftcd	***-1.061***	***-1.353***	***semi-essential***
galk1	-0.173	-0.669	-
galk2	-0.314	-0.315	-
gart	-0.262	-0.016	essential
gck	**1.871**	**-4.162**	-
hkdc1	-0.529	-1.469	-
nme4	**-0.548**	**-1.147**	-
nme6	-0.548	-0.267	-
si:ch1073-100f3.2*	-0.492	-0.277	semi-essential
slc2a11a	**0.068**	**-1.014**	-
slc5a9	-0.788	-0.791	-
tha1*	0.489	0.686	-

Genes marked with an asterisk are not restrictive for 5 dpi. Bold face genes have differential expression for 5 dpi, bold and italic font both 4 and 5 dpi. FC, fold change.

slc2a11a, also associated to a monosaccharide transporter, together with limited availability of flux for the other monosaccharide transporters, puts a large restriction on the model. The low number of only four genes with differential expression (namely *ftcd* at both 4 and 5 dpi, and *gck*, *nme4*, and *slc2a11a* at 5 dpi only) points toward a drawback of this data integration method: it only looks at the mean values of each case, but ignores whether these means are significantly different.

We observed that there was a reduction in growth rate in the infected case, and could ascribe this to a number of restricting genes. However, growth reduction might not be the only difference in metabolic activity; which metabolic pathways are contributing to biomass production can also differ between control and infected. To see if there was also a shift in which metabolic pathways contribute to biomass production, the flux ranges were normalized with the biomass flux. The RFRC was then again computed

with the normalized ranges, and only for 4 dpi were there reactions with $|RFRC| > 1$. These reactions are HISD, IZPN, URCN, and EX_his__L_e, and are involved in the pathway converting histidine into glutamate. The high $|RFRC|$ of these reactions can be directly linked to the differential expression of *uroc1*.

Overall, the addition of GPRs to ZebraGEM 2.0 together with GC-flux allowed us to integrate gene expression data into ZebraGEM 2.0, providing us with novel insights into potential metabolic changes due to *M. marinum* infection. First of all, there is a reduction in growth in the infected cases. This can be attributed to differences in the expression of some essential genes as well as monosaccharide transporter genes. When looking at qualitative changes in metabolism, histidine metabolism is reduced at 4 dpi, due to reduced expression of *uroc1*. Together with the restrictive gene *ftcd* (Table 2.4), which is also involved in the histidine pathway, this could make the histidine pathway an interesting starting point for more research on changes in metabolism upon *M. marinum* infection.

2.4 Discussion

In this work, we have presented ZebraGEM 2.0, an improved version of the genome-scale metabolic reconstruction ZebraGEM [34]. We have made the model available through an xml-file, see Supplementary Materials S2.6. The improvements were the addition of GPRs, significant changes to the stoichiometry by the addition of oxidative phosphorylation and checking the reversibility of reaction, and adhering to the existing standards of genome-scale metabolic reconstructions. To validate the new model, we have shown that it performs better than the previous version on a predetermined list of 160 metabolic tasks. We also determined a minimal feed. ZebraGEM assigns more nutrients to be essential, which is in agreement with what is known about zebrafish nutrition. To test the added GPRs, we did an in silico knockout screening, and found a large agreement between genes causing a phenotype in the model and genes that are known to have a phenotype in vivo in zebrafish or in human.

Altogether, ZebraGEM 2.0 is now suitable to be used with gene expression, which we demonstrated by integrating a gene expression data set of *M. marinum*-infected and noninfected embryos. In this study, our simulations predicted a lowered growth rate for the infected embryos due to changes in essential gene expression as well as monosaccharide transporter gene expression, and a change in histidine metabolism.

Here, we will discuss further improvements and limitations of ZebraGEM 2.0, and briefly discuss the future work.

2.4.1 Blocked reactions

Blocked reactions are reactions that cannot carry any flux due to absence of some or all pathways carrying metabolites toward or away from the reactions. Currently, 1675 out of 3018 (55.5%) of the reactions remain inactive in ZebraGEM 2.0. This number is high in comparison with similar metabolic reconstructions: in Recon 2, 2123 out of 7440 (28.5%) reactions are blocked [30], and in iMM1415, 1294 out of 3726 (34.7%) reactions are blocked [31]. Even if the blocked reactions are currently nonfunctional, we have decided to leave them in ZebraGEM 2.0. This prepares the model for future improvements that can unblock these reactions.

To unblock these reactions, we will need to add a number of missing exchange reactions. These allow the model to import metabolites and excrete waste metabolites. Due to flux balance, the whole metabolic pathway is blocked if excretion or further processing of a metabolite is impossible. One example of such a missing exchange reaction is the exchange reaction for urea; after we added it to the model, it allowed for the production and incorporation into biomass of arginine. For our current needs, further addition of exchange reactions was not needed. Besides that, improvements in the import and export reactions are complicated by three facts. First, there is the food composition, which is not predetermined for free-feeding larvae and adult fish; a solution here would be to add all possible exchange reactions and open or close them depending on fodder composition. Second, there is the unknown factor of exchange with the environment by other means than diet, such as excretion and uptake of metabolites through the skin. Third, there is exchange among cells and tissues of metabolites, such as the uptake of nutrient from the yolk in developing embryos.

Further unblocking of reactions will be achieved by identifying unconnected parts of the network and add the missing metabolic pathways. Such gap-filling can, in part, be automated by finding the minimal set of addition to the network [129, 130, 131], or using novel topology-based methods that can pinpoint missing essential reactions [132]. Such automized gap-filling should be done with care, because the gaps often require reactions that have no or little literature that clearly supports those reactions.

2.4.2 Lipid metabolism

ZebraGEM 2.0 and its predecessor have applied a number of simplifications in the description of lipid metabolism. First, a generic fatty acid is used in most lipid metabolism reactions. Also, the essential lipid linolenic acid has no reaction in the model converting it into this generic fatty acid and hence is not processed further by the model. To further improve the description of lipid metabolism in ZebraGEM 2.0, future description of lipid metabolism should include specific reactions for each type of fatty acid. This improvement would make linolenic acid essential, but because a single reaction would be part of the metabolism of a range of fatty acids, it comes at the cost of increased model size. Most likely, this will double the number of reactions, as the ~600 reactions involved in lipid metabolism will be multiplied by the number of specified fatty acids. This will increase simulation time significantly for some of the modeling techniques, like FVA. The Chinese hamster model iCHOv1 [32], a human platelet model [133], and a human erythrocyte model [90] have parts of lipid metabolism with specified fatty acids and can serve as examples.

An additional factor in lipid metabolism is that many of the associated metabolites are located in the compartment "membrane." This compartment accounts for the plasma membrane, Golgi membrane, endoplasmic reticulum membrane, lysosome membrane, nuclear membrane, and the outer mitochondrial membrane all at once. This compartmentalization into a single compartment does not take into account the required transport processes and associated metabolic processes for such metabolites that take place within the cell. Another effect of this membrane compartment is the tunneling of NADH and NADPH over the membrane due to imbalanced reaction reversibility, as discussed in Reaction Reversibility and Reaction Nature section. We have currently solved this issue by checking reaction reversibility, but a future improvement of the compartmentalization of membrane metabolites into specific membrane parts would solve these problems more accurately.

Improving lipid metabolism is also of interest when looking at the growth conditions of zebrafish. Embryos rely on the abundance of lipids present in the yolk as their source of energy, and as zebrafish are often used for experiments in their embryonal stages, insight into lipid metabolism is relevant. Fraher et al. determined changes in lipid composition of both the yolk and the developing embryo [134]. This study provides interesting information upon which estimates for lipid exchange between embryo and yolk can be made, which can further improve metabolic modeling studies of embryonic stages.

2.4.3 Biomass function and quantitative simulations

The current biomass function is not based upon any data on zebrafish cell composition, but on human and mouse models. Although the metabolites of which a cell consists vary little between animals, as all cells are built from amino acids, nucleic acids, and fatty acids [113], the ratios between the required metabolites can vary as much as 30 million fold [91]. The ratios of biomass precursor metabolites can have a large impact on the model predictions. Therefore, data of zebrafish cell composition, possibly for different cell types, will be of high value for increasing model prediction accuracy. So far, there has been detailed study of lipid composition only [135].

Genome-scale metabolic modeling focuses only on metabolism and hence has a limited scope. For example, 20 genes with a non-normal phenotype in Refs. [118, 119, 120, 121, 122, 123, 124, 125] or ZFIN had no phenotype in ZebraGEM 2.0. They could not be ascribed to blocked reactions, no knockout effect due to the gene being redundant in the model, or the associated reaction diverting flux from the biomass optimization. The optimization for biomass production rate does likely not reflect all the required metabolic outputs of a cell. Alternative objective functions would include specific protein synthesis for antibody producing B-lymphocytes, ATP synthesis for muscle cells, or ROS production upon infection. In addition, bacterial metabolism also plays a role during infection. Therefore, results of *in silico* knockout experiments will deviate from the results of *in vivo* experiments.

A generic problem of flux balance analysis is that it does not consider kinetics and thermodynamics. Gene mutations or knockouts can change the kinetics of metabolic reactions, causing for instance accumulation of toxic compounds. Thermodynamics can also affect the rate of reactions and has been combined with constraint-based methods before [136]. Finally, these genes can cause a phenotype *in vivo* by other means than metabolism, that is, they could be involved in signaling and genetic regulating processes as well, and those aspects are not part of this model.

Last but not least, when using data integration methods, one has to be careful with the distribution of experimental values. As we saw now with our data-integrated simulations, most of the restricting genes were not significantly differentially expressed, which could lead to pinpointing incorrect causes of altered metabolism. The algorithm we used, as well as many others take only a single value for the expression of genes, often just the average; the original distribution underlying that average has to be considered, especially when comparing different situations. Extending

data integration methods for constraint-based metabolic modeling with methods from robust optimization can offer a framework in which such distributions can be taken into account.

Despite these limitations, the improved model combined with the zebrafish embryo data results in the prediction of lowered growth in the case of *Mycobacterium* infection. Furthermore, we showed that metabolism of histidine synthesis was decreased in infected zebrafish embryos. Further improvements on the model as well as the data integration methods and analysis can lead to new applications of ZebraGEM 2.0, such as elucidating yolk and embryo metabolism or exploring the causes of metabolic diseases.

2.5 Supplementary material

The supplementary data can be accessed through the site of the publisher.

Supplementary Data XML-file containing ZebraGEM 2.0

Supplementary Table 2.1 Estimation of the zebrafish biomass function coefficients based on Recon 2, iMM1415 and a human three tissue model.

Supplementary Table 2.2 Maximized effluxes of metabolites out of the original ZebraGEM model.

Supplementary Table 2.3 Results of gene knock out experiments.

Supplementary Table 2.4 Results of metabolic function tests and the performance of the original ZebraGEM and the updated model.

Supplementary Table 2.5 List of reactions with Relative flux range changes (RFRC) greater than 1.

Chapter 3

Cell-matrix adhesion affects cell motility mode: from short-term persistent to long-term subdiffusive modes

Leonie van Steijn, Clément Sire, Loïc Dupré, Guy Theraulaz, Roeland M.H. Merks*

Abstract

Lymphocytes have been described to perform different motility patterns such as Brownian random walks, persistent random walks, and Lévy walks. Depending on the conditions, such as confinement or the distribution of target cells, either Brownian or Lévy walks lead to more efficient interaction with the targets.The diversity of these motility patterns may be explained by an adaptive response to the surrounding extracellular matrix (ECM). Indeed, depending on the ECM composition, lymphocytes either display a floating motion without attaching to the ECM, or sliding and stepping motion with respectively continuous or discontinuous attachment to the ECM, or pivoting behaviour with sustained attachment to the ECM. Moreover, on the long term, lymphocytes either perform a persistent random walk or a Brownian-like movement depending on the ECM composition. How the ECM affects cell motility is still incompletely understood. Here, we integrate essential mechanistic details of the lymphocyte-matrix adhesions and lymphocyte intrinsic cytoskeletal induced cell propulsion into a Cellular Potts model (CPM). We show that the combination of *de novo* cell-matrix adhesion formation, adhesion growth and shrinkage, adhesion rupture, and feedback of adhesions onto cell propulsion recapitulates multiple lymphocyte behaviours, for different lymphocyte subsets and various substrates. With increasing attachment area and increased adhesion strength, the cells' velocity and persistence decreases. Additionally, the model can predict short-term persistent with long-term subdiffusive

*Submitted as Leonie van Steijn et al. "Computational modelling of cell motility modes emerging from cell-matrix adhesion dynamics", available on bioRxiv, `https://doi.org/10.1101/2021.06.09.447692`

motility, showing a pivoting motion. For small adhesion areas, we observe that the spatial distribution of adhesions influences cell motility. Small adhesions at the front allow for more persistent motion than larger clusters at the back, despite a similar total adhesion area. In conclusion, we present an integrated framework to simulate the effects of ECM proteins on cell-matrix adhesion dynamics. The model reveals a sufficient set of principles explaining the plasticity of lymphocyte motility.

Author summary

During immunosurveillance, lymphocytes patrol through tissues to interact with cancer cells, other immune cells, and pathogens. The efficiency of this process depends on the kinds of trajectories taken, ranging from simple Brownian walks to Lévy walks. The composition of the extracellular matrix (ECM), a network of macromolecules, affects the formation of cell-matrix adhesions, thus strongly influencing the way lymphocytes move. Here, we present a model of lymphocyte motility driven by adhesions that grow, shrink and rupture in response to the ECM and cellular forces. Compared to other models, our model is computationally light making it suitable for generating long term cell track data, while still capturing actin dynamics and adhesion turnover. Our model suggests that cell motility is affected by the force required to break adhesions and the rate at which new adhesions form. Adhesions can promote cell protrusion by inhibiting retrograde actin flow. After introducing this effect into the model, we found that it reduces the cellular diffusivity and that it promotes stick-slip behaviour. Furthermore, location and size of adhesion clusters determined cell persistence. Overall, our model explains the plasticity of lymphocyte behaviour in response to the ECM.

3.1 Introduction

Lymphocytes patrol in tissues and are recruited to infected areas to detect and clear the area of pathogens and cancer cells. The type of walk by which lymphocytes patrol determines the efficiency of finding their target depending on the environment [137, 138, 139, 56, 140]. In the absence of obstacles, Lévy walks and persistent random walks outperform Brownian walks. Lévy walks are characterized by long strides in their trajectories such that they cover larger areas than Brownian walks. In environments crowded with obstacles Brownian walks perform better, as the more compact trajectory leads to more thorough local exploration [137]. Even more

local exploration follows from subdiffusive random walkers, which diffuse less far than could be expected from their speed and persistence. Consistent with these characteristics, in the densely populated lymph nodes T lymphocytes perform Brownian walks [53, 54] or persistent random walks [55]. In the brain tissue, T cells perform Lévy walks [56]. In pancreatic islets CD4+ T cells perform subdiffusive random walks, whereas CD8+ T cells perform confined random walks [57]. The characteristics of these different types of motion, including speed distribution and mean squared displacement (MSD), determine how efficiently lymphocytes can find their targets *in vivo*. Therefore it is key to understand what factors give rise to these different types of walks.

The plasticity of lymphocyte motility behaviour is dictated both by environmental factors and by cell intrinsic features [141, 142]. An *in vitro* study has shown that the type of extracellular matrix (ECM) used as cell culture substrate affects the motility patterns of B lymphocytes, possibly due to the attachment strength [143]. On fibronectin, B lymphocytes show higher diffusivity and more effective displacement than on collagen IV substrates where cells move more slowly. The B lymphocytes formed larger adhesive connections with fibronectin than with collagen IV, and on fibronectin the cells changed shape more rapidly than on collagen IV. Similar effects have been found for T lymphocytes. The majority of cells on a casein substrate move through multiple, distinct and temporary adhesion zones, i.e., walking motility, whereas on ICAM-1 substrates, the majority of cells make one continuous contact zone with the substrate, i.e., sliding motility [144].

Cells also show large individual variation in their motility patterns. On fibronectin, individual B lymphocytes showed either floating-like behaviour with little attachment, dynamic attachment leading to stepping/walking behaviour, or sustained attachment leading to cells pivoting around their adhesive area [143]. Similarly, T lymphocytes showed either walking, mixed or sliding behavior, with frequencies depending the type of culture substrate [144].

It is still poorly understood what causes, on the one hand, the consistent differences in motility modes between culture substrates, and on the other hand, the large individual differences between cells on the same type of substrate. To answer this question, here we propose a simplified mathematical model of cell motility and the adhesive interaction with the ECM.

Previous modeling studies have already provided useful insight into this problem. Copos et al. [145] asked what causes the cellular extension and retraction cycles driving the motility of *Dictyostelium discoideum* cells.

They modelled *D. discoideum* movement in a force-based model. Depending on the density of adhesion binding sites in the substrate, or the strength of these adhesions, the model displayed different motility types. For low densities of adhesion binding sites in the substrate and low adhesion strength, gliding motility was observed. Increasing the density of the binding sites or the adhesion strength led to a stepping motility mode of reduced speed. For the highest adhesion densities or adhesion strengths the cells became stationary. Thus the cells moved faster in the gliding motility mode than in the stepping motility mode. Although this agreed with preliminary experimental results on *D. discoideum* cells[145], these results contradict observation in lymphocyes: T cells move faster in stepping motility mode than in gliding motility mode [144]. Furthermore, as a one-dimensional model, it cannot produce two-dimensional cell tracks and it is computationally too heavy for producing the large amounts of cell track data required for our purpose.

Phase field models make it possible to study the effect of ECM on cell motility in two dimensions [65, 67]. In [65], the model includes actin polymerization, explicit dynamics of adhesion site formation and substrate compliance. Simulated cells displayed a gliding motion when substrate stiffness was high, the protrusion strength was large and adhesions formed at a high rate. At intermediate substrate stiffnesses with sufficiently high protrusion strength and intermediate adhesion formation, the cells displayed a stick-and-slip motion. Yet, the computational costs are still too high for the length and number of cell track data we require for statistical analysis.

Yu et al. [146] introduced a computationally efficient, coarse-grained model to study long term cell persistence. The model considered spheroid cells with a fixed pool of focal adhesions. These adhesions were assumed to be widely dispersed within the cells for soft substrates and more narrowly dispersed for rigid substrates. The increased persistence times on rigid substrates led to durotaxis, i.e. preferential movement towards stiffer substrates. However, Yu et al. imposed a direct dependence of cell persistence on adhesion distribution. In our work, we hope to explain this relation emerging from first principles.

Thus, previous models are either too computationally expensive or do not model the effect of adhesion on the microscopic level. The Cellular Potts model is conceptually closely related to phase fields model, but is computationally much lighter. The Act model [73], a recent extension of the Cellular Potts model, provides of phenomenological model of actin dynamics. Interestingly, this model can already display multiple motility

modes: Simulated cells show intermittent (stop-and-go) or persistent random walks. An in-depth characterization revealed that the model displays universal coupling between speed and persistence, and specifically that speed increases linearly with protrusion strength, whereas persistence time increase exponentially with protrusion strength [147].

Here we extend the Act model with cell-ECM interactions. The model combines coarsed-grained actin dynamics, with simplified dynamics of adhesion turn-over and detachment, resulting in a diverse palette of cell motility. In a second version of the model, the cell-ECM adhesions promote cell protrusion by inhibiting retrograde actin flow. Our model can simulate cell motion with sufficient detail on the location and size of adhesive patches, while being computationally light enough for statistical analysis of cell motility. With the actin component and cell-matrix adhesion component of the model, we are able to reproduce a variety of cell motion types, similar to the behaviour seen in other models that also include those two components [145, 65, 67]. In addition to persistent random walks and ballistic cell motility, the extended model can also predict anomalous diffusion with long-term subdiffusive behaviour, showing all three phases of lymphocyte motility on fibronectin found in the experimental work by Rey-Barroso et al.[143]. Our model shows that simple cell-ECM interactions can drastically alter cell motility. Thus, adhesion dynamics can play a key role in the plasticity of motility in response to ECM composition.

3.2 Results

In this section, we present how we model the dynamics of cell-matrix adhesions. We show that this model can reproduce a wide range of lymphocyte motility modes. Next, we extend the model with feedback from the adhesions onto the actin polymerization force and show that we can capture more dynamic motility behaviours. Overall, our model recapitulates the diversity of lymphocyte motility modes and provides insight into the mechanisms underlying such behavioural diversity.

3.2.1 Modelling cell-matrix adhesions

Our computational model is based on the Act model [73], an extension of the Cellular Potts model (CPM, [148, 149]) with an actin-inspired feedback mechanism that results in realistic cell shapes and cell polarization. In short, this extension keeps track of recent “actin activity" through Act values at a subcellular level, and cell protrusion at sites with locally high Act values is favoured. Two important parameters for this are λ_{Act}, the weight of the Act

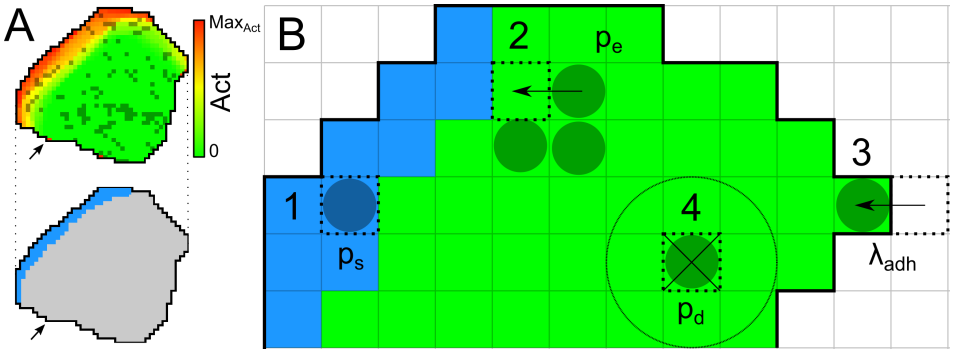


FIGURE 3.1: **Overview of the adhesion processes within the model.** A) Top: overview of a simulated cell. Red to yellow shading indicates the Act-level of each grid point. Darker coloured grid points contain an adhesion. Bottom: same cell with the region where new adhesions can form as the local Act-levels exceeds the $0.75\,\mathrm{Max}_{Act}$ threshold depicted in blue. Both: Arrows point to area with one grid point with high Act-level due to a recent extension of the cell (top, red), but the geometric mean of Act-levels does not exceed the threshold and hence new adhesions cannot form there (bottom, grey). B) Visual summary of adhesion processes. Dark coloured circles indicate grid points containing an adhesion. 1) New adhesions can form spontaneously with probability p_s at cell grid points where the geometric means of Act values exceeds the threshold of $0.75\,\mathrm{Max}_{Act}$ (blue region). 2) An adhesion patch can grow by Eden growth. A random neighbour of an adhesion site is selected. When it does not contain an adhesion yet, the patch extends into that grid point with probability p_e. 3) Adhesions can disbond spontaneously, depending on the number of neighbouring grid points without adhesions and probability p_d. 4) When cell retraction would break an adhesion, this is paired with a energy cost λ_{adh}.

model that can be interpreted as the maximum protrusive force of the actin network, and Max_{Act}, the maximum Act value, interpretable as the lifetime of an actin subunit within the actin network (see Table 3.1). In addition, our model takes cell-matrix adhesion into account and reflects the dynamic processes of such adhesions (Fig 3.1). We will shortly explain the addition of cell-matrix adhesions below. For further details, see the Method section.

Modelled cell-matrix adhesions are monitored at a subcellular level. A subcellular CPM grid point can either contain an adhesion or not. A single grid point is approximately 600 nm wide, considering that simulated cells contain approximately 1000 CPM grid points each, and that B cells and T cells on a substrate cover an area of approximately 360 μm^2[143, 150]. Observations show that single adhesion units in lymphocytes are

approximately 100 nm in diameter [150], so a single adhesion grid point in the model represents a small number (≤ 5) of adhesion units, considering adhesion unit density of 5 clusters/μm^2 [150].

The formation of new adhesions depends on actin polymerization, membrane protrusion and the distribution of integrins at the leading edge of the cell [151, 152, 153]. As actin activity and the leading edge are marked in the Act model by grid points with high Act values, we let new adhesions appear at grid points with a locally high Act-level: i.e. when the geometric mean of Act-level $Act(y)$ of $NB_\sigma(x)$, the Moore neighbourhood of grid point x restricted to the same cell as point x, exceeds a threshold $\left(\prod_{y \in NB_\sigma(x)} Act(y)\right)^{\frac{1}{|NB_\sigma(x)|}} \geq 0.75\,\mathrm{Max}_{Act}$, grid point x receives an adhesion with probability p_s (Fig 3.1A and 3.1B, process 1).

Once an adhesion has formed, it can either expand into an adhesion patch, or disbond. Patch expansion happens due to some membrane properties: membrane fluctuations lessen with membrane-matrix adhesion and hence allow for more integrins to bind the matrix [154], and the curvature of the membrane favours aggregation of integrins [155, 156]. We choose to model the effects of these properties in a phenomenological way, guided by the observations of Jacobelli et al. [144]. They report radial expansion of adhesion patches with some bias in the direction of the cell front. The Eden-growth model [157] gives radially expanding spherical objects, so we decided to use an Eden-like growth process to model adhesion patch expansion. During the update of the adhesion layer of the model, whenever a grid point with adhesion is selected, we also randomly select a neighbour. If that neighbour does not contain an adhesion, it gains one with probability p_e (Fig 3.1B, process 2).

Cell-matrix adhesions are not everlasting and they can disbond spontaneously or by force. We model two distinct disbonding processes. First, we consider a general and spontaneous disbonding of adhesions. As adhesion molecules undergo continuous turnover and experience stress from myosin-II, adhesions are broken constantly. Hence, we associate a probability with this process. Again, following the observations from [144] that patches dissolve concentrically due to the involvement of myosin-II, we let the disbonding probability depend on the local neighbourhood of the adhesion. An adhesion grid point surrounded by other adhesion grid points is likely within the centre of a patch and, hence, less likely to spontaneously disbond, whereas a single adhesion grid point with no neighbouring adhesions is quite likely to disappear. The probability that an adhesion site disbonds is $p_d \cdot \left(\frac{|\{nb \in NB(x) | Adh(nb)=0\}|}{|\{nb \in NB(x)\}|}\right)^2$, where $NB(x)$ indicates the Moore neighbourhood of grid point x (Fig 3.1B, process 3).

The second process that disbonds adhesions is adhesion rupture through retraction of the cell. We model this rupture only at the edge of the cell, where contraction forces of the cell can break bonds. Integrin bonds are known to show catch-slip bond behaviour, meaning that initially the bond strengthens with increase of force, but will still break if enough force is applied. Here we neglect this specific behaviour and associate a single required energy cost of λ_{adh} with the rupture of adhesions at the retracting edge (Fig 3.1B, process 4).

All in all, our model extension for adhesions is quite simple and computationally light. All adhesion dynamics are governed by the four parameters p_s, p_e, p_d and λ_{adh}. An overview of all the relevant parameters is shown in Table 3.1.

3.2.2 Adhesions lead to crawling and pivoting motions

From the newly introduced parameters p_s,p_e,p_d and λ_{adh}, the two parameters p_s and λ_{ahd} are most directly associated with properties of the ECM. We can interpret p_s, the probability with which new adhesions arise at sites with high actin activity, as multiple biological processes. One process is the rate at which integrin molecules bind to their ligands in the extracellular matrix. Higher rates would translate into higher p_s. Another process is the availability of integrins to the cell front. Transportation towards the cell front, integrin production and breakdown can thus all influence p_s. For λ_{adh}, there are two complementing interpretations. As λ_{adh} is defined as the energy required to break an adhesion, it describes both the binding affinity between integrins and their ligands, as well as the number of integrins bound in a single adhesion complex. Since we are interested in how lymphocytes adapt their behaviour to the ECM, we first look at the influence of these two ECM-associated parameters.

The proposed model displays various motility types. Cells crawl when λ_{adh} is low to moderate (Fig 3.2A,3.2C, Supplementary Video S1). Cell persistence decreases as λ_{adh} increases (Fig 3.2B,3.2D, Supplementary Video S1). When both λ_{adh} and p_s are high, i.e., when adhesions easily form and require much energy to break, cells will remain stuck in place on the matrix. However, they can still make protrusions around them, resembling a pivoting motion (Fig 3.2D, Supplementary Video S1).

Comparing these four parameter settings (Fig 3.2), the cell area that is covered with adhesions is mainly regulated by the parameter p_s. The velocity of the cell is fluctuating a lot more than the cell adhesion area, but is mostly affected by the parameter λ_{adh}. These observations are in large agreement with the observations of Rey-Barroso et al. that B cells on

TABLE 3.1: List of parameters involved in adhesion dynamics and values used for simulations.

Parameter	Description	Values Figs. 2-4	Figs. 6,7	Fig. 8
λ_{Act}	Weight of the Act-extension, the maximum protrusive force induced by actin polymerization	240	120, 240	240
Max_{Act}	Maximum value of the Act-field, actin lifetime	120	120	120
-	Act-value threshold above which adhesion formation is possible	0.75	0.75	0.75
p_s	Probability of new adhesion formation	0.004-0.020	0.001-0.004	0.003, 0.001
p_e	Probability of neighbouring grid site to become adhesion site if not already so.	0.0055	0.0055	0.0015, 0.004
p_d	Scaling of probability of disbonding adhesion site	0.0064	0.008	0.0032
λ_{adh}	Energy required to rupture adhesion upon retraction of the cell	20-100	20-100	60
f	Prefactor for the adhesion feedback onto Act model	-	b-1	b-1
b	Base value of f in absence of adhesions	-	0.5	0.5
s	Adhesion area fraction saturation threshold above which $f = 1$	-	0.1	0.12

fibronectin with dynamic adhesion surfaces showed walking behaviour, and cells with large and sustained adhesion surfaces displaced very little as they were unable to relocate the adhesion area [143]. They are also in agreement with the observation of Jacobelli et al. that T cells displaying a gliding motion with higher adhesion area have lower speed compared to cells with a walking motion with lower adhesion area [144].

3.2.3 Adhesions slow down cell motion and diminish dispersal

The examples shown in Fig 3.2 and Supplementary Video S1 indicate that higher adhesion area is correlated with lower speed and lower cell diffusivity. To further look into this relationship, we averaged the cell velocity and adhesion area of 1000 independent runs for different combinations of p_s and λ_{Adh} (Fig 3.3A). Increasing the value of p_s increases the average adhesion area, while increasing the value of λ_{Adh} decreases cell velocity. Moreover, p_s and λ_{Adh} seem to have a synergistic effect. The drop in instantaneous cell speed (from highest to lowest in Fig 3.3A about 50% smaller) is modest compared to the drop in diffusion coefficient (Fig 3.3B, about 380% smaller). The diffusion coefficient drops rapidly with increasing λ_{Adh}.

Highly adhesive cells show subdiffusive behaviour

To investigate the drop of the diffusion coefficient, we analysed the mean squared displacement and fitted the values with a persistent random walker model [47, 158]:

$$MSD(t) = 4\frac{v_{th}^2}{\gamma_1^2}\left(\gamma_1 t - 1 + e^{-t\gamma_1}\right), \tag{3.1}$$

with v_{th} the walker's velocity and γ_1 its persistence time. However, this description fails at the short time scale, at which the CPM is mainly driven by the random fluctuation in grid points. Hence, we extended Eq. 3.1 with translational diffusion [159]:

$$MSD(t) = 4\frac{v_{th}^2}{\gamma_1^2}\left(\gamma_1 t - 1 + e^{-t\gamma_1}\right) + D_T t \tag{3.2}$$

Eq. 3.2 gives good fits, except for the higher $\lambda_{adh} = 80, 100$. For lower λ_{adh}, we obtained the persistence time from fitting Eq. 3.2, (data not shown): The larger drop in dispersal rate compared to instantaneous cell speed can be explained by loss of persistence with higher adhesion energies and adhesive areas.

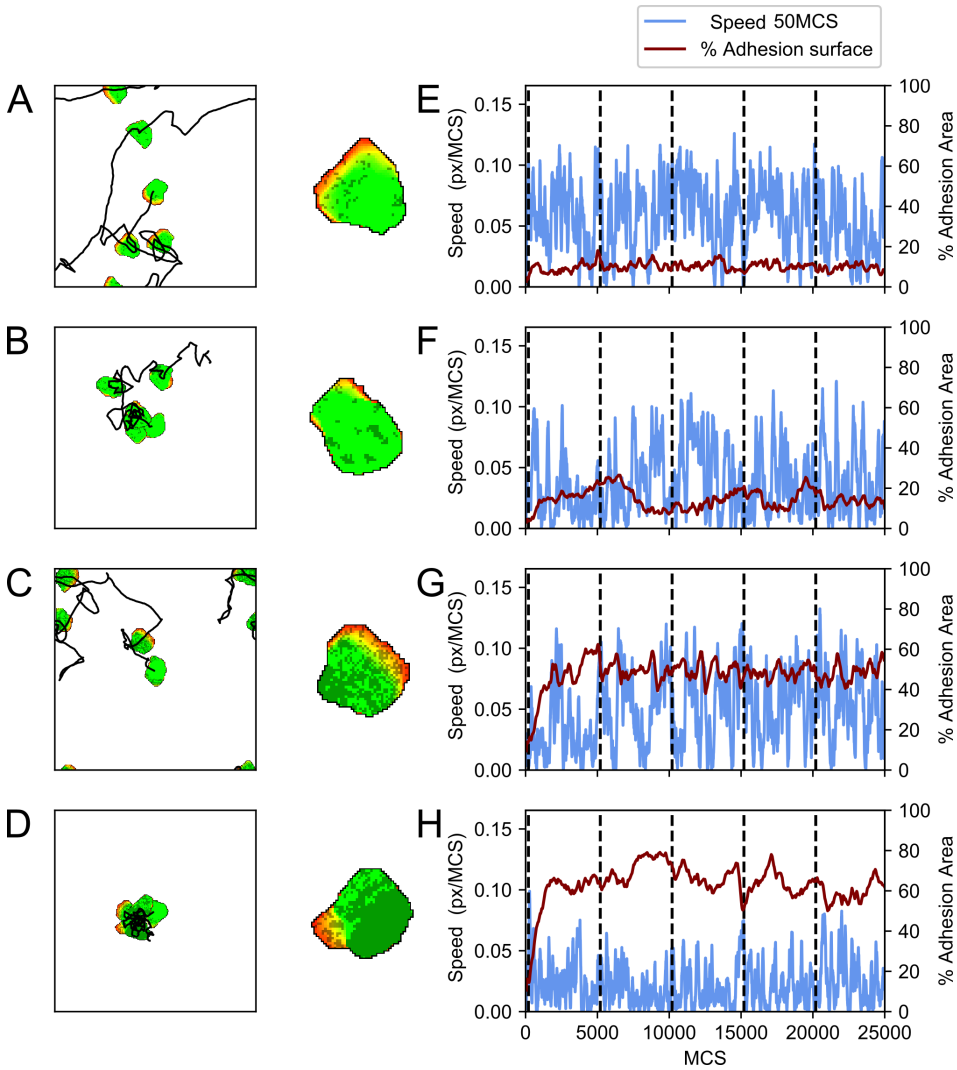


FIGURE 3.2: **Simulations of the model showing different motility types.** On the left, a display of a single cell at 5000 MCS interval snapshots combined with the cell centre's trajectory. Each trajectory starts in the centre of the field and periodic boundaries are used. In the middle, a close-up of the cell with the adhesions displayed in a darker colour. On the right, a plot of the cell's velocity and percentage of the cell's area containing adhesions corresponding to the track on the left. Vertical dashed lines indicate the times of the snapshots on the left. Parameters are: A) $\lambda_{adh} = 20$, $p_s = 0.004$, B) $\lambda_{adh} = 100$, $p_s = 0.004$, C) $\lambda_{adh} = 20$, $p_s = 0.02$, D) $\lambda_{adh} = 100$, $p_s = 0.02$. Furthermore, $p_d = 0.0008$ for A, B, C and D. These simulations are also available as Supplementary Video S1.

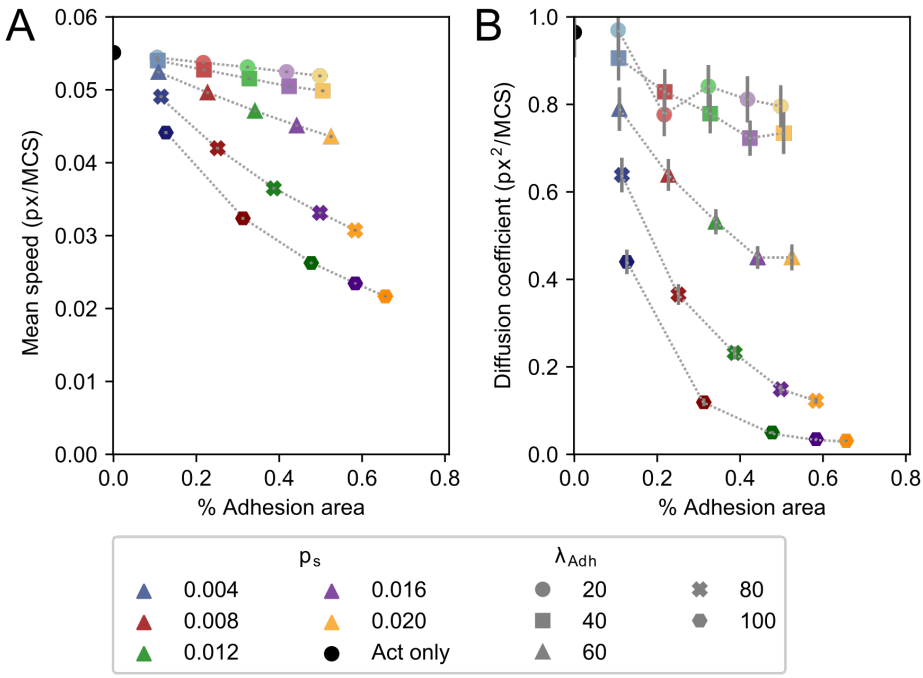


FIGURE 3.3: **Mean velocity, diffusion coefficient, and mean adhesive area change by increasing p_s and λ_{adh}.** Mean velocity (A) and diffusion coefficient (B) plotted against mean percentual adhesion area for different values of parameters p_s and λ_{adh}. Each dot represents the mean of 1000 independent simulations. Different colours indicate different p_s, where shades from light to dark and marker symbol indicate $\lambda_{Adh} \in \{20, 40, 60, 80, 100\}$. For reference, the mean velocity and diffusion coefficient of the Act model without any adhesions are indicated by a black circle. Error bars indicate 95% confidence intervals.

TABLE 3.2: Fitted values of α from Eq. 3.3 for different values of λ_{adh} and p_s

Parameters	α
$\lambda_{Adh} = 20, p_s = 0.004$	1.019
$\lambda_{Adh} = 20, p_s = 0.020$	1.013
$\lambda_{Adh} = 100, p_s = 0.004$	1.024
$\lambda_{Adh} = 100, p_s = 0.020$	1.257

For the parameter regime where Eq. 3.2 was a bad fit, we increased the initialization period left out of the MSD computation, in order to compute the MSD of cells closer to their dynamic equilibrium in both Act model dynamics as well as adhesion-extension dynamics. This barely improves the fit and suggests that cell motion in this regime cannot be correctly described by a persistent random walker with translational diffusion.

In [160] a fractional Klein-Kramers process was suggested as a good description of transformed Madin-Darby canine kidney cell motion. They fitted their data with

$$MSD(t) = 4v_{th}^2 t^2 E_{\alpha,3}(-\gamma_\alpha t^\alpha) + (2\eta)^2, \tag{3.3}$$

where $E_{\alpha,3}$ is the generalized Mittag-Leffler function and η is a noise term. The case where $\alpha = 1$ results in Eq. 3.1 except for the noise term. Since we already determined that translational diffusion plays significant role in the short-time scale of the CPM, we replaced the noise term with the term for translational diffusion, obtaining:

$$MSD(t) = 4v_{th}^2 t^2 E_{\alpha,3}(-\gamma_\alpha t^\alpha) + D_T t. \tag{3.4}$$

which reduces to Eq. 3.2 for $\alpha = 1$. This parameter α describes the long-term diffusive behaviour. For $t \to \infty$, Eq. 3.3, and by extension Eq. 3.4, can be approximated by $MSD(t) \sim \frac{4D_{th}t^{2-\alpha}}{\Gamma(3-\alpha)}$ [160]. So for $\alpha > 1$, long-term behaviour is subdiffusive, whereas for $\alpha < 1$, long-term behaviour is superdiffusive.

In most cases where Eq. 3.2 fits well, we obtain $\alpha \approx 1$ (Table 3.2). However, for the cases were Eq. 3.2 fits badly, Eq. 3.4 has a better fit and $\alpha > 1$ (Fig. 3.4, Table 3.2). This corresponds to the cells stuck to the matrix and pivoting around their adhesion patch, as they are moving persistently on a local scale (a single protrusion front), but moving subdiffusively on a longer timescale.

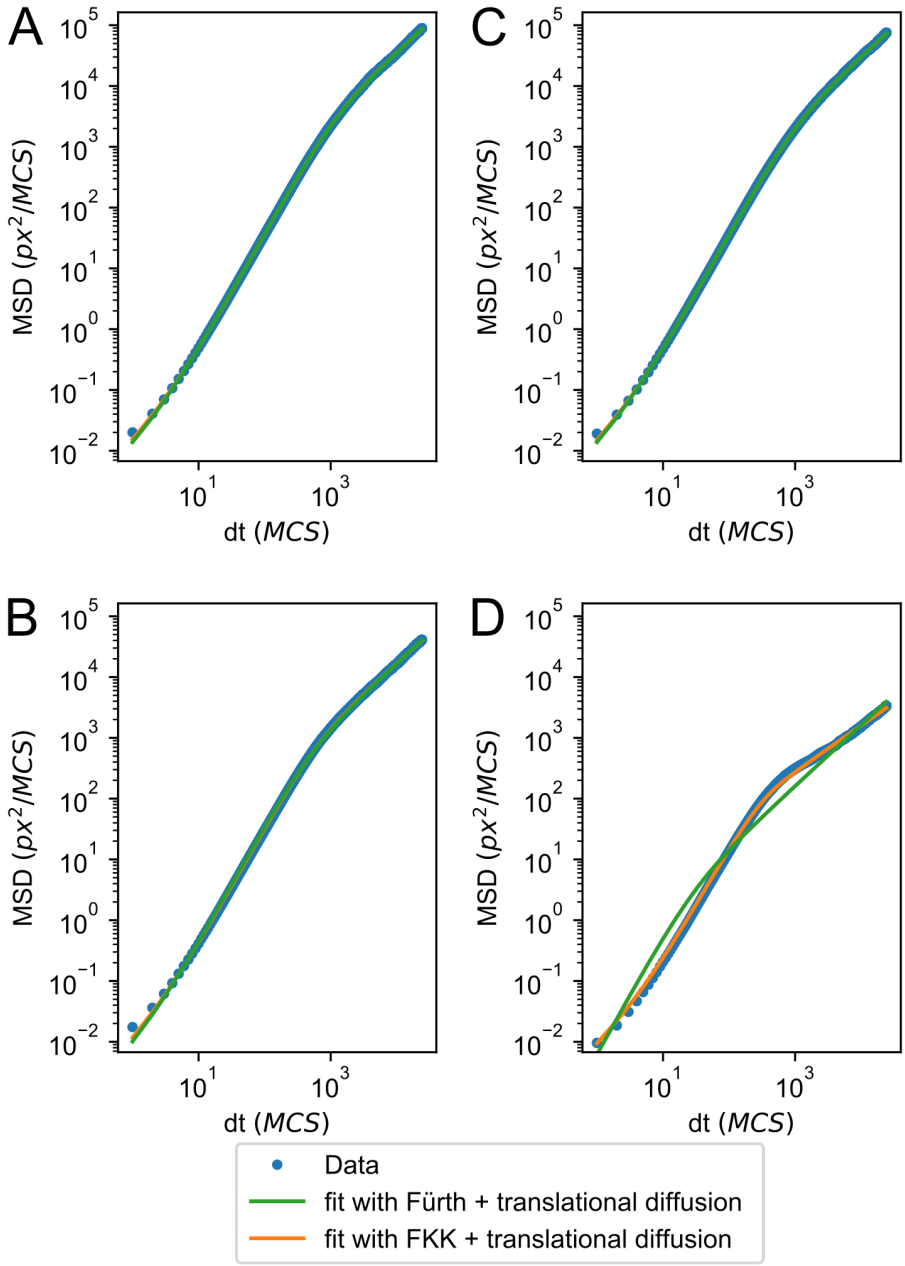


FIGURE 3.4: **MSD fits either a persistent random walker or a subdiffusive persistent random walker.** Log-log plot of MSD for the four scenarios in Fig. 3.2, with fits of Eqs. 3.2 and 3.4. Parameters are: A) $\lambda_{adh} = 20$, $p_s = 0.004$, B) $\lambda_{adh} = 100$, $p_s = 0.004$, C) $\lambda_{adh} = 20$, $p_s = 0.02$, D) $\lambda_{adh} = 100$, $p_s = 0.02$.

3.2.4 Modelling feedback of adhesions onto propulsion efficiency

This current model cannot explain the observation that B cells with a low adhesive area or no adhesive area on a fibronectin substrate show low displacement compared to cells with dynamic attachment [143]. Adhesions allow actin polymerization to lead to more efficient protrusions [161, 162, 163], as the force generated by the actin polymerization is transferred onto the matrix via the adhesion complex, instead of leading to treadmilling. When more of the actin network is connected to integrin complexes, a greater force resulting from actin polymerization can be transferred to the matrix. We mimic such behaviour using positive feedback between the adhesions and actin polymerization. We model this by defining a prefactor f which dynamically alters the weight of the Act-extension, and hence the propulsion force. For simplicity, we assume that protrusion efficiency will increase linearly with the cell's total adhesion area and will reach a point of saturation. Hence, we define:

$$f = \begin{cases} b + \frac{1-b}{s}\frac{A_{adh}(i)}{A(i)} & \text{if } \frac{A_{adh}(i)}{A(i)} \leq s \\ 1 & \text{if } \frac{A_{adh}(i)}{A(i)} > s \end{cases},$$

with b the baseline protrusion efficiency and s the saturation adhesion area. A schematic overview is shown in Fig 3.5. We expect that the feedback between adhesion area and propulsion strength only affects cell motility when the adhesion area is below or near the saturation point s.

New behaviours ranging from slow cells to stick-slip

Here, we will look at parameter combinations which result in adhesion areas below or around the saturation threshold s. We choose $s = 0.1$ and $b = 0.5$, and from the previous simulations we know that p_s is the main parameter controlling adhesion area, so we chose $p_s \leq 0.004$.

We observed different types of behaviour depending on p_s (Fig 3.6, Supplementary Video S2). For very low values of p_s (Fig 3.6A), cells have only a small number of tiny adhesion patches and thus very small adhesive area. Furthermore, they disperse little (Fig 3.6A). Despite their low dispersion, their motion can still be described well with Eq. 3.2 or with Eq. 3.4 with $\alpha = 0.974$, so the type of motion can still be classified as a persistent random walk, albeit with lower persistence time. When $p_s = 0.004$, the mean adhesive area is approaching the saturation point. However, the diffusion and persistence are lower compared to the model without the adhesion-protrusion feedback, but velocity is comparable (Fig 3.6B). In between these

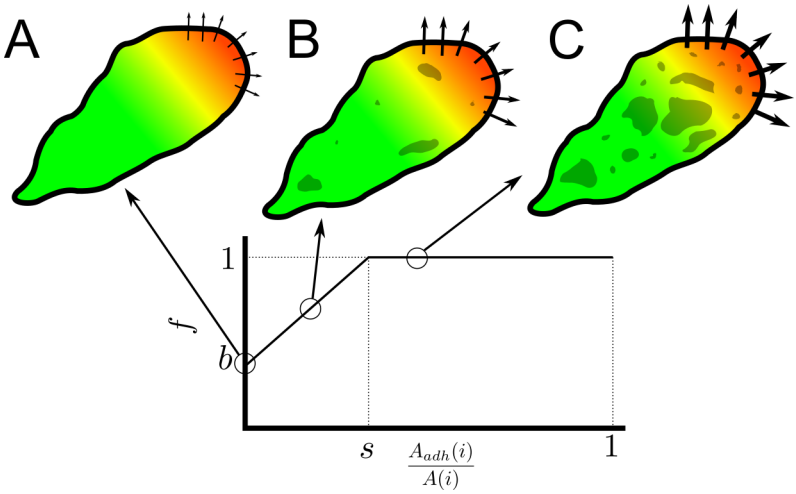


FIGURE 3.5: **Schematic representation of the adhesion propulsion feedback.** Colour schemes are similar to Fig 3.1A. Arrow width corresponds to the effective protrusion strength $f\lambda_{Act}$. A) In the absence of adhesions, the propulsion prefactor f is equal to the base level b. B) Below the saturation point s, f increases linearly with adhesion area. C) Above the adhesion area saturation point s, prefactor f, and thus effective protrusion strength $f\lambda_{Act}$, are maximal.

two adhesive regimes, there is the possibility of stick-slip behaviour (Fig 3.6C), with clear bursts of adhesive area coupled with increased speed.

There is a clear difference between the model with and without feedback of adhesion area onto propulsion strength. For low adhesive areas, we expect the two models to converge to the Act model with an effective propulsion strength equal to the set λ_{Act} for the version without feedback, and an effective propulsion strength of $b\lambda_{Act}$ for the version with feedback. Looking at average velocity and the diffusion constant, this expectation is met (Fig 3.7). Similar to the results in Fig. 3.3, the effect of adhesion energy λ_{adh} is larger on the diffusion of cells than on their velocity for the model both with and without feedback between adhesion and propulsion. Remarkably, there is a slight difference in mean adhesion area between the models. This small effect is likely due to closing the feedback loop between adhesions and propulsion, as *de novo* adhesion formation depends on Act-front presence, which becomes less pronounced when adhesion area is low in the model with feedback.

Our model with feedback shows that cells with low adhesive area have low displacement compared to cells with higher adhesive areas. Overall, by adjusting the parameters p_s and λ_{adh}, the model is able to reproduce

the three behaviours of B cells observed on fibronectin: low attachment with low displacement, dynamic adhesion area with high displacement and sustained attachment with low displacement [143].

3.2.5 Adhesion growth dynamics change persistence time

So far, we have only looked at the effects of new adhesion formation p_s and adhesion strength λ_{adh} on adhesive area and cell motility. However, the adhesive area is also in part determined by the two parameters p_e for adhesion growth and p_d for disbonding adhesions. To gain further insight in the impact of adhesion dynamics on cell motility, we also explored parameter settings resulting in the same adhesion area but with different adhesion cluster size distributions, by varying the formation rates for new adhesions (p_s) and adhesion growth (p_e). Fig. 3.8 shows the results of two such parameter settings resulting in the same average adhesion area. One parameter set obtains its adhesive area mostly through the formation of new adhesions ($p_s > p_e$, Fig. 3.8A: blue), whereas the other more rapidly grows out adhesion clusters ($p_s < p_e$, Fig. 3.8A: orange, see also Supplementary Video S3). This results in different cluster size distribution (Fig. 3.8B), with only small clusters when $p_s > p_e$ (blue line), and small clusters combined with a few large ones when $p_s < p_e$ (orange line). Aside from the different adhesion cluster size distribution, cell motility also differs between the two situations.

First, the velocity distribution has a slightly higher mean for the many-small-cluster (blue) setting, but also appears more bimodal than the few-large-cluster (orange) setting (Fig. 3.8). Analysing the MSD shows more differences: the few-large-cluster (orange) setting shows an earlier start of the final linear regime. The onset of this regime corresponds to the persistence time, which we obtained by fitting the MSD with Eq. 3.2 as well as Eq. 3.4. The fitted persistence times confirm this observation: the few-large-cluster (orange) setting has about 25% lower persistence time than the many-small-cluster (blue) setting. So not only the total adhesion area influences cell motility, but also how that area is distributed over adhesion clusters and where those clusters are located. This further shows that the dynamics of cell-matrix adhesion influence cell motion and can be a key component of cell motion plasticity.

The distribution of the adhesion clusters over the cell is reminiscent of the difference between walking and gliding T cells [144]. The sliding T cells had a large contact area at the cell front, quite similar to the blue setting in Fig. 3.8, which derives its adhesion area mainly from the Act-dependent formation of new adhesions. The walking T cells, in constrast, showed

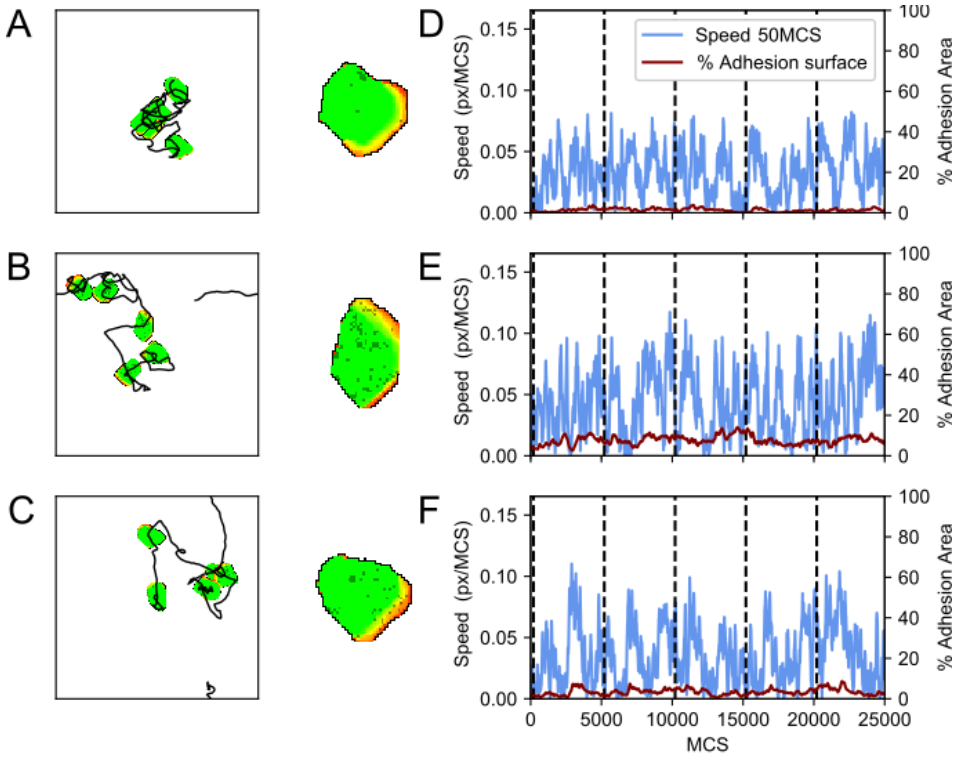


FIGURE 3.6: **Simulations of the model with adhesion-propulsion feedback.** On the left, a display of a single cell at 5000 MCS intervals combined with trajectory of the cell centre. Each trajectory starts in the middle and periodic boundaries are used. In the middle, a close-up of the cell with the adhesion displayed in a darker colour. On the right, a plot of the cell's velocity and percentage of the cell's area containing adhesions corresponding to the track on the left. Parameters are: A) $\lambda_{adh} = 100$, $p_s = 0.001$, B) $\lambda_{adh} = 100$, $p_s = 0.004$, C) $\lambda_{adh} = 100$, $p_s = 0.0025$. Furthermore $p_d = 0.001$ for A, B and C. These simulations are also available as Supplementary Video S2.

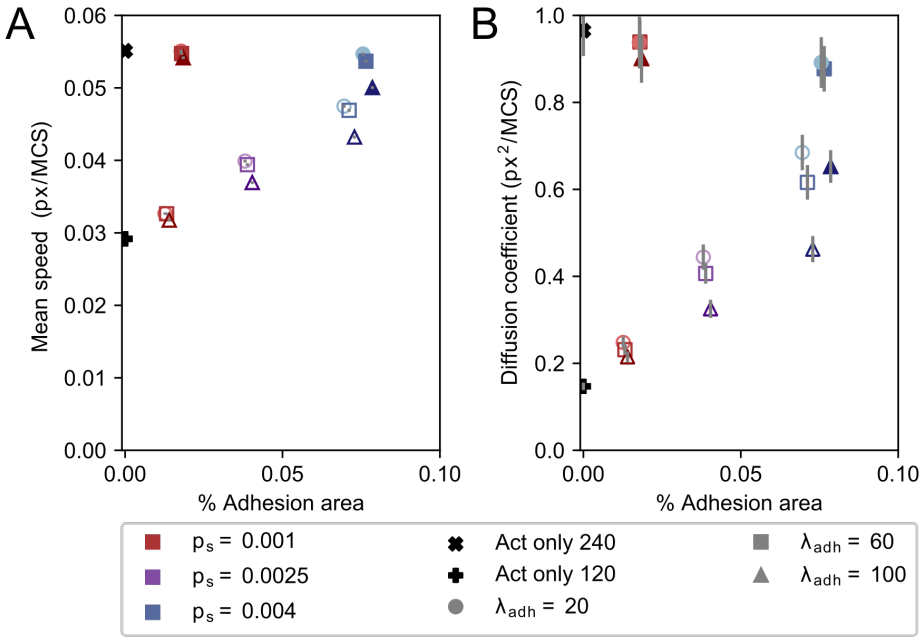


FIGURE 3.7: **Mean velocity, diffusion coefficient and adhesion area differ between the model with and without adhesion-propulsion feedback** Mean velocity (A) and diffusion coefficient (B) plotted against mean percentual adhesion area for different values of parameters p_s and λ_{adh}. Each dot represents the mean over 1000 independent simulations. Filled symbols are results from the model without adhesion-propulsion feedback, and empty symbols show results from the model with adhesion-propulsion feedback. Different colours indicate different p_s, where shades from light to dark and marker symbol indicate $\lambda_{Adh} \in \{20, 60, 100\}$. Error bars indicate 95% confidence intervals.

contact area also at the rear of the cell, and had multiple distinct contact areas. The orange setting from Fig. 3.8, where adhesion area mainly grows by adhesion expansion rather than new adhesions, resembles this adhesion distribution over the cell. However, the mean speed found in [144] is higher for walking cells and lower for gliding cells, opposite to the two parameter settings shown here. Nonetheless, our model suggests that the different processes by which adhesions form, such as *de novo* at the cell front or by extending existing adhesion patches, can underlie the differences between walking and gliding cells.

3.3 Discussion

Here, we have presented an extension of the CPM-Act model with dynamic cell-matrix interactions. In this model, cell-matrix adhesions can develop *de novo* in an Act-dependent manner, and adhesion patches can shrink and grow. Furthermore, adhesions can break for a set energy cost. We first studied the effect of two parameters, namely the energy cost of breaking adhesions λ_{adh}, also interpretable as the strength of an adhesion bond, and the probability with which new adhesions form at the polarized front of the cell. Cells with low adhesion area perform a high speed, highly persistent random walk. The simulated cells slow down for increasing bond strength, but also for increasing *de novo* adhesion formation. For very high bond strengths, the cells can even get stuck. Stuck cells show a different type of motility which is persistent on a short time scale, but subdiffusive on long-time scales. By adding feedback between cell-matrix adhesion area and propulsion strength, a richer behavioural repertoire can be reproduced for low adhesive areas. With this feedback, simulated cells with very low adhesive areas have low dispersion, as their propulsion strength is weakened. Cells with slightly higher adhesive areas can show temporary spurts of increase in adhesion area combined with increase in velocity. Finally, we studied the effect of the processes that form the adhesion area and found that adhesion cluster size distribution can affect cell motility. Cells with many small adhesion clusters at the cell front perform a more persistent motion than cells that have fewer but larger adhesion clusters located at the centre and the back of the cell, even while total adhesion area is equal.

For long-term behaviour in our model, we mostly observed diffusive behaviour or, for more extreme λ_{Act} and p_s values, subdiffusive behaviour. The B cells observed by Rey-Barroso et al. performed long-term diffusive behaviour. Superdiffusive behaviour has also been observed in mammalian

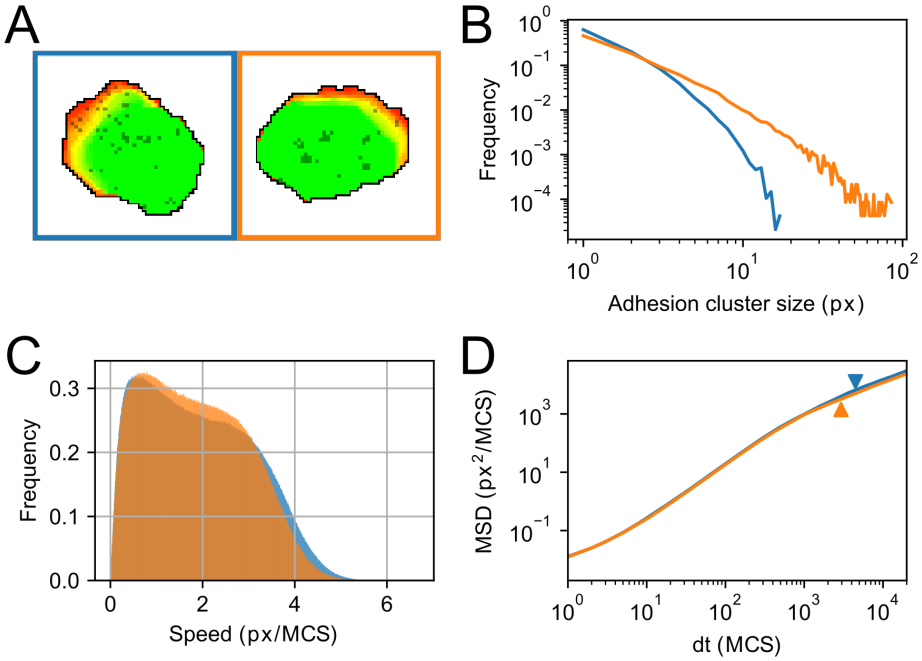


FIGURE 3.8: **Adhesion growth dynamics influence adhesion cluster size, cell speed and MSD** (A) Example of different adhesion cluster size for different parameter values of p_s and p_e. Blue: $p_s = 0.003$, $p_e = 0.0015$ resulting in a multitude of single grid point sized adhesions. Orange: $p_s = 0.001$, $p_e = 0.004$, resulting in a small number of larger clusters. Colours in B,C,D correspond to these parameter settings. (B) Distribution of cluster sizes for 1000 independent simulations for each parameter setting on a logarithmic scale. Distribution of blue does not exceed cluster size 20 (C) Distribution of instantaneous velocity of 1000 independent simulations for each parameter setting. Mean speed for orange is lower compared to blue (D) Log-log plot of MSD. The onset of the second linear regime (log-log slope approximately equal to 1) is marked with an arrowhead in corresponding colours. This regime starts at smaller dt for the orange curve compared to the blue curve, which corresponds to a lowered persistence time compared to blue. Simulations are also available as Supplementary Video S3.

cells. The murine T cells in [56] showed superdiffusive behaviour, but have only been tracked for a relatively short time (∼10 min), so their diffusive behaviour on longer time scales is unknown. The Madin-Darby canine kidney cell in [160] have been tracked for a much longer time (∼1000 min) and perform superdiffusive walks at both short and large time scales, and their velocities show long range correlations in time. What causes these long-time correlations is unclear. As we have not seen our model display long-term superdiffusive behaviour, and because there is no long-term memory in our current model, we think that some molecular memory could play a role in cell superdiffusivity.

Intercellular variability

When comparing the resulting behaviours of this model with the motility patterns described in [143] and [144], we were able to capture the floating, stepping, and pivoting behaviour observed in B cells on fibronectin, by just altering adhesion bond energy cost and *de novo* adhesion formation, as well as the walking and gliding behaviour observed in T cells on ICAM and casein, by adjusting *de novo* adhesion formation and adhesion patch growth. Noteable is that these different types of motility were all observed within the same populations of cells.

Variability among individual isogenic cells has also been described in chemotaxis of *Dictyostelium discoideum* cells [164] and keratocyte shape and motility [39]. Moreover, a single random walk model could not describe the motion of a CD8+ T cell population, but division of the population into Brownian walkers and persistent walkers could describe the motion of the population [165].

Our model can aid in pinpointing what underlies the intercellular diversity. It shows that a small degree of variability in adhesion formation rate, adhesion detachment rate, adhesion distribution, or in the coupling from adhesions to cytoskeleton can lead to distinct migration properties and even modalities. It is therefore tempting to speculate that lymphocyte populations observed *in situ* or in cultures harbour a certain degree of heterogeneity in some of these pivotal parameters. Already, different subsets of differentiated CD4+ T cells (Th1/Th2/Th17 subsets) have been described to harbour distinct motility properties both *in vitro* and *in vivo* [166, 167]. These differences appear to be explained by distinct molecular equipment in terms of adhesion and cytoskeleton dynamics. Interestingly, these differences have been proposed to support distinct search strategies aligning with the fact that these cell subsets target different types of pathogens. Actually, our study provides a mechanistic framework to ask such questions

and address them experimentally (e.g., by measuring integrin expression levels among individual cells by flow cytometry, monitoring size and distribution of adhesions with super-resolution microscopy approaches). For instance, to fully understand the mechanisms underlying the differences between B cell motion on fibronectin versus collagen, more knowledge on this intercellular diversity in integrin dynamics would be beneficial. This knowledge can then be used in further simulations to choose parameters for individual cells from a suitable range and reproduce the fast Brownian motion on fibronectin with more adhesive area as well as the slow persistent motion on collagen IV with lower adhesive area.

How to disentangle velocity and persistence in the model

Although our model provides a plausible explanation for the impact of cell adhesion on subdiffusive cell motility, it does not reproduce the observation that B cells move faster and in a Brownian fashion on fibronectin, but slowly and persistently on collagen IV. This could be due to the fact that in our model, persistence and motility (dispersion) are highly correlated: fast cells also show high persistence, so it is unlikely that simulations will result in slow but persistent motion. An experimental study showed a universal coupling between cell speed and cell persistence (UCSP) to be mediated by actin flow [168], as actin flow stabilizes cell polarity. In our model, the actin flow is modelled phenomenologically by the Act model [73], which displays this UCSP as well [147].

Currently, the Act-extension, and specifically λ_{Act} is the only model component that is influenced by the adhesion dynamics. If adhesions stabilize actin fibres, it is also reasonable to make Max_{Act} or the speed of Act degradation dependent on the presence of adhesions. Otherwise, other aspects of cell locomotion, already captured in our model or not, could be influenced by cell-matrix interactions and lead to slow but persistence motion on collagen IV.

Where the Act-extension mainly models the front of the cell, many locomotion-related processes also involve the rear of the cell. Myosin-II contraction pulls the back of the cell towards the front and can increase cell velocity [144]. Preliminary studies with our model show that cell velocity can be changed by altering the weight of the perimeter constraint, or by changing the contact energy between cell and matrix. Both of these components model myosin-II contractility indirectly. Part of this cortical tension is transferred onto the matrix through adhesions [169, 170]. An interesting question is whether the cortical tension is also influenced by the presence of adhesions. Furthermore, myosin-II is suggested to be a

polarization cue and to be transferred to the back of the cell by retrograde actin flow and could possible also alter persistence of cell polarization [168]. An interesting direction for future research would be to study how the retrograde flow is influenced by cell-matrix adhesions and how this may affect the UCSP.

Matrix rigidity and mechanistic feedback between integrins and matrix

We have modelled cell-matrix adhesions as a simple on/off-switch, with a set amount of energy required to break the adhesion. In reality, the adhesion process is much more complex, involving mechanistic feedback between integrins and the matrix. Hence, both matrix rigidity and the cell's ability to generate force influence cell shape and cell motion. When it comes to modelling this feedback, different approaches have been used already. In Copos et al. [145], adhesions were modelled as mechanosensitive bonds. In Ziebert et al. [65], adhesions ruptured when they exceeded a maximum length. In Shao et al. [66] the probability of adhesion rupture increased with force. In Löber et al. [67], the matrix deformation was also taken into account, leading to non-trivial motion such as bipedal motion.

Modelling matrix deformation or displacement of adhesion sites within the CPM is challenging, but a lot of progress has been made recently. Methods to estimate forces within the CPM cell have been developed, either based on cell shape or on the Hamiltonian [171, 62]. The CPM has also been combined with a finite element method to model matrix traction forces with feedback between the CPM and FEM [171, 172]. Moreover, cell-matrix adhesions were recently introduced into this framework [173] as focal adhesions with force dependent growth, and smaller focal adhesions being easier to dislodge from the ECM. In the observed B cells, however, adhesions are not structured into focal adhesions. Nonetheless, these novel methods can be used to improve the realism of our model, both on the side of cell-matrix bonds, as well as the side of matrix deformation and matrix stiffness.

Adhesion patch detachment

In our current model, adhesion patch detachment occurs through a stochastic process of loss of sites combined with the energy requiring retraction. This part of the model can be refined in several ways. First, it is known that myosin-IIA, besides rear-end retraction, is also involved in the detachment of adhesion patches in T cells [144, 174]. Myosin-II increases the forces

exerted on the adhesions. Combining mechanistic feedback between adhesions and the matrix, as proposed in the previous paragraph, with an explicit model of myosin-II near adhesion patches could result in more realistic patch dynamics. Second, detachment at the rear of the cell is also regulated by other molecular processes. Talin and moesin, both scaffolding proteins between integrin and the cytoskeleton, can compete with each other, but have different properties. While talin connects the cytoskeleton to integrins, moesin inactivates integrin, thereby decoupling the adhesion from the cytoskeleton [175]. This process mainly occurs at the rear of the cell. Integrating the activities of talin and moesin into our model can refine the de-adhesion process at the cell rear currently modelled by an energy threshold.

Conclusion

In conclusion, we have introduced a CPM model combining the Act model [73] with dynamic cell-matrix adhesions. We have shown that our model is capable of extending the repertoire of motility types within the CPM, both from a detailed cellular level, as well as on a statistical level (Figure 3.9). Besides the persistent random walk that emerges from the Act model, our model is also able to reproduce a short-term persistent but long-term subdiffusive random walk. While adhesion dynamics are simplified, our model can show the different types of motion observed in individual B cells on a fibronectin substrate, such as reduced motility for non-attached cells, a walking motion, and pivoting due to sustained attachment, as well as the walking and gliding motion of T cells on ICAM or casein substrate. Here we uniquely link short-term molecular scale to the long-term cell behaviour scale to learn about those molecular parameters that explain the plasticity of immune cell motility upon interaction with varying substrates. In particular, our study highlights that the interplay between adhesion formation, adhesion expansion and adhesion strength determine the turn-over of the adhesion area which regulates cell speed and persistence. Furthermore, the model provides a mechanistic framework for generating experimentally testable hypotheses.

3.4 Methods

In this work, we model the different cell motilities of cells adhering on flat matrix surfaces. The basis of our model is the Cellular Potts model.

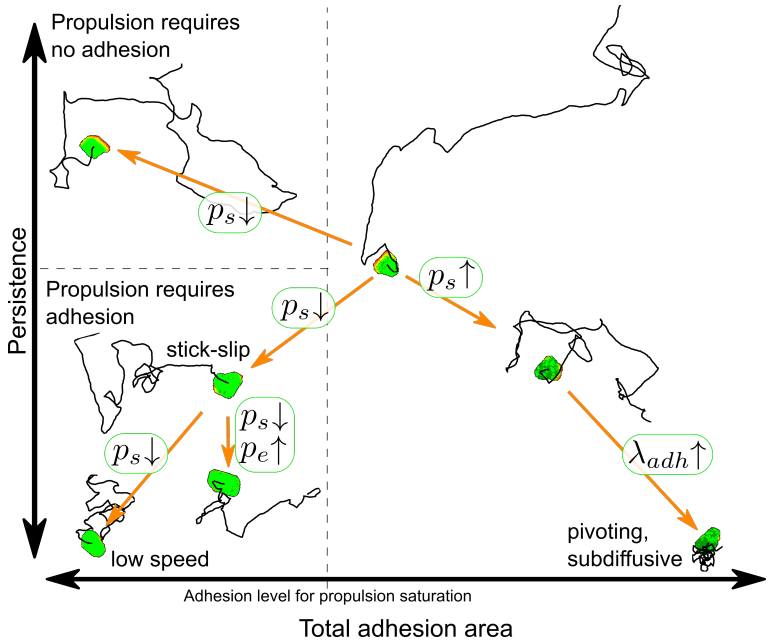


FIGURE 3.9: **Overview of the motility modes possible in the model and which parameters govern the transitions between them.** For each motility mode, a representative cell and its trajectory are plotted in a persistence versus total adhesion area plane.

3.4.1 Cellular Potts model

The Cellular Potts model (CPM) is a grid-based model. Grid points are assigned an identity σ, indicating to which cell they belong. All grid points with the same σ together form a single cell. The model evolves through time by doing copy attempts: two neighbouring grid points are selected and the σ of one grid point is trying to be copied into the other grid point. Whether this copy attempt actually occurs is decided by the Hamiltonian $\mathcal{H}$, which describes the energy of the system. Different Hamiltonians have been proposed for different CPM models. We use the following:

$$\mathcal{H} = \sum_{u,v \in \text{Neighb.}} J_{\tau(\sigma_u),\tau(\sigma_v)}(1 - \delta_{\sigma_u,\sigma_v}) + \lambda_A \sum_{\sigma}(a_\sigma - A_\sigma)^2 + \lambda_P \sum_{\sigma}(p_\sigma - P_\sigma)^2$$

The first term describes the contact energies between cell and medium or between two cells. The second term describes an area constraint, with A_σ being the target area of cell σ and the third term describes in similar ways a perimeter constraint with P_σ the target perimeter of cell σ.

A copy attempt is accepted depending on the change in Hamiltonian $\Delta\mathcal{H}$ it causes:

$$P(\Delta\mathcal{H}_{\mathbf{x}\to\mathbf{y}}) = \begin{cases} 1 & \text{if } \Delta\mathcal{H} < 0 \\ e^{-(\Delta\mathcal{H}/T)} & \text{if } \Delta\mathcal{H} \geq 0 \end{cases}$$

Here, T plays the role of an effective temperature and controls the amount of variability we allow to happen. Higher T will allow more thermodynamically unfavourable copy attempts to succeed.

A measure of time in the CPM is the Monte Carlo Step (MCS). Within one MCS, the expectation is that each grid point has been updated once. We use a rejection-free algorithm that only considers attempts between neighbours of different σ to speed up simulations [69, 176].

3.4.2 Cell motility

Cells move by making protrusions through actin polymerization and form cell extrusions like filopodia, pseudopodia and lamellipodia. Actin polymerization in the CPM has previously been modelled in a phenomenological way by Niculescu et al. [73]. We also use this extension to model the actin polymerization. This extension adds an extra layer to the CPM, described as the Act-layer. The values of this layer vary between 0 and a maximum value Max_{Act}. Grid points that are newly added to a cell obtain Act-value Max_{Act} and each MCS the values of the Act layer are lowered

by 1 until 0. The Act-extension is added to the CPM as an extra term for the change in Hamiltonian $\Delta\mathcal{H}$, namely as $\Delta\mathcal{H}_{Act}$. For $\Delta\mathcal{H}_{Act}$, the local geometric mean of Act-values of both the expanding and retracting grid points are calculated and compared. Then, $\Delta\mathcal{H}_{Act}$ favours the grid point with the highest mean in the following way:

$$\Delta\mathcal{H}_{Act}(u \to v) = \frac{\lambda_{Act}}{\mathrm{Max}_{Act}} \left(\left(\prod_{y \in V(u)} Act(y) \right)^{1/|V(u)|} - \left(\prod_{y \in V(v)} Act(y) \right)^{1/|V(v)|} \right)$$

with $V(x)$ describing the neighbourhood of grid point x in the same cell. Here λ_{Act} is the weight given to this model component.

Adhesion to the matrix makes actin polymerization more efficient in protruding the cell membrane [161, 162, 163], by transmitting the force to the matrix. We add feedback between the cell adhesions and the actin polymerization, by strengthening the force produced by polymerization upon increase in adhesion area. This is done by multiplying λ_{Act} with factor f defined as follows:

$$f = \begin{cases} b + \frac{1-b}{s} \frac{A_{adh}(i)}{A(i)} & \text{if } \frac{A_{adh}(i)}{A(i)} \leq s \\ 1 & \text{if } \frac{A_{adh}(i)}{A(i)} > s, \end{cases},$$

Here $A(i)$ denotes the area of cell i, and $A_{adh}(i)$ denotes the adhesive area of the cell i, b the value of f when there are no adhesions, and s the value at which f saturates.

3.4.3 Adhesions to the substrate

The adhesions of a cell to the extracellular matrix are modelled as a third layer in the CPM. A grid point x in this layer can have either $Adh(x) = 0$, no adhesion, or $Adh(x) = 1$, when it denotes an adhesion patch of the cell. Adhesion dynamics are governed by four processes: spontaneous formation of new adhesions, adhesion patch expansion, adhesion patch retraction, and rupture of adhesion through retraction of a cell. We describe each of these processes below.

New adhesion sites

New adhesions form when the cell membrane comes in close enough contact with the extracellular matrix such that integrins can bind to the matrix. This process is dependent on actin polymerization, membrane

protrusion and polarized distribution of integrins [151, 152, 153]. We model *de novo* adhesion formation through a stochastic process. Each MCS a grid site within a cell can turn from non-adhesion to an adhesion site. This can only happen when the local geometric mean of the Act layer exceeds the value $0.75\,\mathrm{Max}_{Act}$, i.e. when:

$$\left(\prod_{y \in NB(x)} Act(y)\right)^{\frac{1}{|NB(x)|}} \geq 0.75\,\mathrm{Max}_{Act}$$

Then, the probability for a cell to form a new adhesion at that grid site is given by the probability parameter p_s.

Adhesion patch expansion

Once adhesion patches are formed, they can increase in size. Multiple processes underlie this expansion. First, once the cell membrane is attached to the matrix, it fluctuates less, allowing for easier attachment of new integrins [154]. Secondly, the curvature of the cell membrane influences the intermolecular forces, favouring aggregation of integrins [155, 156].

We do not model integrin recruitment and membrane curvature, but choose to model adhesion patch growth phenomenologically. Jacobelli et al. [144] observed that adhesion patches grow radially, with some bias in the direction of the cell front. Hence, we model adhesion patch expansion as an Eden-like growth model [157], known to give roughly circular shapes. While updating the adhesion layer, once a grid point containing an adhesion is selected to be updated, we also select a random neighbour. If that neighbouring grid point contains no adhesion, it can form an adhesion with probability p_e.

Adhesion patch retraction

Aside from patch expansion, patch retraction can also occur. Patches will partially dissolve spontaneously, as they are out-of-equilibrium systems [177]. Also, Myosin-II contraction is involved in patch retraction [144], which occurs concentrically. Following these observations, an adhesion site x in this model can spontaneously decay with a probability depending on the adhesion status of its neighbours.

$$P(x \text{ will de-adhere}) = p_d \cdot \left(\frac{|\{nb \in NB(x) | Adh(nb) = 0\}|}{|\{nb \in NB(x)\}|}\right)^2$$

with $NB(x)$ the neighbourhood of x. Thus, the higher the number of non-adherent neighbours, the higher the probability that the site loses its adhesion.

Adhesion rupture through retraction

Adhesions can also disbond by force. It is know that integrin shows catch-slip bond behaviour [178, 179]. Since we do not directly model forces in the CPM, and since we do not model matrix stresses, we simplify the rupture of an adhesion to a constant amount of energy required to break an adhesion upon cell retraction. This energy is added to the change in Hamiltonian the following way:

$$\Delta\mathcal{H}_{adh}(u \to v) = \lambda_{adh} Adh(v)$$

Here $\sigma(u) \neq \sigma(v)$ and the cell $\sigma(v)$ is retracting.

3.4.4 Order of layer updates

This model has three layers: one for the grid site identities σ, one for the Act-values and one for the adhesions. These three layers are updated in the order σ - adhesions - Act-values. The σ-layer is updated through the rejection free Metropolis algorithm. During the σ-update, Act-values and adhesion updates regarding the relocation of the cell are executed immediately: e.g., for copy attempts that let a cell retract from a grid point, we do directly update the Act-values and adhesions of that site. After the σ-update, we update the adhesion layer asynchronously: we iterate, in random order, over the grid points within the cell and execute the processes described in the Adhesions subsection. Lastly, we update the Act-layer: every Act-value is diminished by 1 until 0. These three updates together constitute one MCS and this cycle is repeated for a set number of MCS.

3.4.5 Simulation parameters

During our different simulations, many parameter values were kept constant (Table 3.3). All simulations were done on a 300×300 pixel grid with periodic boundaries with a single cell. Parameter values that were not constant are shown in Table 3.1. For the simulations in Fig 3.2, 3.3 and 3.4, $p_d = 0.0064$, and p_s and λ_{adh} varied according to the figure legends. For simulations shown in Fig 3.6 and 3.7, $p_d = 0.008$ and again λ_{adh} varied according to the figure legends. The Act-only simulations in Fig 3.3 and Fig 3.7 were run with all adhesion dynamics parameters equal to zero: i.e.,

TABLE 3.3: List of parameter values kept constant during all simulations. Values are arbitrary units, unless specified otherwise.

Parameter	Description	Value
T	temperature	30
A	target area	1000 px^2
λ_A	weight area constraint	50
P	target perimeter	350 px
λ_P	weight perimeter constraint	4
λ_{Act}	weight of Act model	240
Max_{Act}	Act lifetime	120 MCS
$J_{medium,medium}$	adhesion energy between medium	0
$J_{cell,medium}$	adhesion energy between cell and medium	35
Total MCS	simulation duration	25000 MCS

λ_{adh}, p_s, p_e, and p_d were all zero. For all simulations, $\lambda_{Act} = 240$, except for the specific Act-only simulations in Fig 3.7 with $\lambda_{Act} = 120$. For the simulations in Fig 3.8, p_e and p_s were varied, see figure legend. The parameters not mentioned in the figure legend are $p_d = 0.0004$, $\lambda_{adh} = 60$, $b = 0.5$, $s = 0.12$.

Acknowledgments

G.T. gratefully acknowledges the Indian Institute of Science to serve as Infosys visiting professor at the Centre for Ecological Sciences in Bengaluru. L.v.S. would like to thank the French Embassy in the Netherlands for the Bourse d'Excellence Descartes de stage. We also thank SURFsara for the support and computing time in using the Lisa cluster computer.

3.5 Supporting information

The supplementary videos can be accessed online at `https://doi.org/10.1101/2021.06.09.447692`.

S1 Video Cell velocity and persistence drop with increasing values for adhesion formation and adhesion strength.

S2 Video Cell velocity and persistence drop in the model with feedback from adhesions onto cell protrusion when adhesive areas are small.

S3 Video Similar adhesive area sizes lead to different motility when adhesion growth is dominated by the actin-dependent formation of new adhesions versus the growth of existing patches.

Chapter 4

Topotaxis on silicon and *in silico*: Obstacle-induced contact-inihibition of locomotion explains topotactic cell navigation in dense microenvironments

Leonie van Steijn, Joeri A.J. Wondergem, Koen Schakenraad, Doris Heinrich, Roeland M.H. Merks

Abstract

During biological development, cancer metastasis and in the immune system, cells navigate through dense environments filled with obstacles such as other cells and the extracellular matrix. Recently, the term 'topotaxis' has been introduced for the navigation of cells along topographic cues such as density gradients of obstacles. Experimental and mathematical efforts have analysed topotaxis by looking at the migration of single cells in pillared grids, in which topographic gradients can be defined precisely as a pillar density gradient. We have previously introduced a model based on active Brownian particles (ABP) which has shown that ABPs perform topotaxis, i.e., drift, on average, to lower pillar densities, due to the decrease in effective persistence time when pillars are closer together. Whereas topotactic drifts of up to 5% of the instantaneous cell speed have been observed experimentally, the ABP model could only predict topotactic drifts of up to 1%. We hypothesized that the discrepancy between the ABP and the experimental observations is in 1) the deformable cell volume, and 2) the cell-pillar interactions. Here we introduce a more detailed model of topotaxis, based on the Cellular Potts model. To model persistent cells we use two methods: the Act model by Niculescu et al., and a method similar to the ABP model. After fitting our model parameters to yield the same motion as experimentally found for *D. discoideum* on an empty field, we study how topotaxis is performed. For starved *D. discoideum*, our model predictions are close to the experimental results, especially when we correct for the fact that our simulated cells have a higher speed. However,

topotactic efficiency is different between the two persistence methods, with the Act model outperforming the ABP persistent model. We found that the Act model shows a larger reduction in effective persistence time in pillar grids than the vector model. Lastly, we also modeled the slow and less persistent vegetative *D. discoideum* cells, for which our model predicted a small topotaxis drift. Here, both model variants predicted similar topotaxis, contrasting with the starved cell results.

Popular summary

Cell motility is an important function in development and immunity. Knowing how the environment influence cell motility is useful in developing methods to interfere during disease or in tissue engineering. One of the ways the environment affects cell motility is by the presence of obstacles. Previous work has shown that single cells move from a high density of obstacles to a lower density of obstacles, a process called topotaxis. A previous mathematical model, modelling cells as particles showed that the velocity at which cells move from high to low obstacle density, the topotactic drift, can in part be explained by reduced persistence: i.e. the closer the obstacles are together, the quicker the particle looses its direction, resulting in a bias towards less obstacle dense areas. However, the experimentally found topotactic drift was higher.

In this work, we look at cell motility in obstacle gradients with a more realistic cell model with deformable cells and compare those to an experimental setup. We used two different methods to model the persistence of cells: one which models the cytoskeleton and with active reorientation after collisions, the other similar to the previous mathematical model and with passive reorientation after collisions. We fitted our model to experimental cell motion. Both models as well as the experiments show more topotactic drift in steeper obstacle gradients. The model with active reorientation shows a better match in topotaxis efficiency and outperforms the model with passive reorientation. This is due to a larger loss of persistence when collisions feed back to cell motility.

Next steps would be to investigate more complex environments. Interesting directions would be the integration of multiple environmental cues, such as chemotaxis, in the model, and exploration of how obstacle shape influences topotaxis, so that we get a better understanding of cell navigation in the body.

4.1 Introduction

Motile cells within the body encounter many obstacles such as other cells and extracellular matrix, as they move through a tissue. How cells react to the density of obstacles is of importance for many processes, such as cancer cells invading from a dense tumor into looser packed tissues [180, 181], immune cells moving through tissues with different porosity [182, 183] or pathogens such as *Plasmodium* that migrate through different tissue throughout their life cycle [184, 185]. It has been shown that cell can use the topography of the environment as a way to orient themselves [186, 187, 188]. So far, topotaxis has only been shown in *in vitro* environments, but it will also likely play a role in *in vivo* systems.

To study topotaxis quantitatively, a collection of *in vitro* models have been developed to provide well-defined topographic cues to migrating cells. In [186], cells are allowed to move on a subcellular-patterned array coated with fibronectin. Cells follow the gradient in nanopattern, from dense to sparser patterning. In [187], cells were put on polyurethane acrylate nanohairs. These hairs were either vertical, or bent in a specific direction. Cells on top of these hairs moved with a bias towards the bent direction, but without bias on straight hairs. Also larger scale cues are known: in [188], cells were placed on micropatterned substrates. The patterns were either disconnected cell-sized triangles in a line, connected cell-sized triangles in a line and a belt. They showed that cells in disconnected triangles hardly moved from one spot to another, whereas cells in the other two patterns did. Cells in the belt moved without preferential direction over the belt, but the cells on the connected triangles moved with a bias in the direction of the triangles. The rearrangement of actin and formation of the lamellipodium were key players in this bias or ratcheting. An extensive review of other cues in 'ratchettaxis' has been done in [189].

Wondergem et al. showed another setup in which topotaxis arises from a cell-size cue [74]. They let starved *Dictyostelium discoideum* cells move on a grid with micropillars placed with increasing pillar spacing in one direction. As in previous work with isotropic pillar lattices [190][191], the size of these micropillars were in the order of *Dictyostelium* cell size, in contrast to [186] and many of the structures described in [189]. They found that, on average, cells move to the side with larger pillar distance, with a velocity of about 5% of their instantaneous speed.

Because *Dictyostelium* cells are persistent walkers [192], we hypothesized that pillars alter this persistent motion. Once a cell bumps into a pillar, it cannot continue and will have to change its direction, and hence change its persistence. This first hypothesis on this form of topotaxis was tested

by Schakenraad et al. [75]. They tested whether this topotaxis could be explained by altered persistence lengths, similar to durotaxis on a stiffness gradient [193] where cells on softer substrates change direction more often than on stiffer substrates, or bacterial chemotaxis [194] where cells perform more tumbles in lower concentration of the chemoattractant. Using a model with Active Brownian Particles, which also perform persistent motion, Schakenraad et al. showed that indeed topotaxis can be derived from altered persistence lengths due to pillar spacing. However, the ABP model could explain only a topotaxis of 1% of the intrinsic cell speed, significantly lower than the 5% found in *Dictyostelium* experiments. Here we propose a refined model for topotaxis giving a better match with the experiments.

This previous work shows that topotaxis in gradient pillar density can partially be explained by changed effective persistence lengths. However, what are the remaining factors that can explain the gap in topotactic drift between ABPs and *Dictyostelium* cells? Some candidates are easy to point out: ABPs are point-particles, whereas cells have a deformable and non-zero volume. We can therefore expect cells to have a minimum pillar distance through which they can move, as well as larger than zero area of interaction between the cells and pillars. This cell-pillar interaction points to another candidate: how cells reorient after collision. Where ABPs only change their target direction by thermal fluctuation, cells can actively reorient themselves. A well studied example of cell reorientation is Contact Inhibition of Locomotion (CIL), where cells reorient themselves after collision with another cell.

In this work, we model a persistently moving, deformable cell with volume using the Cellular Potts model and test two different schemes for cell persistence. The first scheme used the Act model which models internal cell skeleton rearrangements [73]. This allows for emergent, active reorientation as the cell cannot extend further into a pillar and will polarize elsewhere, similar to CIL. As a control, the second scheme resembles the ABP model and shows no CIL, as we explicitly assign a direction to a cell together with a persistence time, which results in passive reorientation upon collision. Before we look into topotaxis, we search for parameter settings that match experimental *Dictyostelium* cells on an empty field so we can quantitatively compare our model results with the experimental results of persistently moving starved and less persistently moving vegetative *Dictyostelium* cells.

For the starved parameter settings, the two different persistence schemes results in different topotaxis, with the active reorientation model having more efficient topotaxis. Our models slightly overpredict topotaxis compared to the experimental results, but the main difference is cell speed,

which is much higher in the simulations than in the experiments. On a relative level, the model topotaxis and experimental topotaxis match very closely, with the active reorientation model being a closer match. Trying to explain the difference in speed, we test the effect of interaction between pillar and cell, by changing the adhesion energy between them. However, when simulated cell speed approximates experimental cell speeds, the cells are largely sticking to the pillars, which we do not observe in the experiments. Finally, we also predict a minute effect of the pillar grid on vegetative *Dictyostelium*, that is only visible for very steep pillar gradients, and is independent of persistence scheme.

4.2 Results

We characterized topotaxis for *in vitro* starved *D. discoideum* cells, and developed a Cellular Potts model with persistent cell motion with active or passive reorientation upon collisions. Before we compared topotaxis measured in experiments and topotaxis measured in simulations, we made sure to match the motility of simulated cells to *in vitro* starved cells. Futhermore, we also matched the motility of simulated cells to *in vitro* vegetative *D. discoideum* cells. After that, we characterized topotaxis in our starved as well as vegetative simulations.

4.2.1 Model

We used the Cellular Potts model to simulate *Dictyostelium discoideum* cells. The Cellular Potts model is a 2D lattice based model in which a cell is represented by a number of lattice sites. The lattice is updated per site, allowing for deformable cells, and a set number of lattice updates is called a Monte Carlo Step (MCS), the time unit we use in this model. We can set a goal cell area and cell perimeter, so we can control cell size and shape to a desired extent.

In order to obtain persistent cell motion, we used two extensions of the CPM. The first, the Act-model, is based on an actin-inspired feedback mechanism that results in cell polarization [73]. It is controlled by two parameters: λ_{Act}, which determines the weight of the Act-model and can also be interpreted as the maximum protrusive force of the actin network, and $\max_{\mathrm{Act}}$, the maximum activity value, also interpretable as the lifetime of an actin subunit within the actin network. By altering these two parameters, a large variety of persistent random walkers can be achieved in this model. Furthermore, due to the feedback mechanism, the cell can loose

its polarization upon collision and repolarize in a different direction, i.e. it has contact inhibition of locomotion. We indicate this model as the active reorientation model.

The second extension is based on the previous ABP model and has a directional vector at its center. The cell is more likely to move in a certain direction if that movement aligns with the directional vector. The directional vector changes over time by a scaled random noise term. This extension is also controlled by two parameters: $\lambda_{\text{persistence}}$, which determines the weight of this extension and how strictly cells should align to their directional vector, and τ, the persistence time of the directional vector. Combinations of these two parameters result in a large variety of persistent random walkers as well. Notably, this model extension has no CIL as the cell will only reorient once its directional vector has changed. We will further indicate this model version as the passive reorientation model.

4.2.2 Matching model parameters to cell motion

In order to match *in silico* cell motility to *in vitro* cell motily, we determined a number of *in vitro* cell properties for starved and vegetative *Dictyostelium discoideum* on a flat PDMS substrate (Figs. 4.1, 4.2). We measured the mean cell area (Figs 4.1D,4.2B), mean cell speed (Figs 4.1E,4.2C), cell persistence time (Figs 4.1G,4.2E) and cell diffusion constant (Figs 4.1H,4.2F), see also Table 4.1. In general, starved *D. discoideum* cells are highly motile and perform a persistent walk, whereas vegetative *D. discoideum* cells are less motile and less persistent.

From these properties, cell area and cell speed were used to set the length and time scale in the CPM. With the free parameters λ_{Act} and $\max_{\text{Act}}$ for the active reorientation model, and $\lambda_{\text{persistence}}$ and τ for the passive reorientation model, we used a hillclimbing algorithm to fit the diffusion constant and persistence time of simulated cells to the *in vitro* values in Table 4.1. The resulting best fits for both starved and vegetative cell motilities and both persistence methods are also shown in Table 4.1, and the corresponding model parameter values are shown in Table 4.2. Overall, we have achieved a good match for the starved cells and a reasonable match for the vegetative cells.

4.2.3 Introducing pillar gradients

With the model performing similar motion to the starved and vegetative cells, we introduced a pillar grid with gradient [75]. The pillar gradient is

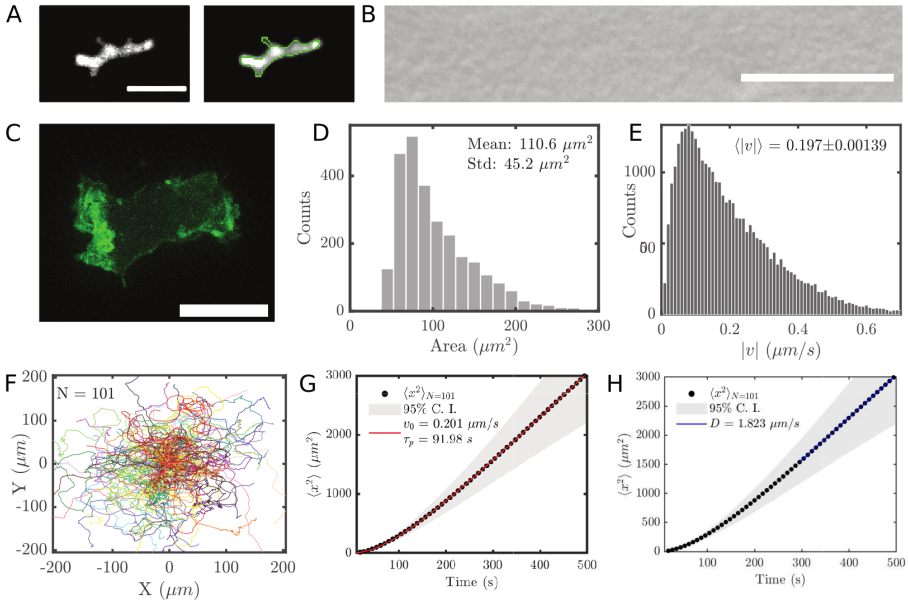


FIGURE 4.1: **Empirical basis of CPM parameter constraints I: starved *D. discoideums***

A) Detection of starved *D. discoideum* motion on flat PDMS was used to calibrate the CPM simulations. Time-lapse fluorescent confocal imaging of cells (left) was used to detect cell edges (green, right). Scale bar is 10 μm. **B)** SEM image of flat PDMS, without engineered topographies the substrate has nanometer scaled features providing traction to cells. Scale bar is 1 μm **C)** Actin polymerization hot spots (visualized through LimE-GFP, z-projection) are highly anisotropically distributed, associated with the high persistence of starved *D. discoideum*. **D)** Areas measured for starved cells on flat PDMS. These determined the area (A_σ) parameter used in the CPM. **E)** Instantaneous velocity distribution of observed cell motility. **F)** Trajectories obtained from live cell imaging. **G)** The MSD (black) of starved *D. discoideum* trajectories is well fit by analytical expression for persistent random particles (red). The average instantaneous velocity ($\langle |v| \rangle$) and persistence time (τ_p) were used to calibrate the vector- and actin based CPMs. **H)** The MSD (black) of starved *D. discoideum* trajectories at larger lag-times is well fit by a line (blue). The slope of the blue line is used to derived the diffusion constant D, which was used to calibrate the vector- and actin based CPMs.

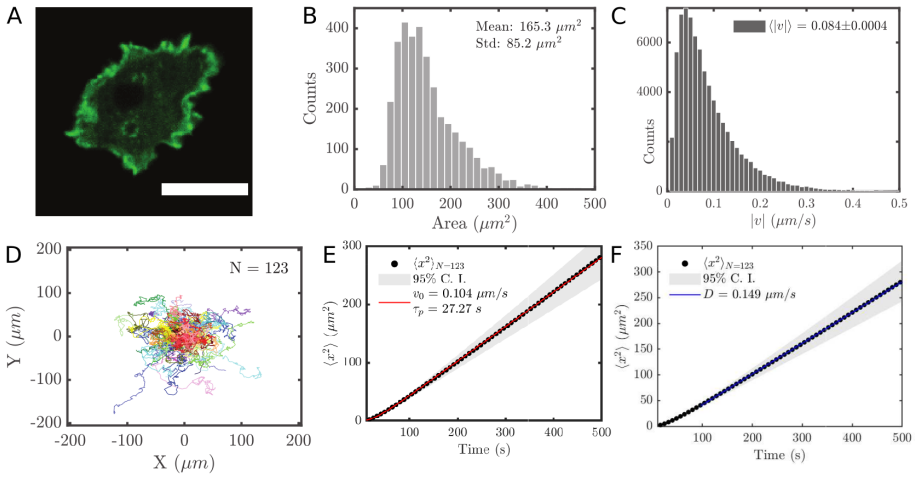


FIGURE 4.2: **Empirical basis of CPM parameter constraints II: vegetative *D. discoideums***

A) Actin polymerization hot spots (visualized through LimE-GFP, z-projection) in the cell are more isotropically distributed in the vegetative state, giving rise to slower and less persistent motion than starved *D. discoideum*. **B)** Areas measured for vegetative cells on flat PDMS. These determined the area (A_σ) parameter used in the CPM. **C)** Instantaneous velocity distribution of observed cell motility. **D)** Trajectories obtained from live cell imaging, motion detection was performed equal to the starved state (see Fig 4.1). **E)** The MSD (black) of starved *D. discoideum* trajectories is well fit by analytical expression for persistent random particles (red, but the parabolic regime for low lag-times is significantly shorter. The average instantaneous velocity ($\langle |v| \rangle$) and persistence time (τ_p) were used to calibrate the vector- and actin based CPMs. **F)** The MSD (black) of starved *D. discoideum* trajectories at larger lag-times is well fit by a line (blue). The slope of the blue line is used to derived the diffusion constant D, which was used to calibrate the vector- and actin based CPMs.

TABLE 4.1: Cell motility properties for the experimental data and best fits for the active and passive reorientation models.

		Starved	
	Experimental	Active	Passive
Length equivalent 1px (μm)	-	0.525	0.525
Time equivalent 1 MCS (s)	-	0.373	0.574
Speed ($\mu m/s$)	0.197±0.001	-	-
Area (μm^2)	110.4±45.1	-	-
Effective diffusion ($\mu m^2/s$)	1.82±0.68	1.76±0.08	1.87±0.08
Persistence time (s)	91.98±0.98	89.30±2.22	89.89±3.33
		Vegetative	
	Experimental	Active	Passive
Length equivalent 1px (μm)	-	0.524	0.524
Time equivalent 1 MCS (s)	-	0.388	0.821
Speed ($\mu m/s$)	0.084±0.0004	-	-
Area (μm^2)	164.6±84.3	-	-
Effective diffusion ($\mu m^2/s$)	$0.149 \pm 4.89 \times 10^{-4}$	0.137±0.007	0.121±0.006
Persistence time (s)	27.41±0.41	25.48±0.35	33.27±1.48

TABLE 4.2: Parameters of the best fits for the starved and vegetative *Dictyostelium* cells for both the active and passive reorientation models.

Best fit parameters	Starved	Vegetative
T	20	20
Area	400	600
λ_{area}	50	50
Perimeter	313	350
$\lambda_{perimeter}$	1	1
$J_{cell,medium}$	20	20
Active model		
λ_{Act}	129	80
$\max_{Act}$	37	33
Passive model		
$\lambda_{persistence}$	159	78
τ	30	7

TABLE 4.3: Pillar gradient grid parameter values for the experimental and simulation setup

Parameter	Description	Experiment (µm)	Model (px)	Model (µm)
R	pillar radius	5	10	5.3
h	pillar height	18 ±2	N/A	
d	pillar center distance at origin	16.8	32	16.8
d_{min}	pillar center distance left cap	13.6	26	13.7
d_{max}	pillar center distance right cap	19.9	38	20.0
s	gradient steepness	0.01, 0.03	0.01 - 0.11	

defined according to the set of pillar centers described by

$$P = \left\{ \vec{x} \in \mathbb{R}^2 \;\middle|\; \begin{array}{l} x_1 = \frac{d}{1-e^{-s}}\left(e^{sn} - 1\right) + \frac{d}{2} \text{ and} \\ x_2 = d\left(m + \frac{1}{2}\right)e^{sn}, \text{ with } n, m \in \mathbb{Z} \end{array} \right\}.$$

Here, d is the distance between pillar centers at the center of the grid, and s a parameter that controls the steepness of the gradient, with higher s indicating steeper gradients. P defines a pillar gradient in the x-direction with increasing pillar distance from left to right. The gradient is capped a both ends with a regular grid of pillar center distance d_{min} and $d_{max} = 2d - d_{min}$ to prevent cells from not being able to pass in between pillars. All measurements, both for the experimental and simulation setup and including pillar radius and pillar height, are shown in Table 4.3.

In the experimental setup, the pillar grid is a molded PDMS pillar grid. Cells are able to navigate between the pillars and resolve collision with pillars (Fig. 4.3A). Because cells were seeded randomly on the pillar grid, only $s = 0.01$ and $s = 0.03$ were used (Fig. 4.4A,B), as for steeper gradients too few cells are seeded in the short pillar gradients. Starved *D. discoideum* cells were tracked while in the gradient (Fig. 4.4E,F,G).

For the simulations, pillars were modeled as immobile obstacles. Cells could not move into a pillar, but were allowed to retract from them. Adhesion between cell and pillar can be controlled by the contact energy parameter $J_{\text{cell,pillar}}$. We assume that cells show no preference for pillars nor medium, so the contact energies between cell-pillar and cell-medium are equal ($J_{\text{cell,pillar}} = J_{\text{cell,medium}} = 20$), unless stated otherwise. In contrast to the experiments, repeated simulations were done with single cells only, starting from the same initial position in the center of the field, see Figs. 4.5,4.7,4.10.

The different ways in which the two persistence models respond to

obstacles is clarified in Fig. 4.3B,C. For the active reorientation model, we see CIL: once the cell collides with a pillar, it looses its polarization and will repolarize in a different direction from before. For the passive reorientation model, the cell will try to continue into the same direction, even if it is directed into a pillar, until the direction vector is no longer pointing into that pillar. The time it takes to resolve a collision is in the order of minutes, similar to that seen in the experimental setup (Fig. 4.3A).

4.2.4 Starved *D. discoideum* cells show topotactic drift increasing with gradient steepness

In order to measure the effect of topotaxis of starved cells, drift in both the x and y direction (v_x and v_y) was calculated for cell tracks in the gradient (Fig. 4.4F,G,H). More specifically, the migratory drift was calculated by averaging over all x or y displacements of all trajectories. As a control, the migratory drifts were also calculated for starved cells on flat surface. Starved *D. discoideum* cells showed a significant drift in the positive x-direction (i.e. towards lower pillar densities) for both pillar gradients $s = 0.01$ and $s = 0.03$. This holds compared to both the corresponding v_y as well as compared to v_x on flat surface (Fig. 4.4H). Moreover, the cells on the steeper gradient ($s = 0.03$) showed a larger v_x relative to the cells on the $s = 0.01$ gradient.

4.2.5 Reorientation mode of starved cells affects topotactic capacity

We simulated the same gradient pillar grid in our model. Examples of cell tracks and the pillar grids for different gradients are shown in Figure 4.5. Cells move within the pillar grid, but only explore a limited space each. We can see that the active reorientation model on the steep gradient (Fig. 4.5C) does not penetrate into the dense side of the grid as much as the passive reorientation model does (Fig. 4.5F). The population mean x and y coordinates over time (Fig. 4.5G,H) show that there is drift along the x-axis, but not the y-axis, indicating that the simulated cell perform topotaxis in both models.

Figure 4.6A shows the migratory drift in the x-direction (v_x) as a function of s, for the experimentally measured topotaxis (Fig. 4.4H) and both reorientation models for the starved parameter set. The experimental observation of increasing v_x with increasing s in starved cells is also seen in the results of our models (Fig. 4.6A). Both reorientation models extrapolate this trend of v_x increasing with s in a linear fashion. However, there is a

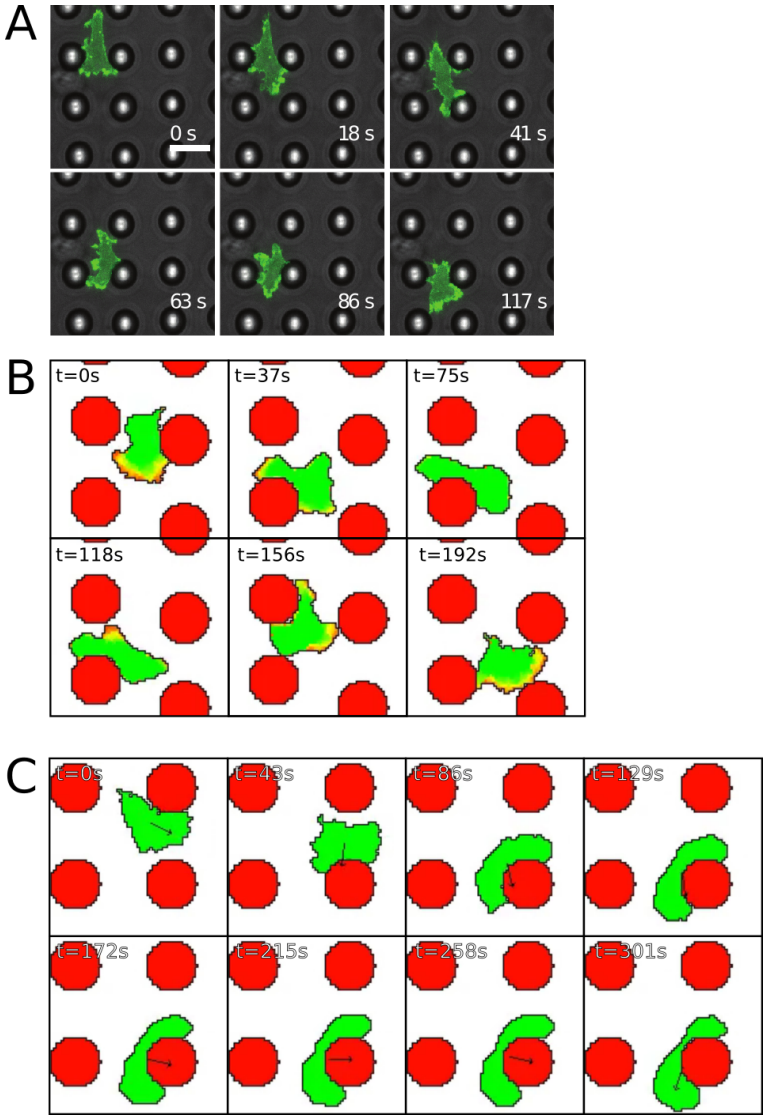


FIGURE 4.3: **Examples of frontal collision of *in vitro* and simulated starved cells with a pillar A)** *in vitro D. discoideum* starved cell in a pillar grid. The arrow indicates qualitatively the direction and magnitude of motion. Cell cytoplasm is labeled fluorescently. **B)** Active reorientation model: Act-levels are colored from red (highest level) to green (Act-level=0). Once the polarized cell hits the pillar, it looses its polarization. A new Act-front appears, the cell polarizes again and moves away from the pillar. **C)** Passive reorientation model simulation: the arrow is the preferential direction vector of the cell, starting at its center of mass. Once the cell collides with the pillar, it can only move away once its directional vector points away from the pillar (final frame).

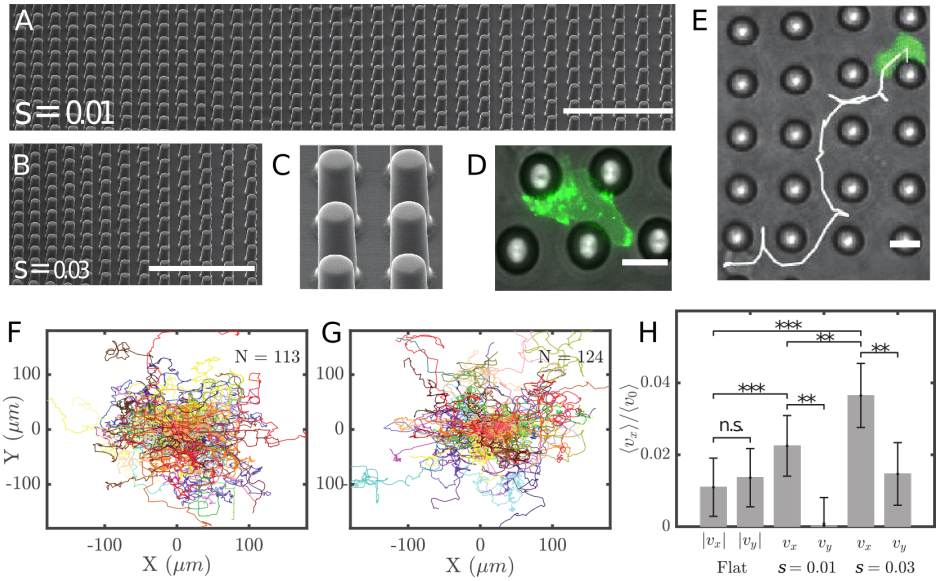


FIGURE 4.4: **Large scale topotaxis for highly-motile amoeboid cell migration (*D. discoideum*) A)** SEM micrograph of the $s = 0.01$ pillar field. Scale bar is 100 μm. **B)** Same micrograph, but for $s = 0.03$. **C)** Detailed micrograph of pillars ($h = 20\mu m$, $d = 10\mu m$). **D)** Actin polymerization hot spots visualized for a cell migrating through a pillar field. Fluorescence is LimE-GFP expressed in LimE null cells [195, 196] after z-projection, scale bar is 10 μm. **E)** Bright field image of a pillar field ($s = 0.03$) overlaid with the trajectory of a migrating starved *D. discoideum* cell (free cytoplasmic GFP in green). Scale bar is 10 μm. **F)** Trajectory plot of cells moving on topotaxis field $s = 0.01$. **G)** Trajectory plot of cells moving on topotaxis field $s = 0.03$. **H)** Drifts ($v_{x,y}$) measured in all live cell experiments (flat, $s = 0.01$ and $s = 0.03$) compared. The anisotropicity in pillar positioning was in the x direction for all non-flat experiments. The topotactic drift (v_x) was found to be significantly higher ($p \leq 0.01$) than noise (v_y) for all topotaxis assays. Additionally, the topotactic drift was found to increase ($p \leq 0.001$) with a rising gradient.

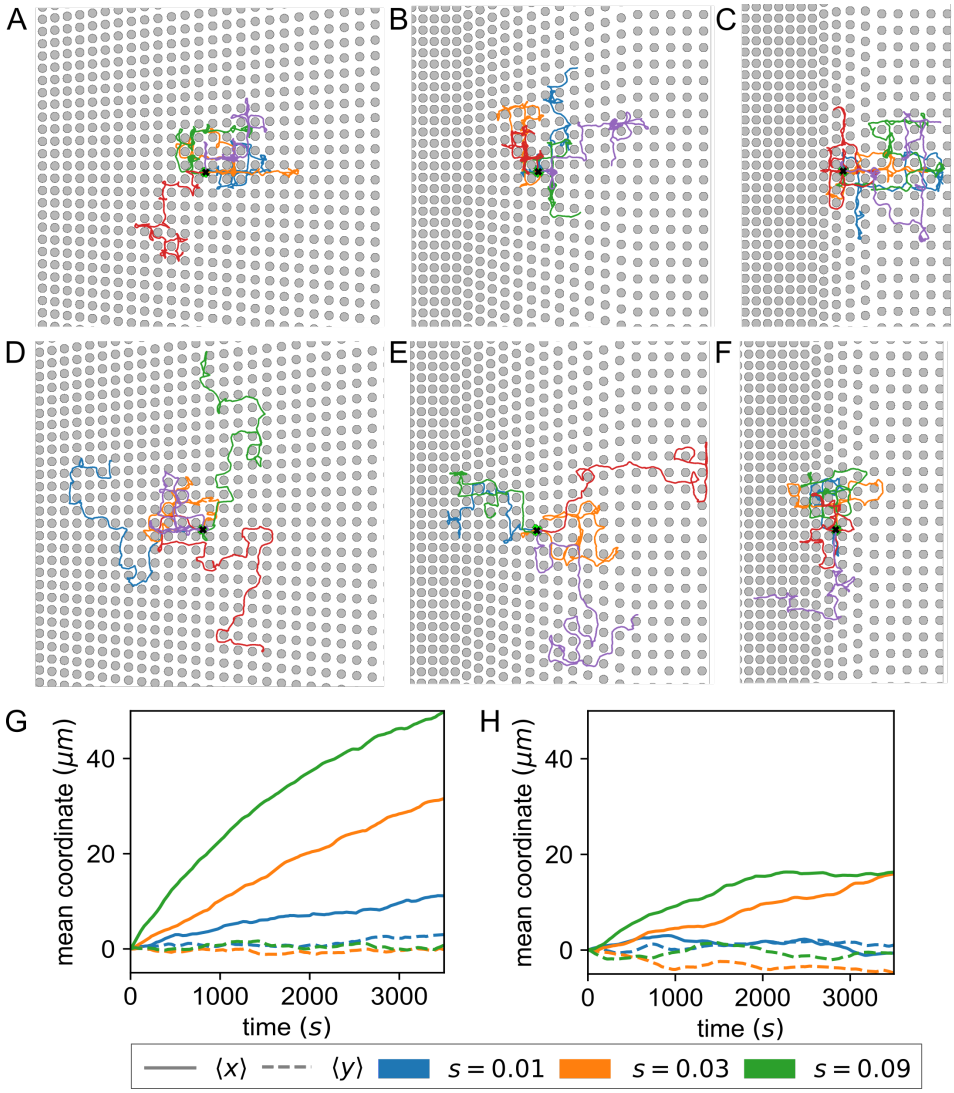


FIGURE 4.5: **Trajectories of starved simulated cells and means of x and y coordinates over time for different gradient steepness parameter (s)-values.** *Top row* active reorientation model example trajectories: **A)** $s = 0.01$, **B)** $s = 0.03$, **C)** $s = 0.09$. *Middle row* passive reorientation model examples: **D)** $s = 0.01$, **E)** $s = 0.03$, **F)** $s = 0.09$. Starting location is marked with a black cross and the depiction of a cell. Each trajectory has its own color. *Bottom row* mean x and y coordinates over time for gradient steepness s corresponding to the gradients shown in the rows above. **G**) Active reorientation model. **H**) Passive reorientation model.

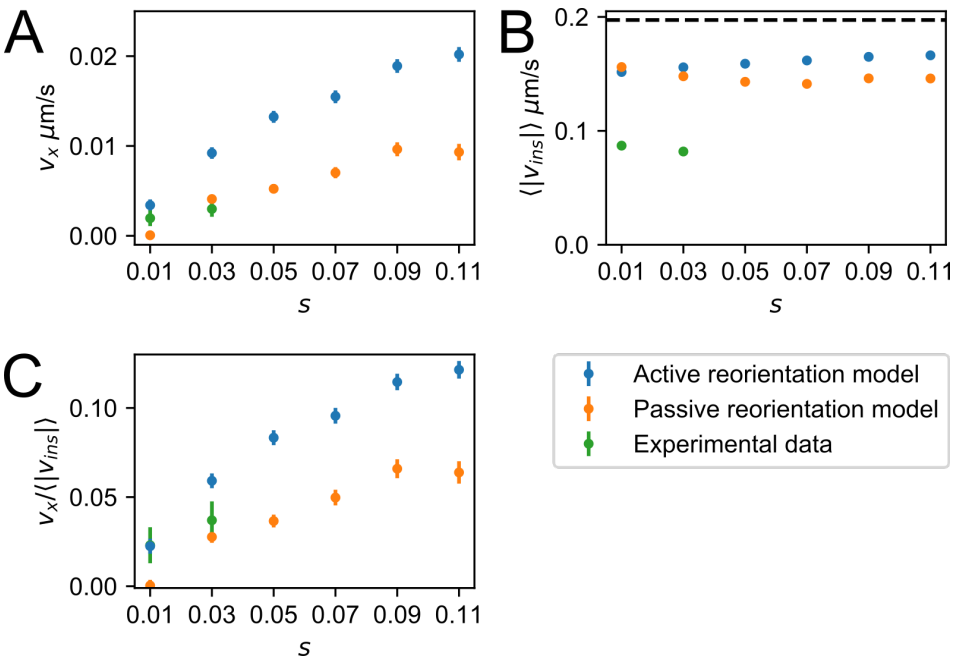


FIGURE 4.6: **Topotactic drift and instantaneous speed of starved parameter set against gradient steepness** s**.** Each orange or blue data point represent the average of a 1000 simulations. Error bars indicate 97.5% CI. **A)** Topotactic drift v_x. **B)** Instantaneous speed $\langle|v_{ins}|\rangle$. Dotted line depicts the instantaneous speed of starved cells in absence of pillars. **C)** Relative topotactic drift $v_x/\langle|v_{ins}|\rangle$).

clear difference in v_x when comparing the two reorientation models: the passive reorientation model shows a lower topotactic drift. Although the fitted parameters of the methods do not result in the exact same persistent random walk, we wouldn't expect such a difference in topotactic drift. Nonetheless, when directly comparing the topotaxis between simulated cells and *D. discoideum* cells, our model predictions are close. For $s = 0.01$, the experimental v_x is in between the active and passive reorientation models' v_x, whereas for $s = 0.03$ both models results in higher v_x, albeit that the passive reorientation model v_x is very close to the experimentally measured v_x.

As Figure 4.3 shows, *Dictyostelium* and simulated cells collide with pillars. This decreases their speed. We measured the mean instantaneous speed $\langle|v_{ins}|\rangle$ of the cells within the pillar grid, shown in Figure 4.6B and Table 4.4. For both experimental and simulated cells, the speed does indeed

TABLE 4.4: Instantaneous speed $\langle|v_{ins}|\rangle$ in $\mu m/s$ on empty field and within pillar grid for starved *D. dictyostelium* cells, starved active reorientation model simulations and starved passive reorientation model simulations. Given error is 97.5% confidence interval.

Field	Experimental	Act based	Vector based
Empty field	0.197±0.00139	0.198±0.00285	0.197±0.001331
$s = 0.01$	0.087±0.00080	0.152±0.00046	0.156±0.000264
$s = 0.03$	0.082±0.00082	0.156±0.00047	0.148±0.000274

drop and is not dependent on pillar gradient steepness s. However, experimental cells show a larger decrease in speed than the simulated cells, about 50% decrease compared to cells on an empty field, versus 20% decrease for the simulated cells. The mean instantaneous speeds of both reorientation models are not far apart.

As our simulated cells move at a higher speed, we would also expect a higher topotactic ddrift. Therefore, aside from comparing absolute topotactic velocities, we also looked at the relative topotactic velocity: $v_x/\langle|v_{ins}|\rangle$, (Fig. 4.6C). For $s = 0.01$, the relative v_x of the experimental data is very similar to that of actively reorienting simulated cells, whereas for $s = 0.03$, the experimental v_x lies between the value of the two reorientation modes, and is closest to the passive reorientation model predictions. We conclude that our models make good prediction for relative topotaxis. Furthermore, the similar mean instantaneous speeds but different topotactic drifts for the active reorientation and passive reorientation model yield different relative topotactic drifts. The active reorientation model's relative topotactic drift is on average twice as large as the passive reorientation model's. This indicates that the persistence method itself causes part of the difference in topotaxis drift, and could be due to CIL in the active reorientation model.

4.2.6 Pillar adhesion changes velocity within the grid

To study what could give rise to the discrepancy between the instantaneous velocity of the experimental and simulated cells, we turned to the interaction between cells and pillars. For the model results in Fig 4.6, we assumed that cells neither adhere to nor are repulsed by the pillar surface, i.e. $J_{\text{cell,medium}} = J_{\text{cell,pillar}}$. However, we can test this assumption and vary the contact energy $J_{\text{cell,pillar}}$. To test whether the adhesiveness of cells to pillars affects topotaxis, we simulated cell movement in a grid with gradient $s = 0.03$ for different contact energies. We let $J_{\text{cell,pillar}}$ range from -60 to 60, where $J_{\text{cell,pillar}} = 20$ is neutral with respect to $J_{\text{cell,medium}} = 20$. Examples

of cell tracks under a subset of this $J_{\text{cell,pillar}}$ range are shown in Fig. 4.7. Clear is that for very low $J_{\text{cell,pillar}}$ (Fig. 4.7A,B,D,E), cell displacement is reduced, and more extremely so for actively reorienting cells than passively reorienting cells. High $J_{\text{cell,pillar}}$ (Fig. 4.7C,F) does not seem to affect cell displacement much compared to neutral $J_{\text{cell,pillar}}$ (Fig. 4.5B,E). A first glance on topotactic drift (Fig. 4.7G,H) shows that the active reorientation model loses it's drift at low $J_{\text{cell,pillar}}$, but not at high $J_{\text{cell,pillar}}$. The passive reorientation model is also affected by pillar adhesiveness, but still displays a drift at low $J_{\text{cell,pillar}}$.

We further quantified the topotactic drift more precisely. Figure 4.8A shows the topotactic drift as a function of $J_{\text{cell,pillar}} - J_{\text{cell,medium}}$ for starved cells. There is an effect of decreasing the adhesion energy $J_{\text{cell,pillar}}$ to negative values, which decreases the topotactic drift. For very negative energy differences, actively reorienting cells eventually do topotaxis in the other direction. However, closer inspection of the cell tracks shows that cells move to the two initial pillars on their left side and get stuck there, see Fig. 4.7A. Increasing $J_{\text{cell,pillar}}$ with respect to $J_{\text{cell,medium}}$ does not seem to have an effect on v_x, suggesting that the topotactic drift saturates with increasing $J_{\text{cell,pillar}}$. Again, we see a difference in the topotactic drift between the two persistence methods (Fig 4.8A).

Aside from collision with pillars, cell-pillar adhesion is now an extra interaction that can influence cell speed. Looking at the $\langle|v_{ins}|\rangle$, we see again differences between the two persistence methods (Fig. 4.8B). For the actively reorienting cells, $\langle|v_{ins}|\rangle$ drops quickly when lowering $J_{\text{cell,pillar}}$, is maximal around $J_{\text{cell,pillar}} = 10$ and slightly decreases for further increase in $J_{\text{cell,pillar}}$. The large drop in $\langle|v_{ins}|\rangle$ confirms the observation on the examples of cell tracks (Fig 4.7A), as cells like to adhere more and more to the pillars and hardly displace. For the passively reorienting cells, $\langle|v_{ins}|\rangle$ doesn't decrease as dramatically as for actively reorienting cells for negative $J_{\text{cell,pillar}}$. Also the optimum in $\langle|v_{ins}|\rangle$ isn't as clear and appears for higher $J_{\text{cell,pillar}} = 20$.

When we compare the results to the experimental data, we can make two observations. First is that the model prediction is close to the v_x of *D. discoideum* cells for $J_{\text{cell,pillar}} - J_{\text{cell,medium}} = -30$. Second is that for $J_{\text{cell,pillar}} - J_{\text{cell,medium}} = -30$, $\langle|v_{ins}|\rangle$ is still higher for the simulated data than the experimental data, resulting in underestimation of the relative topotactic velocity. The values of $J_{\text{cell,pillar}}$ for which the instantaneous speeds are indeed similar between model and experiment are $J_{\text{cell,pillar}} - J_{\text{cell,medium}} = -80$ and -50 for the passive reorientation model and active reorientation model respectively, and are already extreme in limiting cell displacement in general.

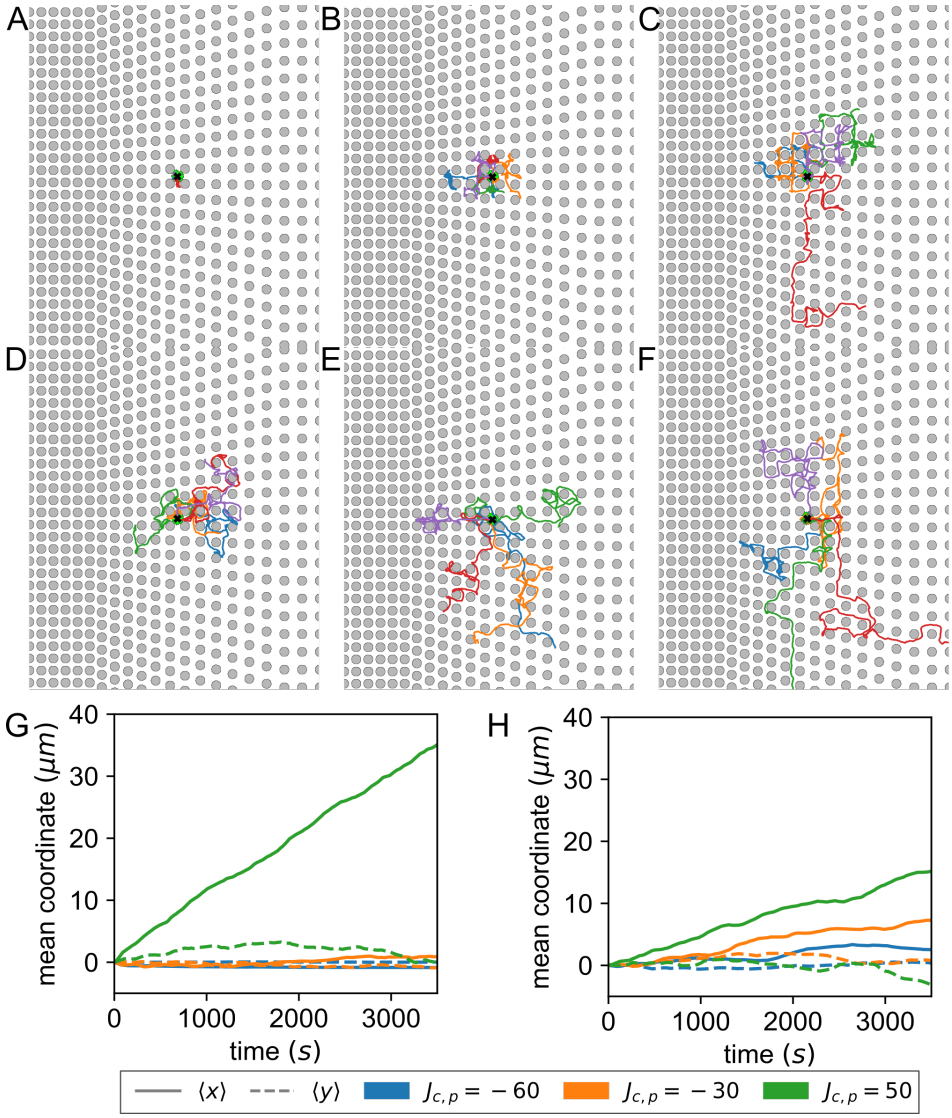


FIGURE 4.7: **Trajectories of starved simulated cells and mean x and y coordinates over time for different pillar adhesion energy $J_{\text{cell,pillar}}$.** *Top row* active reorientation model example trajectories: **A)** $J_{\text{cell,pillar}} = -60$, **B)** $J_{\text{cell,pillar}} = -30$, **C)** $J_{\text{cell,pillar}} = 50$. *Middle row* passive reorientation model example trajectories: **D)** $J_{\text{cell,pillar}} = -60$, **E)** $J_{\text{cell,pillar}} = -30$, **F)** $J_{\text{cell,pillar}} = 50$. Starting location is marked with a black cross and the depiction of a cell. Each trajectory has its own color. *Bottom row* mean x and y coordinates over time of a 1000 independent simulations. **G)** Active reorientation model. **H)** Passive reorientation model.

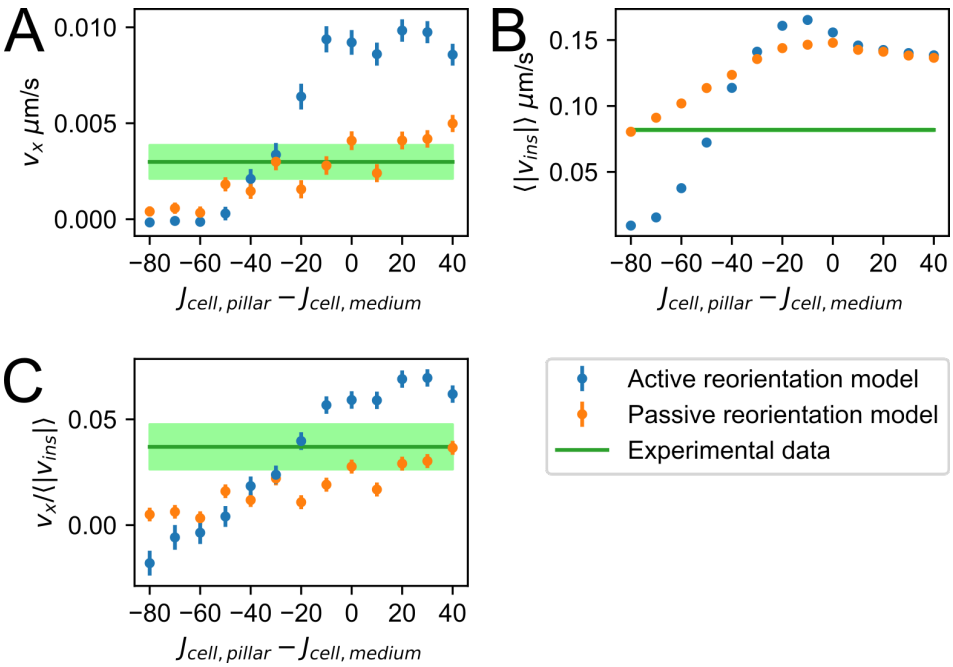


FIGURE 4.8: **Topotactic drift and instantaneous speed of starved parameter set against adhesion preference $J_{\mathbf{cell,pillar}} - J_{\mathbf{cell,medium}}$.** For all points $J_{\text{cell, medium}} = 20$. Each data point represent the average of a 1000 simulations, error bars indicate 97.5% confidence interval. Line with shaded area represents the experimental data with 97.5% confidence interval. **A)** Topotactic drift v_x. **B)** Instantaneous speed $\langle|v_{ins}|\rangle$. **C)** Relative topotactic drift $v_x/\langle|v_{ins}|\rangle$.

Overall, we conclude that adhesion energy between cell and pillar can indeed lower the velocity of cells. However, the drop in speed is paired with a drop in displacement as well, and the values at which it occurs at are not in agreement for when v_x coincides with the experimental v_x.

4.2.7 Effective persistence decreases sooner in active reorientation model than in passive reorientation model

In Schakenraad et al. [75], topotaxis was explained as a change in effective persistence length $l_{\text{eff}} = v_{0,\text{eff}} \cdot \tau_{\text{eff}}$ and effective diffusion D_{eff} due to the presence of pillars, which they verified by simulating a regular grid for a range of different pillar distances d. To further delve into the difference in topotaxis efficiency between the active and passive reorientation model, we also looked into the effective persistence length and effective diffusion over a range of pillar distances. We obtain τ_{eff} by fitting the MSD with Fürth's equation with translational diffusion (Eq. 4.10, see Methods section) and D_{eff} by fitting the MSD at larger time lags with a straight line, similarly to how we obtained them for fitting the empty field data, and obtain $v_{0,\text{eff}}$ by computing $\langle |v_{ins}| \rangle$.

Figure 4.9 shows the effective parameters of the fitted random persistent walk of both models using the starved parameter setting. The two persistence methods react differently to the presence of pillars in regular grids. The effective diffusion coefficient, persistence time and persistence length show a similar trend (Fig. 4.9ABD): they increase as pillar distance increases and saturate to the corresponding values of the empty field measurements. However, this happens more rapidly for the passive reorientation model than for the active reorientation model. The difference in persistence length is mostly due to the difference in persistence times between the two models, as the differences in velocity are minor (Fig. 4.9C). As τ_{eff} is the only measure we determine by fitting Eq. 4.10, we checked whether Eq. 4.10 fits the MSD well. Except for both models on $d = 13.1\mu m$ and the active reorientation model with $d = 14.7\mu m$, the fits seem good. For $d = 13.1\mu m$, the active reorientation model's MSD indicated subdiffusive behaviour: i.e. the cells get stuck between the pillars, whereas the passive reorientation model's MSD did show long term diffusive behaviour but its fit was off for the shorter time scale. The fit for the active reorientation model at $d = 14.7\mu m$ seems to overestimate the persistence time. Nonetheless, at larger d the fits are good and τ_{eff} is decreased more in the active reorientation model. So although the two models behave similarly in an empty field, their behaviour in regular pillar grids is very different, and the persistence of the active reorientation model is more affected by the presence of pillars.

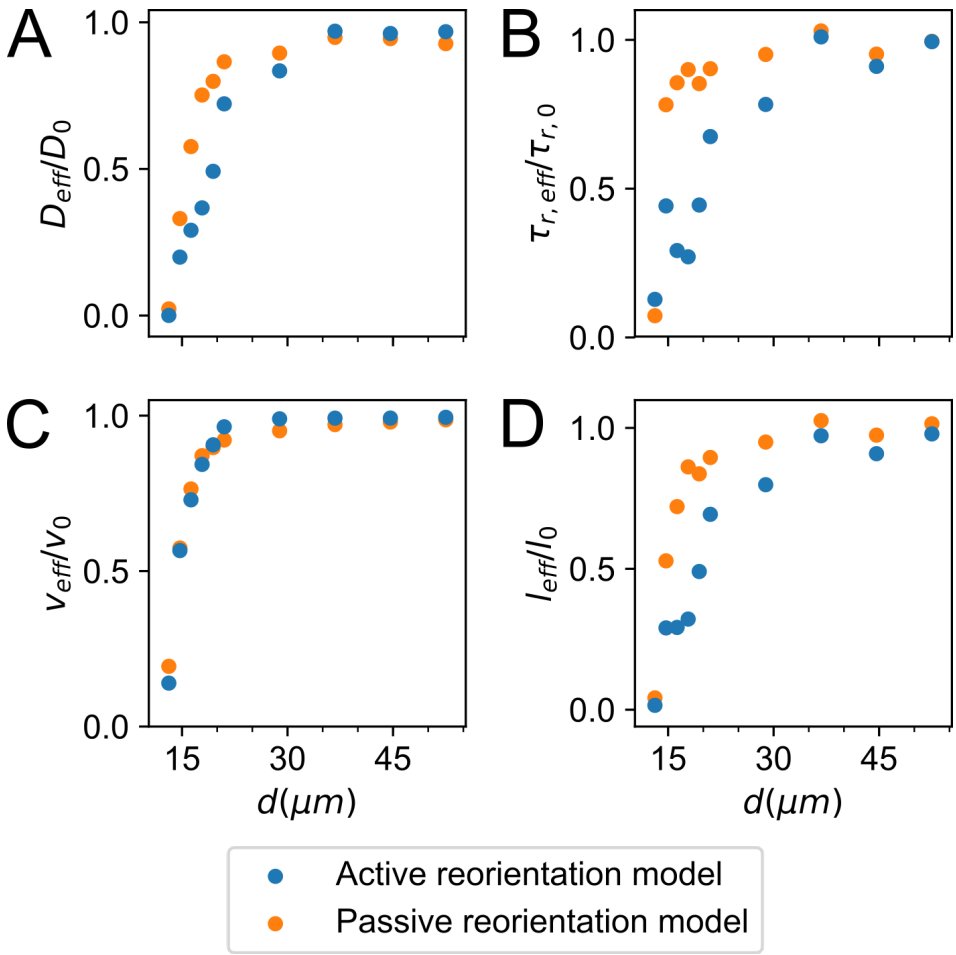


FIGURE 4.9: **Normalized effective parameters of the persistent random walk in regular pillar lattices for the active and passive reorientation model with the starved parameter settings. A)** Effective diffusion coefficient D_{eff} normalized by the diffusion coefficient of starved simulations on an empty field as a function of the distance d between pillar centers. **B)** Effective persistence time τ_{eff} normalized by the persistence time of starved simulations on an empty field as a function of d. **C)** Effective instantaneous velocity v_{eff} normalized by the instantaneous velocity of starved simulations on an empty field as a function of d. **D)** Effective persistence length $l_{\text{eff}} = v_{\text{eff}}\tau_{\text{eff}}$ normalized by the persistence length of starved simulations on an empty field as a function of d.

4.2.8 Model predicts small topotactic drift for vegetative cell topotaxis, but independent of persistent mode

Although no experimental data is available on topotaxis in vegetative cells in these pillar gradients, we are still interested in whether vegetative cells could perform topotaxis. The largest difference in motility between starved and vegetative cells is the lower speed and lower persistence time in the vegetative cells (Table 4.1), and hence, a lower persistence length of about 2.3 µm. The minimum distance between pillars in our pillar grids is significantly larger than that. Hence, we would not expect the persistence of the cell to be altered much in the presence of pillars and therefore deem it unlikely that persistence driven topotaxis contributes much to topotaxis of vegetative cells.

We use our model to predict whether vegetative cells topotax or not. We can see the lower motility of vegetative cells reflected in the example trajectories of our simulated vegetative cells (Fig. 4.10) where cells only move in the order of magnitude of 1-3 cell lengths on the same time scale as in Figs. 4.5,4.7. Still, there is a visible increase in the mean x coordinate but not mean y coordinate (Fig. 4.10), so vegetative *Dictyostelium* cells can likely perform topotaxis as well.

With the lowered displacement of vegetative cells, also v_x is an order of magnitude smaller than that of starved cells (Fig. 4.12A). For $s = 0.01$, the topotactic drift does hardly exceed the drift in the y-direction, (student's t-test $p = 0.01$ for actively reorienting cells, $p = 0.67$ for passively reorienting cells) which indicate how small the topotaxis effect is in a shallow gradient. As the v_x confidence intervals for the starved cell simulations are much smaller compared to those of the experimental starved cells, we expect that it will probably be very hard to measure topotaxis in the experimental system with vegetative cells. The topotactic drift does still increase for higher s , similar to the starved cells. Unlike for the starved cells, the different persistence methods results in similar v_x (Fig. 4.12A) and similar v_{ins} (Fig. 4.12B) and hence also in similar relative topotactic velocity (Fig. 4.12C) for the vegetative cell parameter set. Interestingly, the relative topotactic drift of vegetative cells lies in the range of those of starved vector persistent cells (Fig. 4.6C).

As we did see topotaxis occurring in the vegetative cells, we also looked into the behaviour of these cells in regular grids. We again checked their effective diffusion, effective instantaneous velocity, effective persistence time and effective diffusion length as a function of distance between pillar centres, similar to previously done for the starved cells. Again, we obtained τ_{eff} by fitting Eq. 4.10 to the MSD. For the vegetative cells, we observed

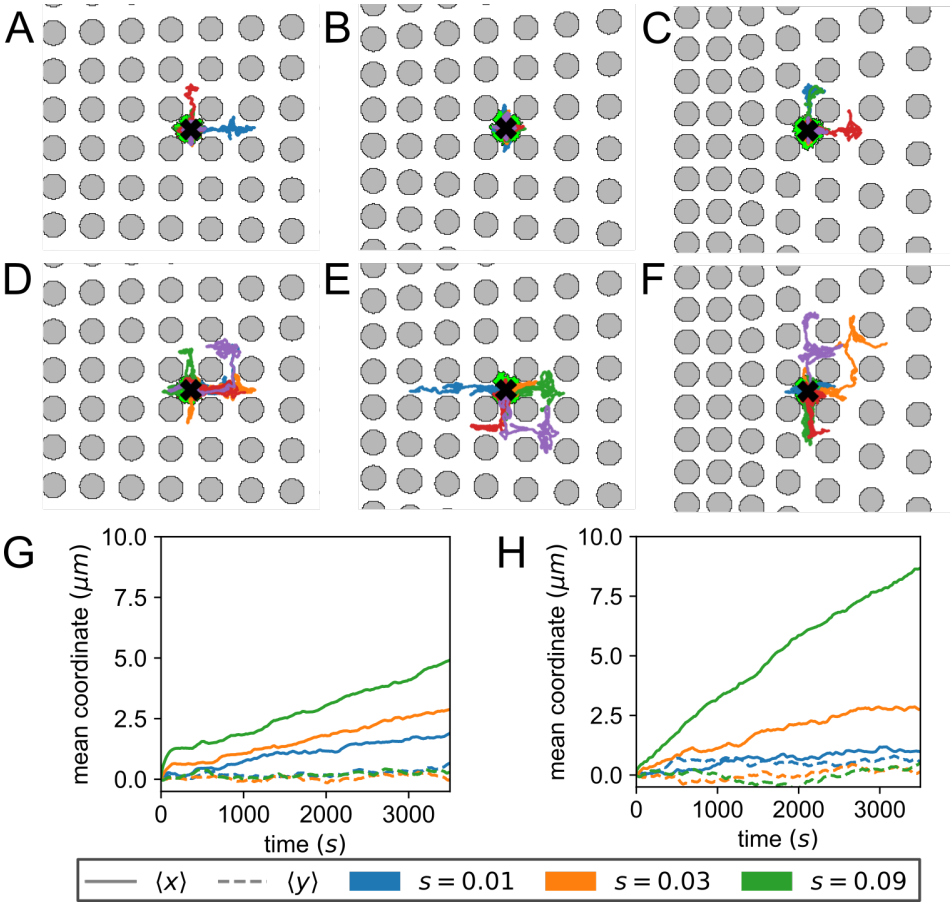


FIGURE 4.10: **Trajectories of vegetative simulated cells and means of x and y coordinates over time for different gradient steepness s.** *Top row*: active reorientation model trajectory examples: **A)** $s = 0.01$, **B)** $s = 0.03$, **C)** $s = 0.09$. *Middle row*: passive reorientation model. Trajectory examples: **D)** $s = 0.01$, **E)** $s = 0.03$, **F)** $s = 0.09$. For the trajectory plots: starting location is marked with a black cross and the depiction of a cell. Each trajectory has its own color. *Bottom row* mean x and y coordinates of 1000 independent simulations. **G**) Active reorientation model. **H**) Passive reorientation model.

that for the active reorientation model, Eq. 4.10 is not a good descriptor of the MSD for low pillar distance $d = 13.1\mu m$ to $d = 16.2\mu m$. Hence, we only show the results in Figure 4.11 for $d > 16.2\mu m$. The effective parameters are more similar between the persistence methods than observed for the starved parameter set (Fig. 4.9). Only on the smallest displayed pillar distances d, we see that the effective persistence time of the active reorientation model is more affected than of the passive reorientation model. Compared to the differences between the persistence methods for the starved cells, this difference here is smaller in both relative and absolute manner. Overall, these effective parameters do not indicate a clear difference between the persistence methods for vegetative cells.

4.3 Discussion

In this paper we have measured topotaxis by *Dictyostelium discoideum* cells in a pillar gradient as well as simulated topotaxis in a similar grid using the Cellular Potts model. Persistent *Dictyostelium* motion was modelled through two distinct persistence modules in our model: the Act model, and a vector-based persistence of which the direction changes over time. We fitted our model parameters to match *Dictyostelium* motion on an empty field before we simulated topotaxis in a pillar grid.

For starved cells, we saw that, for shallow gradients, increasing the gradient steepness increases topotaxis, both in the experiments and the simulations. This is in agreement with [75]. The topotactic velocity of both persistence methods follows a linear trend in gradient steepness, but the actively reorienting cells are more efficient in topotaxis than the passively reorienting cells, regardless of having similar instantaneous velocity.

We further looked into the cause of difference between the active and passive reorientation model and found that the effective persistence length on regular pillar grids is more affected by short pillar distances in the active model than in the passive model. The difference in effective persistence length arises from the difference in persistence time. We conclude that the active reorientation model is more likely to change orientation upon collision than the passive reorientation model, explaining the lower effective persistence time and effective persistence length.

Aside from the difference between our two models, the instantaneous velocity is a major point of difference between the experimental and simulation results. By lowering the adhesion energy between cell and pillar, we were able to lower the instantaneous velocity of the simulated cells. However, this also resulted in reduced displacement, to the point that

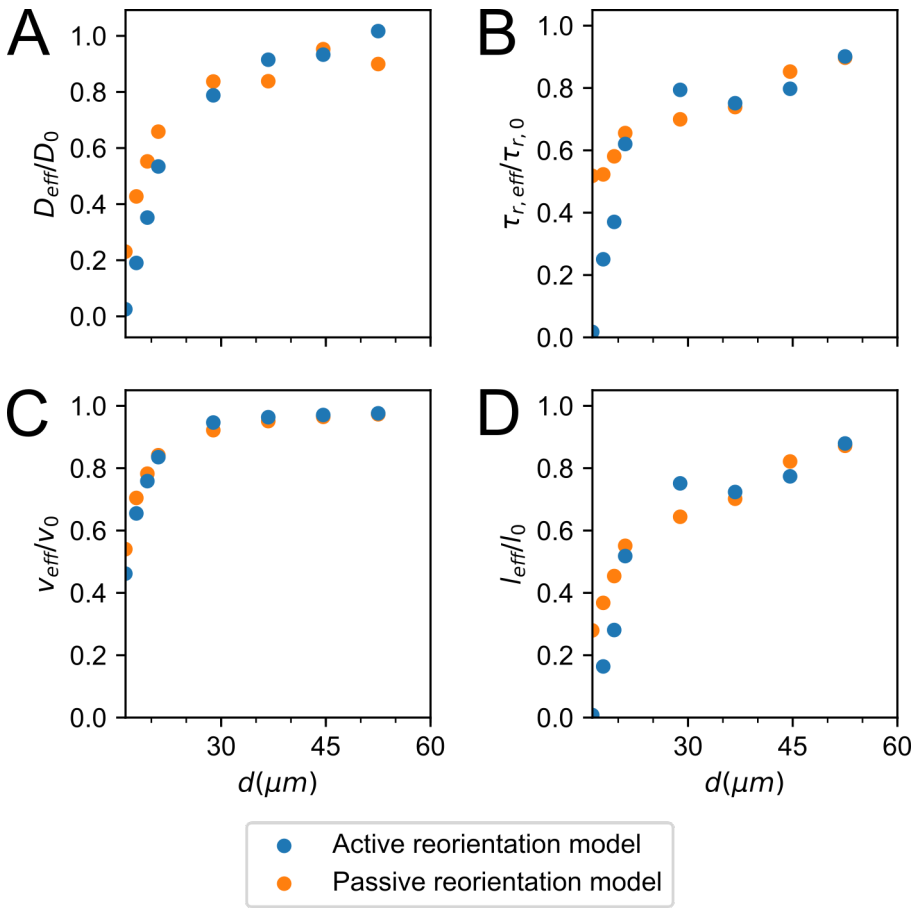


FIGURE 4.11: **Normalized effective parameters of the persistent random walk in regular pillar lattices for the active and passive reorientation model with the vegetative parameter settings. A)** Effective diffusion coefficient D_{eff} normalized by the diffusion coefficient of vegetative simulations on an empty field as a function of the distance d between pillar centers. **B)** Effective persistence time τ_{eff} normalized by the persistence time of vegetative simulations on an empty field as a function of d. **C)** Effective instantaneous velocity v_{eff} normalized by the instantaneous velocity of vegetative simulations on an empty field as a function of d. **D)** Effective persistence length $l_{\text{eff}} = v_{\text{eff}}\tau_{\text{eff}}$ normalized by the persistence length of vegetative simulations on an empty field as a function of d.

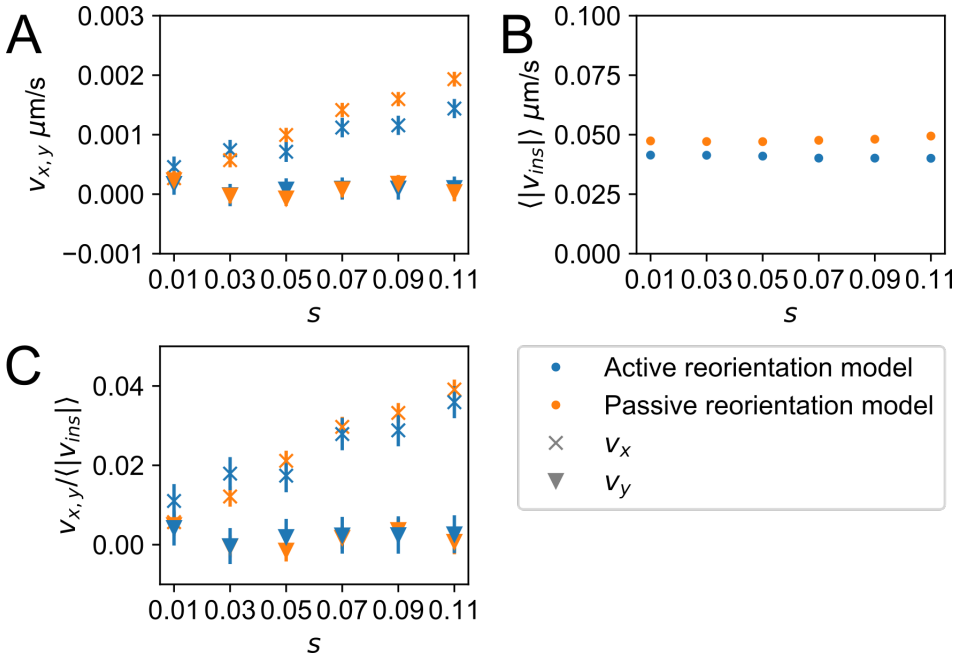


FIGURE 4.12: **Migratory drifts and instantaneous speed of vegetative parameter set against gradient steepness** s**.** Each data point represent the average of a 1000 simulations. Error bars indicate 97.5% CI. **A)** Migratory drifts v_x (lighter colors) and v_y (darker colors). **B)** Instantaneous speed $\langle|v_{ins}|\rangle$. **C)** Relative drifts $v_x/\langle|v_{ins}|\rangle$ (lighter colors) and $v_y/\langle|v_{ins}|\rangle$ (darker colors).

cells are sticking to the pillars. Hence, pillar adhesivity is unlikely the only mechanism to explain the velocity difference between experiment and simulation.

For the vegetative simulations, our model predicts that cells do undergo topotaxis distinct from directional noise, albeit an order of magnitude smaller than the starved cells, with the exception of the shallowest gradient. Also, both persistence modes resulted in similar outcomes, suggesting that there is a difference in topotaxis by quick and persistent cells and slower, less persistent cells.

4.3.1 Topotaxis and the changing effective persistence length

The physical principle underlying topotaxis presented in Schakenraad et al. [75] states that the change in effective persistence length caused by the pillar grid causes ABP to turn towards regions where they display higher effective persistent length. This idea is similar to that of how durotaxis, i.e. motion up a stiffness gradient, arises [193, 146].

When we compare the efficiency in topotaxis of both the *D. discoideum* cells and the CPM simulations to the ABP simulations, both the cells' and CPM simulations' relative topotactic velocities well exceed the 1% topotactic velocity of ABPs, the passive reorientation model at the shallowest gradient excluded. For the models, we can explain this difference by the difference in effective persistence length in a pillar grid. In our model, the effective persistence length is decreased more at small pillar distances than the ABP model. As this is true for both reorientation models, we conclude that this is an effect of the dynamic volume of our simulated cells.

Furthermore, when comparing the persistence length of the active and passive reorientation models, the difference in persistence length between the two models could explain the difference in relative topotaxis. Interesting to note here is that the change in persistence length can be ascribed to the change in persistence time and not change in velocity. We can interpret this as actively reorienting cells changing direction more drastically upon colliding with a pillar. We hypothesize that, for the active reorientation model, the new direction after collision is more uniformly distributed along the directions not pointing towards the pillar center, i.e. contact inhibition of locomotion, whereas reorientation direction for the passive reorientation model is more skewed towards the previous direction. This reorientation effect can likely be exaggerated by looking at concave obstacle shapes.

In this perspective, we can also look at the vegetative simulations. As the persistence length and persistence time for vegetative cells is very short,

the differences in how the models reorient upon collision becomes negligible. We also observed hardly any difference in the effective parameters within regular grids between the two persistence methods, except for small distances between pillars. The lack of difference in these parameters is likely why the resulting topotaxis is so similar between the models.

The reorientation of cells upon obstacle collision has also been subject of other studies [197, 198, 199]. In an experiment with fish keratocytes, actin flow was disturbed upon collision with an obstacle, making the keratocyte change direction [197]. Modelling studies also indicate that intercellular molecular dynamics are important in cell reorientation upon collision. In Nishimura et al. [198] the change in actin retrograde flow plays an important role in the formation of a new protrusion after collision with a wall, allowing the cell to move away from the wall. In Campbell et al. [199] many different cell behaviours after collision were observed in a immersive-boundary method with reaction-diffusion equations on the cell surface. They observed 1) 'freezing' when the cell did not have new pseudopods for a short while after collision, 2) 'doubling back' when the cell return in the direction it came from and 3) 'tug of war' were the new pseudopods competed for leading the cell's new direction.

4.3.2 Dictyostelium on silicon and *in silico*

We saw a disparity between the instantaneous speed of the *D. discoideum* cells and the simulated cells when they were within the pillar grid. This difference is unlikely to be caused by cells adhering to the pillars. Some other mechanisms could be at play, such as dimensionality, curvotaxis and chemotaxis. We will shortly address each of them.

Dimensionality

Our Cellular Potts model is a 2D model. However, the experimental cells crawled in a 3D pillar grid and hence can use the third dimension while moving around the obstacles. Experimental observation show that *D. discoideum* cells sometimes crawl up a pillar, mainly displacing their center of mass vertically. They can also move through narrower pillar distances than our CPM simulation by spreading their volume vertically. Although vertical displacement is not modeled directly in the CPM, the weight of the area and perimeter constraint can be seen as a measure of 3D flexibility. A possibility would be to extend our models into a 3D CPM, or use other computational methods, such as in [199]. Furthermore, how our results translate to 3D matrix environments is still unclear. As shown in [200], cells

migrate very differently in 1D and 3D matrix environment compared to 2D matrix environments, so translating topotaxis from our experiments to more complex structures is non-trivial.

Curvotaxis

Aside from the dimensionality difference, there is another aspect to the experimental setup. The interpillar distance is slightly curved, as the pillars are broadening at their feet. It has been shown that cells are responsive to substrate curvature on the cell-scale [201, 202, 203], and tend to move towards concave “valleys". The interpillar spaces with concave curvature fit this description, and could alter the speed of *Dictyostelium* cells.

Chemotactic sensing

More guiding principles can be at play. *Dictyostelium discoideum* cells are known to secrete cAMP when starved, but can also degrade cAMP when in high concentrations [204]. This chemical functions as a chemoattractant. Usually, this will lead to multiple cells to find each other and aggregate. In a modeling study it has been shown that these aggregates can avoid obstacles through the perturbations in the chemical field caused by such obstacles [205]. However, not many studies have been done on how this affects single *Dictyostelium* cells and whether they can sense their own secreted cAMP. Nonetheless, multiple studies have shown how chemical sensing of the environment can guide cell movement around large obstacles [206], or through mazes [207]. Self-secreted chemoattractant can also trap cells within containing environments such as dead ends in mazes [208]. The effect of chemical sensing in a field with a high density of obstacles as well as its effect in topotaxis is still open for study.

Concluding, our model is a much closer match to the behaviour of *Dictyostelium discoideum* cells in pillar grid with a distance gradient than the previous, analytical ABP model. The deformable cell volume within our models allows for more efficient topotaxis as it lowers the effective persistence length, especially when pillar distance is close to the cell length. We also showed that how cells interact upon collision plays an important role in the effective persistence length and hence, topotaxis.

4.4 Materials, Methods, and Model

For this study, we have used an experimental setup with live cells and a computational model. Experimental data from cells on a flat surface were

used to set the model parameters. Topotaxis measurements were done in both the experimental as well as the computational setup.

4.4.1 Live cell experiments

Cell culture and experiment preparation

For all migration experiments, Axenic *D. discoideum* (Ax2) with a cytoplasmic green fluorescent protein (GFP) insertion was used (strain HG1694, obtained from Dr. G. Gerisch, MPI for Biochemistry, Germany). Cells were grown at 20 °C in HL5 medium, supplemented with 20 $\mu g ml^{-1}$ gentamicin (Gentamycin solution, Merck, Netherlands) as a selection antibiotic. The cells were cultured in 100 mm petri dishes (100 mm TC-treated culture dish, Corning, USA) and confluency was kept below 70% during culturing. For visualizing actin polymerization, a LimE-GFP in lim0 cell line was used [195]. These cells were cultured similarly, but with two selection antibiotics, 10 $\mu g ml^{-1}$ Gentamycin and 10 $\mu g ml^{-1}$ Blasticidin, necessary to maintain the double mutation (LimE-GFP and Lim0).

In preparation for imaging experiments, cells were harvested by pipet induced flow and collected in a conical tube. To remove the culture medium the cells were centrifuged at 1500 rpm for 3 min. In case of vegetative experiments, resulting cell pellets were thrice washed using non-fluorescent buffer (3.6mM KH_2PO_4, 2.9mM Na_2HPO_4, PH 6.7) and, after resuspension, transferred onto (un)structured polydimethylsiloxane (PDMS) surfaces placed inside an imaging chamber (see Sec. 4.4.1). In case of starved experiments a pulsation procedure was started before imaging instead. Cell pellets were thrice washed with 17 mM K-Na-phosphate buffered saline (PBS, pH 6.0) and placed on a shaker for 1 hour. Then, to induce cAR1 expression, cells were pulsed with 150 nM cyclic adenosine-monophosphate (cAMP, Merck, The Netherlands) applied in 6 minute intervals over 4 hours while shaking. After pulsation, any residual cAMP was removed by centrifugation and resuspension. Cells were left to shake in a conical tube with PBS for another 30 min before being loaded onto PDMS surfaces in imaging chambers. For both experiments cells were left to adhere for 1 hour after insertion into the imaging chamber, leading to a 6-7 hour starvation period. Cell seeding concentrations were kept below $1 \cdot 10^{-4}$ μm^{-2} to limit cell-cell interaction and enable studying migration of individual cells.

Obstacle and flat PDMS surfaces

PDMS (Sylgard 184 Silicon Elastomer Kit, Dow Corning, USA) was mixed 1:10 resulting in a 1.72 MPa stiffness for all surfaces [209]. Flat surfaces were

prepared by spin-coating pre-mixed PDMS onto polished Silicon wafers (Siegert Wafer, Germany) and then baked for 4 hrs at 110 °C. Before spin-coating, wafers were silanized by Trichloro(1H,1H,2H,2H-perfluorooctyl) silane deposition under vacuum (50 mbar) for 1 hour to ensure proper PDMS detachment later. Before use, PDMS was cut, peeled off, and washed with ethanol (70%). For migration experiments PDMS was cast around a 250-500 µm thickness, for limE-GFP imaging, PDMS was cast ultra-thin (<50µm) enabling 100x (WD= 130µm) imaging.

Pillar obstacle fields were prepared using a molding process. The pillar molds were prepared by two-photon direct laser writing (DLW) using the Photonic Professional GT (Nanoscribe, Germany). First, a negative of the topotaxis pillar lattices ($s = 0.01$-0.03, see Section 4.4.2) were designed using Inventor (Autodesk, USA) and, via a stereolithography format (.stl), imported to DeScribe (Nanoscribe, Germany) to prepare for DLW. Then, two-photon crosslinking was performed using the IP-S resin (Nanoscribe, Germany) deposited on a silicon wafer. Different laser powers and scan speeds were chosen for bulk and edges of the structure, 27% and 42% (of 140 mW), 10 and 40 mm/min respectively. To remove excess resin, molds were developed for 45 min in polyglycidylmethacrylate (PGMA). After blow drying with nitrogen, wafers were silanized as described previously. To produce the pillar field casts, PDMS was deposited over the wafer (with mold on top), baked, cut, peeled off, and washed with ethanol (70%).

PDMS inserts were hydrophilicitized by 15 min of UV/Ozone exposure (UVO-42, Jelight Company, U.S.A) and placed inside an imaging slide (0.8 sticky-Slide I Luer, Ibidi, Germany), to be used immediately. Before loading cell suspensions, imaging slides were washed with ethanol and then PBS.

To determine quality of the mold, each pillar field was imaged using a nanoSEM (FEI/Thermo Fisher, The Netherlands) scanning electron microscope (SEM). Samples were imaged at 10kV, with a spot size of 4.0. Before imaging, PDMS structures were coated with 2-8 nm Pt/Pa using a plasma magnetron sputter coater (208HR, Cressington, Watford, UK) to enhance conductance.

Live-cell imaging and tracking

Cells were imaged every 8-10 seconds for experiments on flat PDMS and every 20 seconds for topotaxis assays. Measurements lasted for 1-3 hours and were performed with either 10x or 20x air objectives (Plan Fluor, Nikon) on a Nikon Eclipse Ti microscope equipped with a confocal spinning disk unit operated at 10,000 rpm (Yokogawa). The cytoplasmic GFP was excited at 488nm by a solid-state diode laser (Coherent, U.S.A.) supported in an

Agilent MLC4 unit (Agilent Technologies, U.S.A.), at reduced intensity (25% of 2.4mW) controlled by an Acousto-Optic Tunable Filter. Emission was filtered by a quad-band fluorescence filter (TR-F440-521-607-700, IDEX LLC, Rochester, New York, U.S.A.). Images were captured using an exposure time of 200ms by an Andor iXon Ultra 897 High Speed EM-CCD camera (Andor Technology, U.K.). Images of higher magnification were produced with the same setup but using different objectives: Figure 4.4e with 40x air (PlanFluor, Nikon), limE-GFP Figures 4.4d and 4.3a with 60x water (Plan Apo VC, Nikon) and 4.1c and 4.2a with 100x oil (CFI plan Apo, Nikon).

Image tracking was performed using ImageJ (http://imagej.nih.gov/ij/). Microscopy time-lapse images were contrast and brightness adjusted, and ran through a Gaussian filter ($\sigma = 2$) to enable optimal cell body recognition. The ImageJ plugin CellEvaluator was used to determine the x, y-coordinates of the center-of-mass of each cell body in each frame [210]. By linking all these x, y-positions together, cell trajectories were obtained.

Empirical measurement of cell area

Fluorescence microscopy of cell migration experiments on flat PDMS were used to fit the cell surface area (A_σ) parameter in the CPM. Frames used for area analysis were chosen sufficiently far apart ($f^{-1} = \delta t = 200$ s) to avoid correlations. The two-dimensional projection of the cell body was determined using Sobel edge detection, applied to cytoplasmic GFP images obtained using the 20x air objective (0.657 µm/pix). Brightness and contrast were adjusted to rescale pixel intensities (i.e. only use the range of 16-Bit intensities of GFP signal detected). Image analysis was performed using Matlab, and the Image Processing Toolbox (Matlab v2019a, MathWorks, U.S.A.) in particular. First, edges were detected (Sobel) using an appropriate threshold, then the resulting binary edge-image was dilated, holes were filled, borders cleared and the image eroded (equal to the initial dilation). Then all groups of non-zero, adjacent pixels forming a cell were identified and properties (using *regionprops*) like area extracted.

Amoeboid movement on flat and pillar surfaces

We analysed cell migration trajectories, measured on flat and pillar PDMS, with an in-house Matlab code (version 2019b, The Mathworks, U.S.A.). The empirical instantaneous velocities (v_{inst}), persistence times and diffusion constants of cell movement measured on flat PDMS were used to fit the CPM model for both cell types. These were calculated from the cell trajectories obtained in combination with the known frame rate ($f^{-1} = \delta t = 20$

s) between images. Dying or otherwise immotile cells (see [74]) and insufficiently long trajectories (flat $N < 100$, topotaxis $N < 30$) were discarded for analysis. N was chosen higher for the flat data set, to prevent noise in mean-squared displacement (MSD) values at long-lag times.

The displacement (r) of the cell between frames is given by $\bar{r}(t) = \bar{R}(t + \delta t) - \bar{R}(t)$, where $\bar{R}(t)$ are the vectors described by the x, y-coordinates of the cell center in each image. Then, the instantaneous velocity ($v(t)$) and MSD were calculated for each trajectory by,

$$v(t) = \frac{|\bar{r}(t)|}{\delta t} \text{ and } \langle \bar{r}(\tau)^2 \rangle = \frac{1}{N-k} \sum_{i=1}^{N-k} (\bar{R}(t_i + \tau) - \bar{R}(t_i))^2. \tag{4.1}$$

Here, $\tau = k\delta t$ is the lag time, N the number of points in a trajectory, and k the frame number ($k = 1, 2, \ldots N-1$). The MSDs were averaged over all trajectories, and subsequently fit to Fürths formula for persistent random motion. From this fit the characteristic persistence times (τ_r) and effective diffusion constants (D_{eff}) for both vegetative and starved *D. discoideum* cells moving on flat PDMS were extracted and used to fit the CPM.

For flat and topotaxis assays, migratory drift ($v_{x,y}$) was calculated by averaging over all displacements of all trajectories,

$$\langle v_{x,y} \rangle = \frac{1}{n-1} \sum_{i=1}^{n-1} (\bar{R}(t_i + \delta t) - \bar{R}(t_i)) \tag{4.2}$$

where n is the total number of displacements measured.

4.4.2 Model

The model is based on the 2D Cellular Potts model and hence models cells as flexible and dynamically shaped objects in two dimensions. Cell persistence is obtained using two methods: one phenomenologically models actin dynamics, the other has a set preferential direction for the cell to move in that changes over time. We fit the parameters of these methods such that the simulated cells perform the same persistent random walk as measured experimentally in the starved and vegetative *Dictyostelium discoideum* cells on the flat PDMS surfaces. After fitting the parameters, we use the found parameter settings in our simulations with a pillar grid. The grid either contains a gradient in pillar spacing in the x-direction, or is regularly spaced. In the gradient grid, we measure the average step size in the x-direction and speed of the cell. For the regular grids, we determine the effective

persistence time, persistence length, speed and diffusion coefficient. In each of our simulations, we simulate a single cell.

Cellular Potts model of persistently moving cells

The model of cell movement is based upon the Cellular Potts model (CPM) [68] with either of two extensions for modeling persistent cell movement (see Section 4.4.2). The CPM represent cells on a regular square lattice $\Lambda \subset \mathbb{Z}^2$. Each lattice site, $\vec{x} \in \Lambda$, is associated with a spin value $\sigma(\vec{x}) \in \{-2, 0, 1\}$, or *cell ID* that uniquely identifies the lattice site with the cell ($\sigma = 1$), the medium ($\sigma = 0$), or a pillar ($\sigma = -2$). The cell is represented as a collection of lattice sites marked with $\sigma = 1$.

Cell motion is modelled by updating the grid through random copy attempts. In a copy attempt a lattice site x is selected randomly, as well as one of its neighbours y. If $\sigma(x) \neq \sigma(y)$, the copy attempt can change the energy of the system. Whether a copy attempt is accepted depends on the energy change associated with it. The energy of the cell is described by the Hamiltonian $\mathcal{H}$, that contains cell-medium and cell-pillar interactions and two cell constraints: the cell area and the cell perimeter [70].

$$\mathcal{H} = \sum_{u,v} J_{\sigma_u,\sigma_v}(1 - \delta_{\sigma_u,\sigma_v}) + \lambda_{\text{area}}(a_\sigma - A_\sigma)^2 + \lambda_{\text{perimeter}}(p_\sigma - P_\sigma)^2 \quad (4.3)$$

The first term in the Hamiltonian describes the adhesion energy of the cell. Here, J_{σ_u,σ_v} describes the interaction energy between two neighbouring lattice sites u, v of types $\sigma(u), \sigma(v)$. As $J_{\text{pillar,medium}} = 0$, we only take the adhesion energy between cell and medium or cell and pillar into account. The second term describes the area constraint and penalizes deviation of the cell area a_σ from its target area A_σ. The third term is the perimeter constraint and penalizes deviation in cell perimeter p_σ from the target perimeter P_σ. The λ's indicate the weight of both constraints.

The probability of a successful copy attempt depends on the change in Hamiltonian:

$$P(\Delta\mathcal{H}_{\mathbf{x}\to\mathbf{y}}) = \begin{cases} 1 & \text{if } \Delta\mathcal{H} < 0 \\ e^{-(\Delta\mathcal{H}/T)} & \text{if } \Delta\mathcal{H} \geq 0 \end{cases},$$

with T denoting "the temperature", a term that allows for noise, as cells are able to use energy to do energetically unfavourable moves by energy expenditure. To keep track of time, the model time is expressed in Monte Carlo steps (MCS). Within a single MCS, the expectation is that each lattice site has been updated once. Since we only model a single cell in a large

field, many lattice site neighbouring pairs will not lead to cell movement. To speed up simulations, we use a rejection-free algorithm ignoring such unfruitful copy attempts.

Pillars

The pillars in our model have their own $\sigma = -2$, as not to be confused with the σs of the medium or cells. Updates involving a pillar are handled differently from other updates. As pillars are static, we do not allow a copy attempt into a pillar site. However, we allow for copy attempts from a pillar site into a cell. Although, in this case, we do not copy $\sigma(pillar)$ into the site, but $\sigma(medium)$. This allows the cells to retract from the pillars and can be seen as medium flowing back in from the third dimension which we do not model explicitly. The interaction energy $J_{\text{cell, pillar}}$ can be adjusted to represent highly adhesive or slightly repulsive pillar surfaces. In this work, our base value is set to $J_{\text{cell,pillar}} = J_{\text{cell,medium}}$, such that pillar adhesion is neutral in respect to adhesion to the medium, but we also vary $J_{cell,pillar}$ to see the effect of adhesion.

We model two different pillar grids: a regularly spaced pillar grid and a pillar grid with a gradient. The regularly spaced grids are defined by the distance between pillar centers d and the pillar radius R. We set $R = 10px \approx 5.25\mu m$ in our simulations. The pillar grid with gradient consist of three different parts: a regularly spaced part on the left of the field, a part with a gradient in pillar distance in the middle and another regularly spaced part on the right. The gradient of pillar in the middle is defined by the following set of pillar centers P,

$$P = \left\{ \vec{x} \in \mathbb{R}^2 \;\middle|\; \begin{array}{l} x_1 = \frac{d}{1-e^{-s}}(e^{sn} - 1) + \frac{d}{2} \text{ and} \\ x_2 = d\left(m + \frac{1}{2}\right) e^{sn}, \text{ with } n, m \in \mathbb{Z} \end{array} \right\}, \tag{4.4}$$

where d is the distance between pillar centers at the origin $(0, 0)$, the center of the field. We used $d = 32px \approx 16.8\mu m$ in our simulations and $d = 16.8\mu m$ in the experimental setup. The parameter s defines the steepness of the gradient. A lattice site $\vec{x}$ is assigned to a pillar if it is within a distance R from one of the pillar centers, $\vec{p} \in P$. Hence the set of pillar lattice sites becomes,

$$\{\vec{x} \in \Lambda | (\exists \vec{p} \in P)[|\vec{x} - \vec{p}| < R]\} \tag{4.5}$$

with P the set of lattice centers (Eq. 4.4), and R the pillar radius. Again, we set $R = 10px \approx 5.25\mu m$ in our simulations. For the experimental setup $R = 5\mu m$.

This gives a gradient in the horizontal axis. However, to limit the gaps between pillars to get too small or too big, we have the two regularly spaced grids at the left and right of the gradient part of the grid. The pillar spacing of those two parts is set to $d_{min} = 2R + 6px \approx 13.65\mu m$ on the left and $d_{max} = 2d - d_{min}px \approx 19.95\mu m$ on the right in our simulations, and in our experimental setup $d_{min} = 13.63\mu m$ and $d_{max} = 19.93$. See also Table 4.3 for an overview of all grid measurements.

Persistent random walker

Our Hamiltonian on its own does not lead to persistent cell motion. In order to model persistent cells, we use two different extensions of the Cellular Potts model. First, we use the Act-extension [73], which models internal cell skeleton rearrangements and is also capable of describing persistent cell motion. Secondly, there is the vector-based persistence, which inherits most of the properties of the persistence model by Schakenraad et al. [75].

Act-extension - Active reorientation The Act-extension for the Cellular Potts model adds an extra layer to this model that resembles the assembly of actin machinery [73]. Each lattice site has an extra value, called the Act-value, which can range from 0 to $\max_{Act}$, a parameter value. The Act-value outside the cell is always 0, and can vary inside the cell. If the cell has recently made an extension, the site of extension will get the Act-value of $\max_{Act}$. Each Monte-Carlo step, the Act-values will be decreased by 1, until they become 0.

The Act-values play a role in the change in Hamiltonian. If there is a copy attempt extending a lattice site $\vec{c}$ from the cell into a lattice site $\vec{m}$ containing medium, then we look at the geometric mean of the Act-levels in $\mathrm{NB}(\vec{c})$, the Moore neighbours of $\vec{c}$ that are within the cell, such that:

$$\Delta\mathcal{H}_{Act}(\vec{c} \to \vec{m}) = \frac{\lambda_{Act}}{\max_{Act}} \sqrt[|\mathrm{NB}(\vec{c})|]{\prod_{y \in \mathrm{NB}(\vec{c})} Act(y)}. \tag{4.6}$$

Otherwise, if there is an attempt of a cell retracting, then the sign changes, so $\Delta\mathcal{H}_{Act}(\vec{m} \to \vec{c}) = -\Delta\mathcal{H}_{Act}(\vec{c} \to \vec{m})$. As a consequence, cell areas with high Act levels are more likely to extend outwards of the cells and less likely to retract. This simulates the polarized actin structure of a cell in a phenomenological way. The two parameters λ_{Act} and $\max_{Act}$ can be tuned to obtain different cell motilities.

Vector-based persistence - Passive reorientation The vector-based persistence is a hybrid method between [72] and [75]. First of all, the cell has a direction indicated by θ. Cell movement along this direction is favoured. For each copy attempt extending the cell, the angle α between the displacement caused by that move and θ is computed, and coupled to the Hamiltonian as by [72]:

$$\Delta\mathcal{H}_{\text{persistence}} = -\lambda_{\text{persistence}} \cdot \cos(\alpha). \tag{4.7}$$

$\Delta\mathcal{H}_{\text{persistence}}$ is added to the general $\Delta\mathcal{H}$. For the updating of θ, we deviate from [72], and use the differential equation, the same as in [75].

$$\frac{d\theta}{dt} = \sqrt{\frac{2}{\tau}}\xi(t). \tag{4.8}$$

Here τ stands for the persistence time, and $\xi(t)$ is a stochastic white noise term, modelled by a Gaussian distribution with mean 0 and variance $\sigma^2 = 2\frac{\Delta t}{\tau}$. The term Δt couples the time of equation 4.8 to the MCS, and we choose $\Delta t = 0.1\tau$. At initialization the cell is assigned a θ from a uniform random distribution. We update θ each MCS. We can tune the cell motility of this model extension through the parameters τ and $\lambda_{\text{persistence}}$.

Empty field fitting

In order to quantitatively compare the topotaxis of simulated cells to *Dictyostelium discoideum* cells, we tried to find parameters such that the actively reorienting and passively reorienting cells behave similarly to the starved and vegetative *D. discoideum* cells on a surface without any pillars (Figs. 4.1 and 4.2). We only changed λ_{Act} and $\max_{Act}$, and $\lambda_{\text{persistence}}$ and τ freely.

The values we fitted our cells to are the surface area, instantaneous speed, effective diffusion coefficient and persistence time. The surface area and speed were used to determine the length and time scale of the CPM, whereas the effective diffusion coefficient and persistence time were the objective of our fit. We fitted for starved and vegetative *D. discoideum*. Starved cells on an empty surface displayed a persistent random walk (Fig. 4.1, Table 4.1). Vegetative cells had a larger area and were less motile (Fig. 4.2, Table 4.1)

Determining model persistence time and diffusion coefficient For determining the persistence time in our models, we first have computed the mean squared displacement of 1000 simulated cell tracks from 120 MCS

onwards:

$$\langle r(t)^2 \rangle = \frac{1}{1000} \sum_{i=1}^{1000} (R(120+t) - R(120))^2 \tag{4.9}$$

We then fitted to $\langle r(t)^2 \rangle$ the formula

$$MSD(t) = 4D_T t + 2v_0^2 \tau_r t - 2v_0^2 \tau_r^2 (1 - e^{-t\tau_r}), \tag{4.10}$$

which describes the active Brownian motion with translational noise [159]. Here τ_r is the persistence time, v_0 the constant speed, and D_T is the translational diffusion caused by thermally induced fluctuation, which is inherent in the CPM. The corresponding effective diffusion constant is described by $D_{\text{eff}} = D_T + \frac{1}{2} v_0^2 \tau_r$. However, we decided to obtain D_{eff} independently from Eq. 4.10 by fitting a line through the square displacements over time starting from $t = 2000$ MCS and divide its slope by 4.

Scoring the fit For scoring the fit we used a weighted least squares objective. We used a hillclimbing algorithm with multiple restarts to obtain a shortlist of possible parameter values. The best fifteen parameter sets were then scored ten more times to obtain the best parameter set. The resulting effective diffusion coefficient and persistence time are shown in Table 4.1. The fits for starved cells are better than the fits for vegetative cells, most likely due to the discretization of the parameter space. The optimal parameters are shown in Table 4.2.

Measuring model topotaxis

To measure topotactic drift, we have run multiple simulations for different values of s, the parameter indicating the steepness of the pillar gradient. For each parameter value, we did 1000 simulations. To measure topotaxis we computed the mean velocity in the x-direction at $20s$ intervals for all cell tracks while the cell was in the gradient part of the pillar grid. This gives our migratory drift v_x.

Acknowledgements

We thank SURFsara for the support and computing time in using the Lisa cluster computer. J.A.J.W and D.H acknowledge the Fraunhofer Society for the Fraunhofer Attract grant “3DNanoCell” for partly funding this work and thank Dr. G unther Gerisch (MPI forBiochemistry,Martinsried,

Germany) for providing axenic *D. discoideum (Ax2)* with both free GFP and limGFP in lim0 insertion.

Chapter 5

A novel function of TLR2 and MyD88 in the regulation of leukocyte cell migration behavior during wounding in zebrafish larvae

Wanbin Hu, Leonie van Steijn, Chen Li, Fons J. Verbeek, Lu Cao, Roeland M.H. Merks, Herman P. Spaink *

Abstract

Toll-like receptor (TLR) signaling via myeloid differentiation factor 88 protein (MyD88) has been indicated to be involved in the response to wounding. It remains unknown whether the putative role of MyD88 in wounding responses is due to a control of leukocyte cell migration. The aim of this study was to explore *in vivo* whether TLR2 and MyD88 are involved in modulating neutrophil and macrophage cell migration behavior upon zebrafish larval tail wounding. Live cell imaging of tail-wounded larvae was performed in *tlr2* and *myd88* mutants and their corresponding wild type siblings. In order to visualize cell migration following tissue damage, we constructed double transgenic lines with fluorescent markers for macrophages and neutrophils in all mutant and sibling zebrafish lines. Three days post fertilization (dpf), tail-wounded larvae were studied using confocal laser scanning microscopy (CLSM) to quantify the number of recruited cells at the wounding area. We found that in both *tlr2*$^{-/-}$ and *myd88*$^{-/-}$ groups the recruited neutrophil and macrophage numbers are decreased compared to their wild type sibling controls. Through analyses of neutrophil and macrophage migration patterns, we demonstrated that both *tlr2* and *myd88* control the migration direction of distant neutrophils upon wounding. Furthermore, in both the *tlr2* and the *myd88* mutants,

*Modified from Wanbin Hu et al. "A novel function of TLR2 and MyD88 in the regulation of leukocyte cell migration behavior during wounding in zebrafish larvae". *Frontiers in Cell and Developmental Biology* 9 (2021), 210

macrophages migrated more slowly towards the wound edge. Taken together, our findings show that *tlr2* and *myd88* are involved in responses to tail wounding by regulating the migratory behavior of leukocytes *in vivo*.

5.1 Introduction

Acute inflammation is characterized by the directed migration of leukocytes, which can be triggered by tissue damage [212, 213]. The function of directed leukocyte migration is to eliminate cell debris and invading pathogens, with the aim of maintaining homeostasis upon tissue damage [214]. Neutrophils and macrophages are the two crucial immune cells that engage in this process [213, 215]. Neutrophils are the first cells to rapidly respond to the site of injury, and produce cytokines and chemokines to mediate the recruitment of other cells [216, 215]. However, persisting neutrophil recruitment can release toxic granule contents to further damage tissue, and thereby is a hallmark of chronic inflammatory disease [217, 214, 218]. In comparison, distant macrophages move slower and accumulate later at the wounded area and are considered to play a role in eliminating the debris of apoptotic cells and assist in regeneration of wounded tissue [219, 213, 215, 220]. Leukocyte migration must be tightly regulated to avoid negative effects on tissue repair or further damage. Despite myriad studies on leukocyte migration in response to wounding, the underlying mechanisms are not yet completely understood [221].

Neutrophils and macrophages depend on membrane-localized pattern recognition receptors (PRRs) to sense invading microbes and associated tissue damage [222]. PRRs play a crucial role to recognize pathogen associated molecular patterns (PAMPs) of invading microbes in open wounds and damage associated molecular patterns (DAMPs) released by lysing cells [223, 224]. Toll-like receptors (TLRs) are prominent recognition factors for PAMPs and DAMPs to regulate inflammatory responses [225, 226]. Extensive studies have demonstrated that cellular distribution is different for each TLR. TLRs recognize different classes of PAMPs and trigger the production of cytokines and chemokines during infection. Two typical examples are TLR2, which senses bacterial lipoproteins [227] , and TLR4, which recognizes bacterial lipopolysaccharide (LPS) [228]. Accumulating evidence shows that high-mobility group box 1 protein (HMGB1), which is the best well known endogenous danger signal, activates inflammation by forming complexes with other DAMPs (such as single-stranded DNA, nucleosomes and LPS) to be recognized by IL-1R as well as TLR2, TLR4

and TLR9 to induce inflammatory responses [229, 230, 213]. After interacting with these PAMPs and DAMPs, TLRs initiate downstream signaling cascades that ultimately result in producing cytokines and chemokines. Importantly, the activation of downstream signaling pathway by HMGB1 has been shown to be dependent on the TLR down-stream signaling mediated by myeloid differentiation factor 88 protein (MyD88) [213, 231].

TLR2 is one of the best known PRRs and acts as a heterodimer with TLR1 or TLR6 to recognize gram positive bacteria including mycobacteria, presumably based on the specific binding to their cell wall components, such as glycolipids and glycoproteins [227, 232]. TLR2 is expressed and activated after tissue injury even in the absence of infections, like in acute ischemic injury as well as in acute liver and kidney injury [233, 234, 235, 236]. In the study of Mojumdar *et al.* (2016), it was shown that macrophage infiltration was reduced into normal muscle following acute injury in TLR2 deficiency mice [237]. In addition, Kim et al. demonstrated that TLR2 contributes to macrophage infiltration in the dorsal root ganglia after peripheral nerve injury in mice [238]. Such injury-induced TLR2 expression and activation has therefore been hypothesized to be important for human health [239, 240, 236]. Following ischemic injury in mice, TLR2 activation promotes cell permeability, lymphocyte invasion and endothelial cell migration and mediates the release of TNF-α and IL-6 [235]. TLR2-deficient mice have a defective ability to recruit neutrophils to an injured liver and fail to induce the neutrophil chemokine CXCL-2 [236]. Additionally, TLR2 contributes to chronic liver disease in a mouse model by mediating MAPK and NF-$\varkappa$B signaling pathways [241]. However, there is little knowledge of the function of Toll-like receptor signaling in cell migration of myeloid cells to epithelial wounding sites [242].

MyD88 is an essential adaptor protein for all TLRs, except TLR3 [243, 244]. MyD88 is responsible for activating downstream signaling through binding to the TIR domain of TLRs [243, 244]. A few studies have shown changes in MyD88 expression after tissue injury. Similar to *Tlr2*, the gene expression of *Myd88* is upregulated following ischemic injury in mice [245]. Moreover, the expression of *Myd88* and *Tlr2* is significantly increased in diabetic wounded mice [246]. In addition, indirect evidence indicates that Myd88 is involved in the modulation of wound healing [247, 248], but the underlying mechanism is still unclear. Although TLR signaling is important for chemokine production, little is known about the role of MyD88 in leukocyte migratory responses to tissue injuries in the absence of pathogenic infections.

In this paper we use zebrafish larvae as a model for studying leukocyte cell migration after tail wounding. The zebrafish model has become an

important vertebrate model for studying human diseases. The small size and transparency of their larvae are useful characteristics for the screening and imaging of transgenic reporter lines [249]. Zebrafish larvae are a popular model for studying functions involved in wound repair [250, 251, 252, 253, 254, 255]. The availability of mutants in Toll-like receptor signaling genes *tlr2* and *myd88* make it possible to study their roles in leukocyte migratory behavior upon tail wounding in zebrafish [256, 257, 258, 252, 255]. Tlr2 and Myd88 show a highly conserved structure in mammals and zebrafish [259]. In a previous study, we demonstrated the conserved role of *tlr2* in zebrafish as a PRR to recognize the mammalian TLR2 ligand Pam3CSK4, and identified a set of genes that are specifically expressed genes by activation of the downstream pathway of zebrafish *tlr2* [260].

In the present study, live fluorescent imaging was used to investigate the effect of the *tlr2* mutation and the *myd88* mutation on leukocyte migration upon tail wounding. We found reduced numbers of recruited neutrophils and macrophages at the wounding area in both *tlr2* mutants and *myd88* mutants, compared to their sibling controls. Leukocyte migration of the *tlr2* and *myd88* mutations upon wounding was analyzed using quantitative analyses of cell migration tracks. Our results demonstrate that the *tlr2* and the *myd88* mutations affect distant neutrophil migration upon wounding by negatively affecting their directional persistence, but not their migration speed. Not only the directional persistence of distant macrophage was significantly decreased in the *tlr2* and the *myd88* mutants, but also their migration speed. This study shows for the first time that TLR signaling is directly involved in controlling behavior of cell migration of neutrophils and macrophages during wounding, stimulating further studies also in other model systems.

5.2 Materials and methods

5.2.1 Zebrafish maintenance and strain construction

All animal experiments described in this study were performed at the University of Leiden according to standard protocols (zfin.org) and adhered to the international guidelines specified by the EU Animal Protection Directive 2010/63/EU. The culture of adult fish was approved by the local animal welfare committee (DEC) of the university (License number: protocol 14,198). No adult zebrafish were sacrificed for this study. All experiments were done on 3 days post fertilization (dpf) fish, therefore prior to the free-feeding stage and did not fall under animal experimentation law according to the EU Animal Protection Directive 2010/63/EU. Eggs

and larvae were grown at 28.5°C in egg water (60 g/ml Instant Ocean sea salts). For living imaging and tail wounding experiments, 3 dpf larvae were anesthetized with egg water containing 0.02% buffered 3-aminobenzoic acid ethyl ester (Tricaine, Sigma-Aldrich, the Netherlands).

The *tlr2*^{sa19423} mutant and *myd88*^{hu3568} mutant lines were identified by the sequencing of an ENU-mutagenized zebrafish library [257, 258]. Both homozygous mutants were outcrossed with the double transgenic line *Tg (mpeg1:mCherry-F);TgBAC (mpx: EGFP)* [250, 261]. Subsequently, their heterozygous offspring with both positive GFP and mCherry fluorescence were incrossed to produce the homozygous mutants and wild type siblings. In the present study, the double transgenic lines were used for the quantification of cell numbers, recruitment assay upon wounding and leukocyte living imaging experiment.

To investigate the effect of the *tlr2* and the *myd88* mutations on leukocyte development, double fluorescent lines *tlr2*$^{+/+}$ *Tg (mpeg1:mCherry-F);TgBAC (mpx: EGFP)*, *tlr2*$^{-/-}$ *Tg (mpeg1:mCherry-F);TgBAC (mpx: EGFP)*, *myd88*$^{+/+}$ *Tg (mpeg1:mCherry-F);TgBAC (mpx: EGFP)*, *myd88*$^{-/-}$ *Tg (mpeg1:mCherry-F);TgBAC (mpx: EGFP)* were used.

5.2.2 Tail wounding

In the present study, a caudal fin wounding model was applied as previously described [250, 262, 252]. 3 dpf *tlr2* zebrafish larvae were anesthetized with egg water containing 0.02% tricaine (Sigma Aldrich). Subsequently, the caudal fins of larvae were wounded by using a 1 mm sterile sapphire blade scalpel (World Precision Instruments) on a 2% agarose covered petridish. To avoid damaging the notochord and other tissues of zebrafish larvae, all of the wounding experiments were performed under a MZ16FA Fluorescence Stereo Microscope (Leica Microsystems, Wetzlar Germany) equipped with a DFC420C color camera (Leica Microsystems). After the wounding, the egg water with 0.02% tricaine was changed with untreated egg water. Wounded larvae were put back into an incubator at 28.5°C. Subsequently, the wounded larvae were collected or fixed for follow up experiments.

5.2.3 Imaging and quantification

For the quantification of the recruited cell number upon wounding, the double transgenic *tlr2* and *myd88* larvae were wounded with the method described before. 1, 2, 4 and 6 hour post wounding (hpw), larvae were collected and fixed with 4% paraformaldehyde (PFA) in PBS overnight

at 4°C and washed with PBS the next day. The wounded tail area of fixed samples from each group were imaged by using a Leica MZ16FA fluorescence stereo microscope equipped with a DFC420C color camera. Cells localized within an area of 200 µm from the wounding edge toward the body trunk were counted as recruited cells. Analysis was performed by combining three independent experiments.

For detailed cell migration behavior analyses, larvae (3 dpf) were mounted into 1% low melting point agarose (Sigma Aldrich) with 0.02% tricaine and imaged under a Leica TCS SP8 confocal microscope (Leica Microsystems) with a 10× objective (N.A. 0.40). Data were saved as maximum projection images for further cell counting. The number of neutrophils and macrophages in the tail region were manually quantified.

5.2.4 Live imaging

All time-lapse imaging was performed on 3 dpf larvae. Larvae for each condition (unchallenged/ wounded) were mounted in the method described before and visualized in the CLSM with 1 min time interval for 2 h image capture using a 20× objective (N.A. 0.75). For the cell tracking analysis, all time-lapse images were saved as maximum projection images.

We first defined the role of *tlr2* and *myd88* in leukocyte migration under the unchallenged condition. The caudal hematopoietic tissue (CHT) of double transgenic lines was imaged using the CLSM with unchallenged condition. To investigate the effect of the *tlr2* and *myd88* mutations on leukocyte migration upon wounding, the double transgenic line *Tg (mpeg1:mCherry-F);TgBAC (mpx: EGFP)* larvae in the *tlr2*, *myd88* mutant or their wild type background were wounded and performed for real time imaging from 1 hpw to 3 hpw.

5.2.5 Cell tracking and its quantification

The cell tracking of macrophages and neutrophils was performed manually by using a manual tracking plug-in from Fiji [263, 264]. To analyze the behavior of leukocyte tracks more accurately, we defined categories of distant and local resident cells. Cells localized further than 200 µm from the cut edge towards the trunk were categorized as distant cells. Cell localized within a distance of up to 200 µm from the cut toward to the trunk were categorized as local resident cells.

The distance to the wound, mean speed, net displacement, meandering index (M.I.), mean square displacement (MSD), cell diffusivity (D), velocity in anteroposterior direction (V_{AP}) and V_{AP} over time were calculated in

different groups by manual tracking data. The calculation and explanation of the parameters are shown in Figure 5.1. The distance to the wound is defined as the shortest Euclidean distance to the manually traced wound edge (Figure 5.1A). For the velocity in the anteroposterior direction, tracks were rotated such that the spines of the larvae were aligned (Figure 5.1B). Then, for each cell the average velocity in the anteroposterior axis was calculated. For V_{AP} over time, the V_{AP} of all cells within a group was averaged over three consecutive time frames. Net displacement, total displacement, meandering index and mean speed are shown in Figure 5.1C and Table 5.1 (Eqs. 5.1-5.4). The net displacement is the distance of the cell between the first and final time frame (Figure 5.1C), i.e., the Euclidian distance traveled being: $d_{\text{net}} = d\left(p_i, p_N\right)$ (Table 5.1. Eq. 5.1). The total displacement is the length of the total cell track, i.e., the sum of the net displacements between two successive frames ($d_{\text{tot}} = \sum_{i=1}^{N-1} d\left(p_i, p_{i+1}\right)$) (Figure 5.1C) (Table 5.1. Eq. 5.2). Cells can reorient between two frames, such that this measure may underestimate the actual distance traveled. However, we used the same frame rate of 1 min in all experiments, such that the results are comparable with one another. Meandering index is most simply defined as the net distance traveled divided by the total distance traveled (M.I $=\frac{d_{\text{net}}}{d_{\text{tot}}}$) [265] (Figure 5.1C) (Table 5.1. Eq. 5.3). Mean speed is the total displacement divided by traveled time ($\overline{v} = \frac{1}{N-1}\sum_{i=1}^{N-1} v_i$) (Table 5.1. Eq. 5.4). The MSD at time t was calculated for each group by averaging the squared displacement from starting time t_1=1hpw to time t over all cells (K) within that group (MSD $(t) = \frac{1}{K}\sum_{i=1}^{K} d\left(p_{i,1}, p_{i,1+t}\right)^2$) (Figure 5.1D) (Table 5.1. Eq. 5.5). The MSD curve can be used to distinguish migratory behaviours such as Brownian walks and persistent random walks, and analytical expressions for the MSD are known for theoretically ideal random walks. For persistent random walkers, the analytical expression for the MSD is: $MSD\ (t) = 2v^2\ \tau t - 2(v\tau)^2\ (1 - e^{\frac{-t}{\tau}}\ )$) (Table 5.1. Eq. 5.6), with v the intrinsic cell velocity and τ the persistent time, which can be fitted to the MSD calculated from cell tracks [158]. The cell diffusivity constant D and MSD (t) at large t are related through $D = 1/2n\ \frac{dMSD(t)}{\text{dt}}$, with $n = 2$ the dimension, which for persistent random walkers results in $D = 1/2\ v^2\ \tau$ (Table 5.1. Eq. 5.7). Regarding the shape of the MSD curve of our data and the persistent motion reported in leukocyte migration upon tail wounding [266], we assume that distant neutrophils and macrophages behave like persistent random walkers during the time span of imaging. We fit Eq. 5.6 to the MSD curve (Table 5.1. Eq. 5.5) using a non-linear least squares method. The obtained parameters v and τ are then used to compute the approximated cell diffusivity D. For distant neutrophils, the

TABLE 5.1: Formulas of calculated track measures and derived measures

Measure	Definition	
Net displacement (µm)	$d_{\text{net}} = d(p_i, p_N)$	(5.1)
Total displacement (µm)	$d_{\text{tot}} = \sum_{i=1}^{N-1} d(p_i, p_{i+1})$	(5.2)
Meandering index	$M.I. = d_{\text{net}}/d_{\text{tot}}$	(5.3)
Mean speed (µm/min)	$\overline{v} = \frac{1}{N-1} \sum_{i=1}^{N-1} v_i$	(5.4)
Mean squared displacement (µm^2)	$\text{MSD}\ (t) = \frac{1}{K} \sum_{i=1}^{K} (d\ (p_{i,1}\ ,\ p_{i,\ 1+t}))^2$	(5.5)
Fitted mean squared displacement (µm^2)	$\text{MSD}\,(t) = 2v^2\ \tau t - 2\,(v\tau)^2\ \left(1 - e^{\frac{-t}{\tau}}\right)$	(5.6)
Cell diffusivity constant (µm^2/min)	$D = 1/2\ v^2\ \tau$	(5.7)

fit was performed on the first 80 min of tracking, for distant macrophages, the entire 2h tracking period was used.

5.2.6 Statistical analysis

Graphpad Prism software (Version 8.1.1; GraphPad Software, San Diego, CA, USA) was used for statistical analysis. Computations of distance to the wound, MSD and V_{AP} were performed using a Python script including the SciPy stats library for statistical testing. Shaded regions of MSD and V_{AP} over time indicate standard error of mean, the other experiment data are shown as mean ± SD. Statistical significance of differences was determined by using an unpaired, two-tailed t-test for comparing the difference between wild type and *tlr2* and *myd88* mutant. (ns, no significant difference; *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001).

5.3 Results

5.3.1 Tlr2 and myd88 mutations do not affect development and basal motility of leukocytes.

To determine the leukocyte development in *tlr2* and *myd88* mutants, the double-transgenic line *tlr2*$^{+/+}$ *Tg (mpeg1:mCherry-F);TgBAC (mpx: EGFP)*,

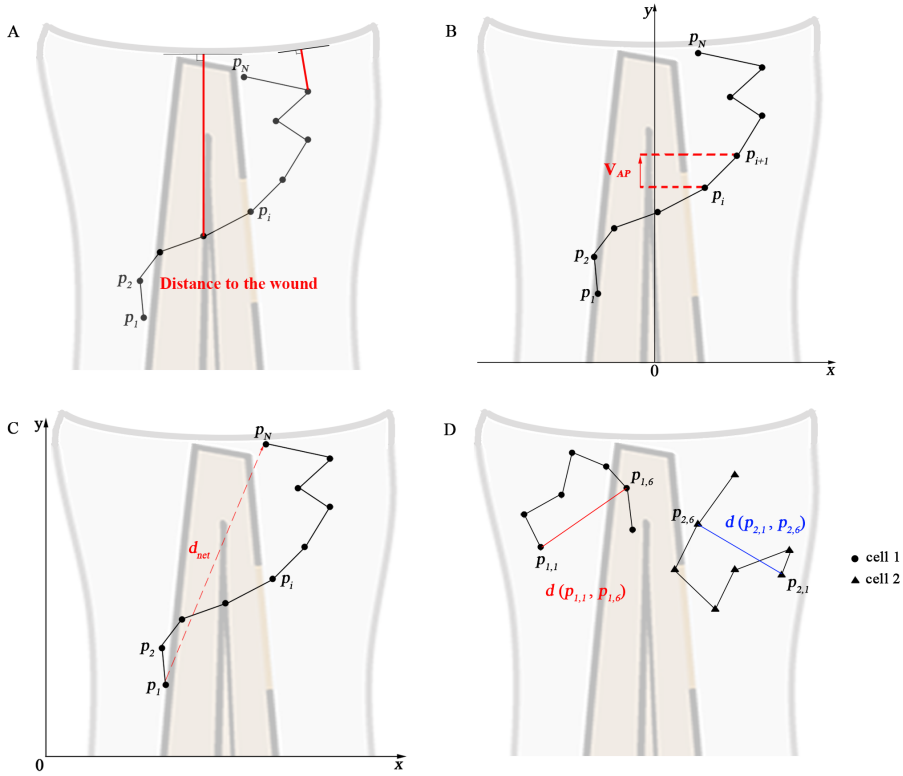
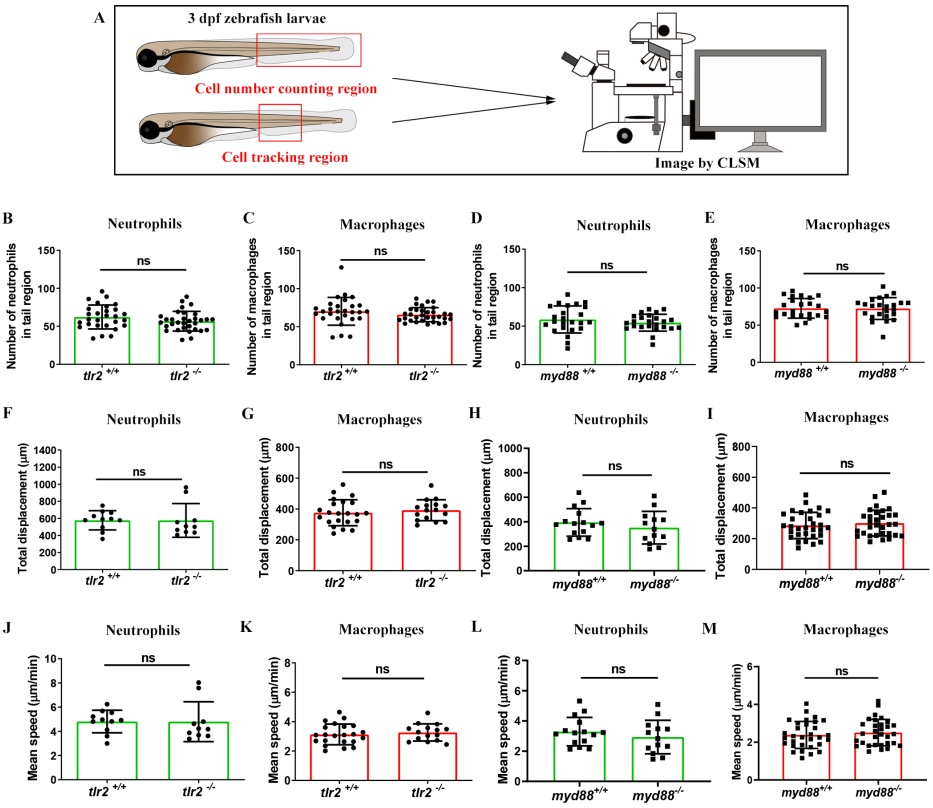


FIGURE 5.1: **Calculated track measures. (A)** Depiction of distance to the wound. It measured for each frame as the shortest distance from the cell's current position to the entire line of the wound, i.e. the orthogonal projection to the wound. **(B)** Depiction of V_{AP}: velocity in anteroposterior axis direction. The visible part of the spine is taken as the y-axis. **(C)** Depiction of the net displacement, total displacement, meandering index and mean speed: the net displacement is the distance of the cell between the first and final time frame. Total displacement is the sum of the net displacement between 2 successive frames. Meandering index corresponds to the net displacement divided by the total displacement and measures the trajectory straightness. Mean speed is the total displacement divided by traveled time. Formulas are shown in Table 5.1. (Eqs. 5.1-5.4). **(D)** Depiction of the construction of the mean squared displacement: the displacement between the first time frame and time frame t from all cells is squared and averaged, see Table 5.1. (Eq. 5.5).

tlr2$^{-/-}$ *Tg (mpeg1:mCherry-F);TgBAC (mpx: EGFP), myd88*$^{+/+}$ *Tg (mpeg1:mCherry-F);TgBAC (mpx: EGFP)* and *myd88*$^{-/-}$ *Tg (mpeg1:mCherry-F);TgBAC (mpx: EGFP)* were constructed. The lines were imaged at 3 dpf to count the number of macrophages and neutrophils in their tail region, and then compared with their wild type siblings (Figure 5.2A). Embryos of the *tlr2* and *myd88* mutants showed similar numbers of macrophages and neutrophils as their wild type siblings (Figures 5.2B- E). This result is consistent with our previous studies of the same *myd88* mutant at 3 dpf and the *tlr2* mutant at 2 dpf [257, 258]. With the aim of investigating the importance of the *tlr2* and the *myd88* mutations for leukocyte behavior under unchallenged condition, the CHT region was analyzed in the double transgenic lines of *tlr2* and *myd88*

FIGURE 5.2 *(following page)*: **Quantification of macrophage and neutrophil numbers and their basal migratory capability in the 3 dpf *tlr2 and myd88* mutants and wild sibling controls larvae. (A)** Experimental scheme. At 3 dpf, numbers and basal migratory capability of GFP-labeled neutrophils and mCherry-labeled macrophages in tail region were quantified using Leica TCS SP8 confocal laser scanning microscopy (CLSM). Red boxes show the area in which cells were counted or tracked. **(B- E)** The quantification of neutrophil and macrophage numbers in tail region by using *tlr2* and *myd88* zebrafish larvae. Data (mean ± SD) are combined from three pools of zebrafish larvae. No significant differences (ns) in the number of neutrophils (B, D) and macrophages (C, E) was detected with an unpaired, two-tailed t-test. Sample size (n): 28, 32 (B, C); 24, 24 (D, E). **(F- G, J- K)** Quantification of basal migratory capability in 3 dpf *tlr2* zebrafish. The total displacement and mean speed of individual neutrophils (F, J) and macrophages (G, K) were quantified by using a manual tracking plugin. Data (mean ± SD) are combined from three experiments of *tlr2*$^{+/+}$ *Tg (mpeg1:mCherry-F);TgBAC (mpx: EGFP)* and *tlr2*$^{-/-}$ *Tg (mpeg1:mCherry-F);TgBAC (mpx: EGFP)* larvae. No significant differences (ns) in the total displacement and mean speed of neutrophils (F, J) and macrophages (G, K) were detected with an unpaired, two-tailed t-test. Sample size (n): 11, 10 (F, J); 22, 16 (G, K). Cell tracking movies are shown in Supplementary Movie S1-4) **(H- I, L- M)** Quantification of basal migratory capability in 3 dpf *myd88* zebrafish. The total displacement and mean speed of individual neutrophils (H, L) and macrophages (I, M) were quantified by using a manual tracking plugin. Data (mean ± SD) are combined from three experiments of *myd88*$^{+/+}$ *Tg (mpeg1:mCherry-F);TgBAC (mpx: EGFP)* and *myd88*$^{-/-}$ *Tg (mpeg1:mCherry-F);TgBAC (mpx: EGFP)* larvae. No significant differences (ns) in the total displacement and mean speed of neutrophils (H, L) and macrophages (I, M) were detected with an unpaired, two-tailed t-test. Sample size (n): 15, 13 (H, L); 31, 31 (I, M). Cell tracking movies are shown in Supplementary Movie S5-8)

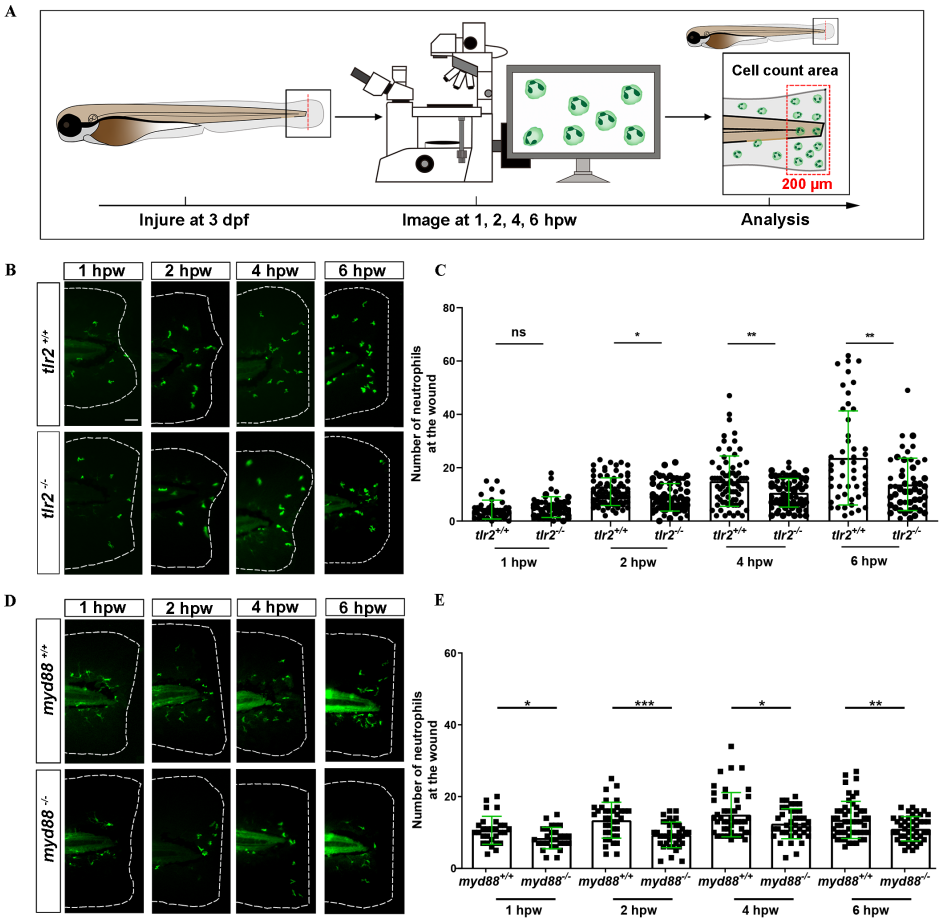

A
3 dpf zebrafish larvae
Cell number counting region
Cell tracking region
Image by CLSM
B
Neutrophils
Number of neutrophils in tail region
ns
tlr2 +/+
tlr2 -/-
C
Macrophages
Number of macrophages in tail region
D
Neutrophils
myd88 +/+
myd88 -/-
E
Macrophages
F
Neutrophils
Total displacement (μm)
G
Macrophages
H
Neutrophils
I
Macrophages
J
Neutrophils
Mean speed (μm/min)
K
Macrophages
L
Neutrophils
M
Macrophages

using CLSM by taking time-lapse images (Figure 5.2A). No significant effect was observed on leukocyte basal motility in the CHT tissue in the *tlr2* and *myd88* mutants compared with their wild type sibling control (Figures 5.2F-M). Representative images are shown in Supplementary Figure S5.1 and Supplementary Figure S5.2.

5.3.2 Tlr2 and myd88 regulate neutrophil recruitment to a tail wound

To study the effect of the *tlr2* and *myd88* mutations on the recruitment of neutrophils towards a site of injury, a tail wound method was used in 3 dpf zebrafish larvae as a model for inflammation. To quantify the number of recruited neutrophils to the wound, we counted the number of neutrophils that were located in a range closer than 200 µm from the wound edge of the tail at 1, 2, 4 and 6 hpw (Figure 5.3A). Our results show that the *tlr2* mutation had a significant negative effect on the recruitment of neutrophils after 2, 4 and 6 hpw (Figures 5.3B,C). However, there is no significant difference in recruited neutrophil numbers between wild type and *tlr2*$^{-/-}$ at 1 hpw (Figures 5.3B,C). Notably, a significant difference of recruited neutrophil numbers was already observed at 1 hpw in *myd88* zebrafish larvae and remained significant until 6 hpw (Figure 5.3D,E).

FIGURE 5.3 *(following page)*: **The number of neutrophils recruited to the wounded area in the *tlr2* and *myd88* mutants and wild type sibling controls larvae. (A)** Experimental scheme. *Tlr2* and *myd88* homozygous mutants and sibling control larvae were wounded at 3 dpf. Their tails were wounded to the tip of the notochord. The red dashed line shows the site of wounding. Recruited neutrophils at the wound were imaged at 1, 2, 4 and 6 hpw by using CLSM. For recruited cell counting analysis, cells localized within an area of 200 µm from the wounding edge toward the body trunk were counted as recruited cells. The red dashed box shows the area where neutrophils were counted as recruited neutrophils. **(B, D)** Representative images of 3 days dpf larvae at 1, 2, 4 and 6 hours post-wounding (hpw). Scale bar: 50 µm. **(C)** Quantification of recruited neutrophil numbers to the wounded area at 1, 2, 4 and 6 hpw in 3dpf *tlr2*$^{+/+}$ and *tlr2*$^{-/-}$ larvae. Sample size (n): 45, 46, 82, 72, 74, 68, 50, 50. **(E)** Quantification of recruited neutrophil numbers to the wounded area at 1, 2, 4 and 6 hpw in 3dpf *myd88*$^{+/+}$ and *myd88*$^{-/-}$ larvae. Sample size (n): 29, 28, 37, 38, 45, 39, 51, 45. In all cases, statistical analyses were done from 3 independent experiments. An unpaired, two-tailed t-test was used to assess significance (ns, no significant difference, *P < 0.05, **P < 0.01, ***P < 0.001) and data are shown as mean± SD.


A
Cell count area
200 µm
Injure at 3 dpf
Image at 1, 2, 4, 6 hpw
Analysis
B
1 hpw
2 hpw
4 hpw
6 hpw
tlr2 +/+
tlr2 -/-
C
Number of neutrophils at the wound
ns
*
**
**
tlr2+/+
tlr2-/-
1 hpw
2 hpw
4 hpw
6 hpw
D
myd88 +/+
myd88 -/-
E
*
***
*
**
myd88+/+
myd88-/-

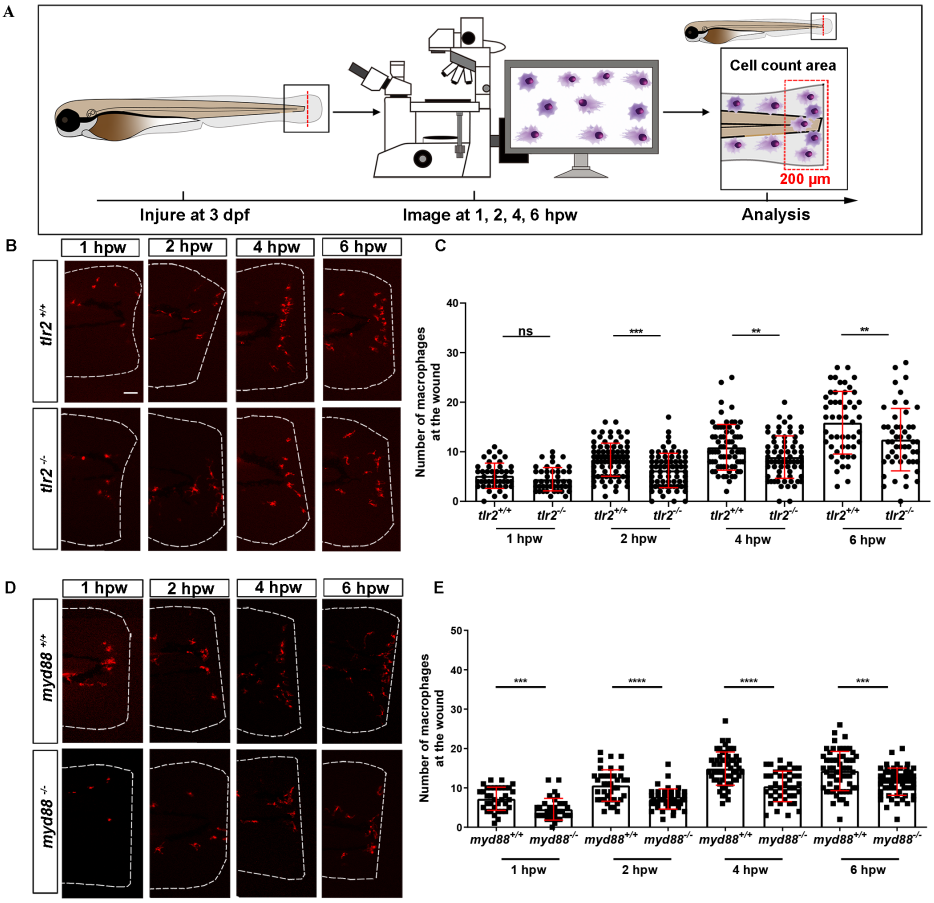
5.3.3 Tlr2 and myd88 regulate macrophage recruitment to a tail wound

To assess the role of the *tlr2* and *myd88* mutations in regulating the recruitment of macrophages to a site of the tail wound, we counted the recruited macrophage numbers by the same method as for measuring the neutrophil recruitment to the wound (Figure 5.4A). Both *tlr2*$^{-/-}$ and *myd88*$^{-/-}$ mutant zebrafish larvae displayed diminished macrophage responses upon wounding (Figure 5.4). Significantly decreased numbers of recruited macrophages toward the injury were measured in the *tlr2*$^{-/-}$ group at 2, 4 and 6 hpw (Figures 5.4B,C). Similarly, there is no significant difference in recruited macrophage numbers between wild type and *tlr2*$^{-/-}$ at 1 hpw (Figure 5.4C). A significant difference of recruited macrophage numbers was already observed from 1 hpw in *myd88* zebrafish larvae, the same as was observed with neutrophil recruitment (Figures 5.4D,E).

5.3.4 Live imaging reveals that the *tlr2* and *myd88* mutations affect distant neutrophil directional persistence, but not migration speed upon tail wounding

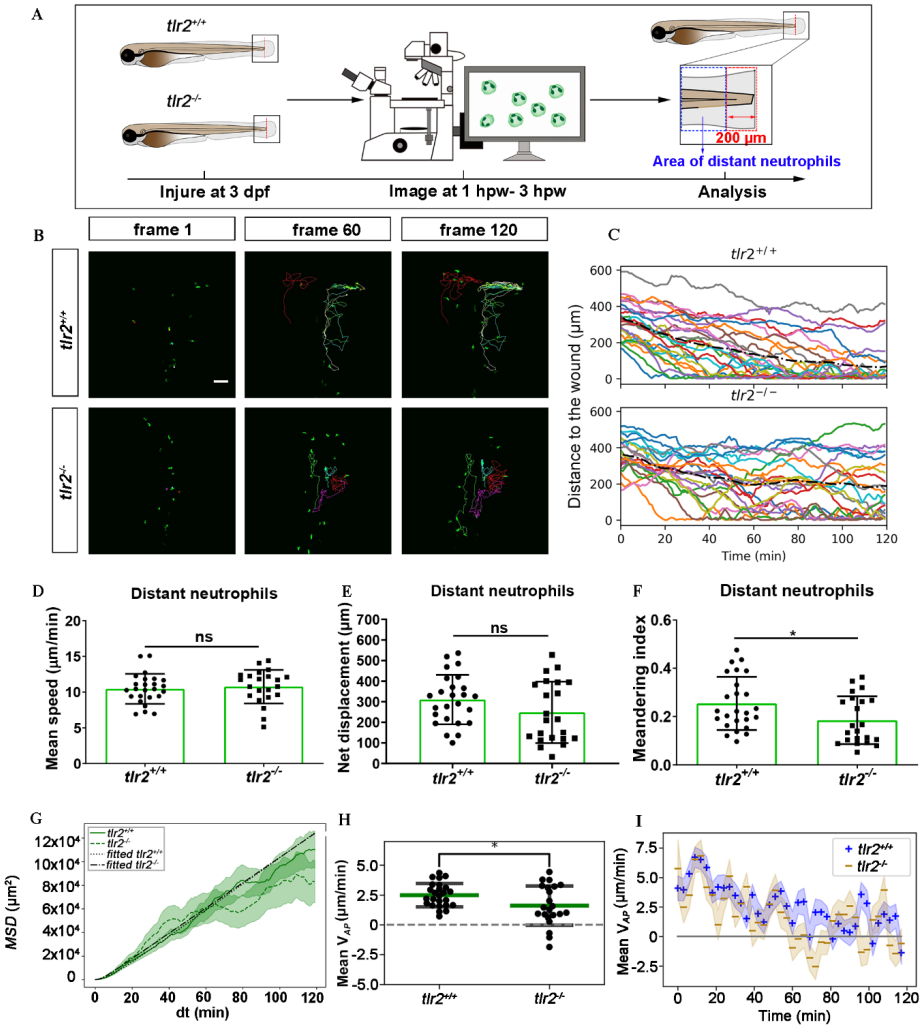
To investigate how neutrophils migrate in the absence of *tlr2* or *myd88* after tail wounding, a time-lapse microscopy experiment was performed

FIGURE 5.4 *(following page)*: **The number of macrophages recruited to the wounded area in the *tlr2* and *myd88* mutants and wild type sibling controls larvae. (A)** Experimental scheme. *Tlr2* and *myd88* homozygous mutants and sibling control larvae were wounded at 3 dpf. Their tails were wounded to the tip of the notochord. The red dashed line shows the site of wounding. Recruited macrophages at the wound were imaged at 1, 2, 4 and 6 hpw by using CLSM. For recruited cell counting analysis, cells localized within an area of 200 µm from the wounding edge toward the body trunk were counted as recruited cells. The red dashed box shows the area where macrophages were counted as recruited macrophages. **(B, D)** Representative images of 3 days dpf larvae at 1, 2, 4 and 6 hpw. Scale bar: 50 µm. **(C)** The quantification of recruited macrophage numbers to the wounded area at 1, 2, 4 and 6 hpw in 3dpf *tlr2*$^{+/+}$ and *tlr2*$^{-/-}$ larvae. Sample size (n): 45, 45, 82, 71, 69, 68, 51, 50. **(E)** The quantification of recruited macrophage numbers to the wounded area at 1, 2, 4 and 6 hpw in 3dpf *myd88*$^{+/+}$ and *myd88*$^{-/-}$ larvae. Sample size (n): 35, 34, 40, 43, 56, 42, 60, 58. In all cases, statistical analyses were done with data of 3 independent experiments. An unpaired, two-tailed t-test was used to assess significance (ns, no significant difference, **P < 0.01, ***P < 0.001, ****P < 0.0001) and data are shown as mean± SD.


A
Cell count area
200 µm
Injure at 3 dpf
Image at 1, 2, 4, 6 hpw
Analysis
B
1 hpw
2 hpw
4 hpw
6 hpw
tlr2 +/+
tlr2 -/-
C
Number of macrophages at the wound
ns
***
**
**
tlr2+/+
tlr2-/-
1 hpw
2 hpw
4 hpw
6 hpw
D
myd88 +/+
myd88 -/-
E
Number of macrophages at the wound
***
****
****
***
myd88+/+
myd88-/-

by using CLSM between 1 hpw to 3 hpw (Figures 5.5,5.6). The definition of distant and local resident neutrophils was shown in panel A of Figures 5.5,5.6 and Supplementary Figures 5.7,5.8. Neutrophils located closer than 200µm to the wound were defined as local resident neutrophils and further than 200µm were defined as distant neutrophils. Measurement of the distance to the wound over time of all distant neutrophils in the *tlr2*$^{-/-}$ group indicated a trend of impaired infiltration towards the wound (Figures 5.5B,C up panel). In total, the group of distant neutrophils in the *tlr2*$^{+/+}$ group that arrived at the wound edge and stayed within a distance of 20 µm to the wound comprises 84 % of a total of 25 tracked neutrophils (Figure 5.5C up panel). The local resident neutrophils in this group all remained at the wound (Figures 5.7B,C up panel). In contrast, the group of the distant neutrophils in the *tlr2*$^{-/-}$ group that arrived at the wound within 2 h time-lapse cell tracking comprises only approximately 36 % (Figures 5.5B,C bottom panel). Moreover, approximately 33 % of local resident neutrophils in the *tlr2*$^{-/-}$ group already migrated away from the wound edge within 3 hpw (Figures 5.7B,C bottom panel).

FIGURE 5.5 *(following page)*: **Quantification of distant neutrophils behavior in wounded *tlr2* mutant and sibling control larvae. (A)** Experimental scheme. *Tlr2*$^{+/+}$ and *tlr2*$^{-/-}$ larvae were wounded at 3 dpf. The red dashed line shows the site of wounding. Neutrophils of wounded zebrafish larvae were tracked for 2 h and images were taken every 1 min by using CLSM. For cell tracking analysis, cells localized outside an area of 200 µm from the wounding edge toward the body trunk were counted as distant cells. Blue dashed box shows the area where distant neutrophils were tracked. **(B)** Representative images of distant neutrophil tracks in the wounded tail fin of 3 dpf *tlr2*$^{+/+}$ or *tlr2*$^{-/-}$ larvae at frame 1, frame 60 and frame 120. Time interval between two successive frames is 1 min. Each color track represents an individual neutrophil. Cell tracking movies are shown in Supplementary Movie S9-10). Scale bar: 50 µm. **(C)** Distance to the wound. Black dash line represents average distance to the wound. Each color line represents one cell. **(D-I)** Quantification of distant neutrophil tracks. There was no significant difference between the groups in terms of mean speed (D), net displacement (E) and MSD (green) and fitted MSD (black) (G). However, meandering index (F) and mean V_{AP} (H) of neutrophils at the wound in *tlr2*$^{+/+}$ is greater than in *tlr2*$^{-/-}$ larvae. The fitted MSD (G, black) was fitted for $dt < 80$ min. The shaded regions in MSD (G) and mean V_{AP} over time (I) indicate standard error of the mean. Statistical analyses were done with 7 and 8 fish respectively for each group. An unpaired, two-tailed t-test was used to assess significance (ns, non-significance, *P < 0.05) and data are shown as mean± SD. Sample size (n): 25, 22 (D, E, F, H).


A
tlr2+/+
tlr2-/-
200 µm
Area of distant neutrophils
Injure at 3 dpf
Image at 1 hpw- 3 hpw
Analysis
B
frame 1
frame 60
frame 120
tlr2+/+
tlr2-/-
C
tlr2+/+
tlr2-/-
Distance to the wound (µm)
Time (min)
D
Distant neutrophils
Mean speed (µm/min)
ns
E
Distant neutrophils
Net displacement (µm)
ns
F
Distant neutrophils
Meandering index
*
G
MSD (µm²)
dt (min)
H
Mean VAP (µm/min)
*
I
Mean VAP (µm/min)
Time (min)

In general, distant neutrophils in the *myd88*$^{+/+}$ group showed more chemotaxis to the wound compared to *myd88*$^{-/-}$ neutrophils (Figures 5.6B,C). Approximately 96% distant neutrophils arrived at the wound (within a distance of 20 µm to the wound) in the *myd88*$^{+/+}$ group in total (Figure 5.6C up panel). However, only 70% distant neutrophils arrived to the wound (within a distance of 20 µm to the wound) in the *myd88*$^{-/-}$ group. (Figure 5.6C bottom panel). The local resident neutrophils in this group all remained at the wound except for a few outliers (Figures 5.8C). In summary, the general trend of distant neutrophils migration in the *myd88* mutant and sibling zebrafish groups was consistent with the result in the *tlr2* mutant and sibling zebrafish groups, respectively (Figure 5.6C).

To quantify differences in neutrophil migration behavior between *tlr2* and *myd88* mutants and their wild type siblings, we first analyzed whether the deficiency of *tlr2* and *myd88* can affect neutrophil mean migration speed upon wounding. The results showed that the *tlr2* and the *myd88* mutations do not affect the mean speed of both distant and local resident neutrophils

FIGURE 5.6 *(following page)*: **Quantification of distant neutrophils behavior in wounded *myd88* mutant and sibling control larvae. (A)** Experimental scheme. *Myd88*$^{+/+}$ and *myd88*$^{-/-}$ larvae were wounded at 3 dpf. The red dashed line shows the site of wounding. Neutrophils of wounded *myd88* zebrafish larvae were tracked for 2 h and images were taken every 1 min by using CLSM. For cell tracking analysis, cells localized outside an area of 200 µm from the wounding edge toward the body trunk were counted as distant cells. Blue dashed box shows the area where distant neutrophils were tracked. **(B)** Representative images of distant neutrophil tracks in the wounded tail fin of 3 dpf *myd88*$^{+/+}$ or *myd88*$^{-/-}$ larvae at frame 1, frame 60 and frame 120. Time interval between two successive frames is 1 min. Each color track represents an individual neutrophil. Cell tracking movies are shown in Supplementary Movie S11-12). Scale bar: 50 µm. **(C)** Distance to the wound. Black dash line represents average distance to the wound. Each color line represents one cell. **(D-I)** Quantification of distant neutrophil tracks. There was no significant difference between the groups in terms of mean speed (D). However, the net displacement (E), meandering index (F),), MSD (green) and fitted MSD (black) (G) and mean V_{AP} (H) of neutrophils at the wound in *myd88*$^{+/+}$ is greater than in *myd88*$^{-/-}$ larvae. The shaded regions MSD (G) and in mean V_{AP} over time (I) indicate standard error of the mean. The fitted MSD (G, black) was fitted for $dt < 80$ min. Statistical analyses were done with 6 and 4 fish respectively for each group. An unpaired, two-tailed t-test was used to assess significance (ns, non-significance, **P < 0.01) and data are shown as mean± SD. Sample size (n): 23, 10 (D, E, F, H).

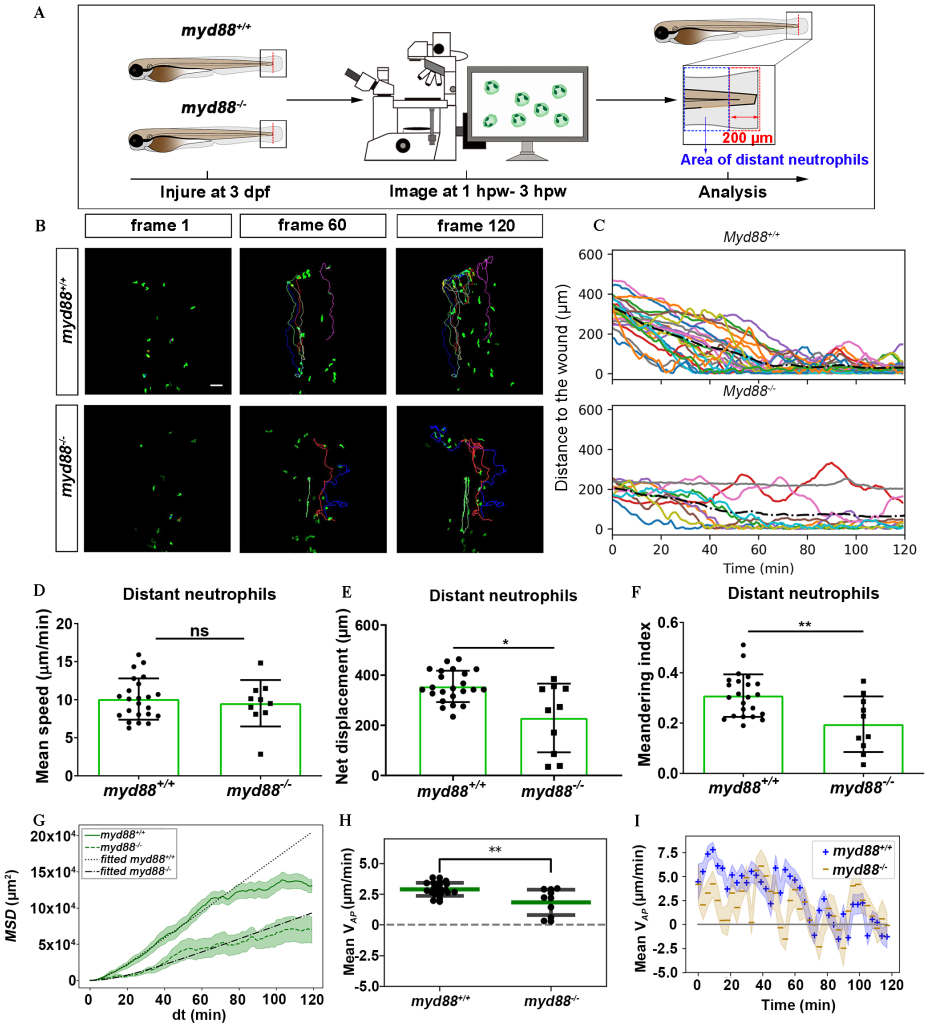

A
myd88+/+
myd88-/-
200 µm
Area of distant neutrophils
Injure at 3 dpf
Image at 1 hpw- 3 hpw
Analysis
B
frame 1
frame 60
frame 120
myd88+/+
myd88-/-
C
Myd88+/+
Myd88-/-
Distance to the wound (µm)
Time (min)
D
Distant neutrophils
ns
Mean speed (µm/min)
E
Distant neutrophils
*
Net displacement (µm)
F
Distant neutrophils
**
Meandering index
G
MSD (µm²)
dt (min)
H
**
Mean VAP (µm/min)
I
Mean VAP (µm/min)
Time (min)

TABLE 5.2: Persistence time, intrinsic cell velocity and diffusivity obtained by fitting Eq. 5.6 to the MSD curves.

Neutrophils	τ (min)	v (μm/min)	D (μm^2/min)	Fitted time-frame
tlr2$^{+/+}$	6.88 ± 0.23	8.98 ± 0.12	277	$t \leq 80$
tlr2$^{-/-}$	3.24 ± 0.35	12.80 ± 0.59	265	
myd88$^{+/+}$	6.83 ± 0.34	10.97 ± 0.22	410	
myd88$^{-/-}$	5.04 ± 0.49	10.33 ± 0.42	269	

upon the wounding (Figure 5.5D; Figures 5.6D; Supplementary Figures 5.7D and Supplementary Figures 5.8D).

We also tested the effect of the *tlr2* and the *myd88* mutations on the movement direction of neutrophils upon wounding by the quantification of net displacement, whose definition is shown in Figure 5.1 and Table 5.1. We observed that the net displacement of distant neutrophils had a decreased trend in the *tlr2*$^{-/-}$ group compared to the *tlr2*$^{+/+}$ group (Figure 5.5E). Moreover, cell diffusivity D determined by fitting Eq. 5.6 to the MSD curve (Table 5.1) did not differ much between the *tlr2*$^{-/-}$ group and the *tlr2*$^{+/+}$ group (Figure 5.5G, Table 5.2). A significant decrease in net displacements was consistently observed in the *myd88* mutant group (Figure 5.6E). Also, *myd88*$^{-/-}$ neutrophils have lower diffusivity than *myd88*$^{+/+}$ neutrophils as measured from fitting Eq. 5.6 to the MSD plots (Figure 5.6G, Table 5.2). As the cell speed of *myd88*$^{-/-}$ neutrophils does not differ from that of *myd88*$^{+/+}$ neutrophils (Figure 5.6D), the reduced diffusivity may be due to more frequent or sharper changes of direction of the *myd88*$^{-/-}$ neutrophils. Preliminary results in analysing the angles between consecutive time frames suggest that the *myd88*$^{-/-}$ neutrophils indeed have sharper turns than their wild type siblings. As neutrophils reach the wound edge, their diffusivity is limited in space. With the initial distance of the distant neutrophils in min, the maximum of MSD(t) is expected around 12×10^4μm^2/min, which is also visible in the flattening of the MSD at later time frames. We therefore limited fitting Eq. 5.6 to the MSD curve to $dt < 80$.

To further study the effect of the *tlr2* and *myd88* mutations on the neutrophil migration direction, we determined the meandering index and mean V_{AP} (Figures 5.5F,H and Figures 5.6F,H). The meandering index and mean V_{AP} are all significantly decreased in the distant neutrophils of both *tlr2*$^{-/-}$ and *myd88*$^{-/-}$ mutants compared to their wild type sibling controls (Figures 5.5F,H and Figures 5.6F,H). However, no significant difference of meandering index was found in local resident neutrophils of the *tlr2*$^{-/-}$

and *myd88*$^{-/-}$ mutants compared to the wild type siblings (Figures 5.7F and 5.8F). The mean V_{AP} over time qualitatively shows again the impaired chemotaxis of *tlr2*$^{-/-}$ and *myd88*$^{-/-}$ neutrophils compared to the *tlr2*$^{+/+}$ and *myd88*$^{+/+}$ neutrophils, respectively (Figure 5.5I and Figure 5.6I). As more and more neutrophils approach the wound (Figure 5.5C,5.6), the mean V_{AP} drops. For almost every time point, mean V_{AP} of *tlr2*$^{+/+}$ exceeds mean V_{AP} of *tlr2*$^{-/-}$(Figure 5.5I). Similar results were observed for *myd88*$^{+/+}$ and *myd88*$^{-/-}$ distant neutrophils (Figure 5.6I).

5.3.5 Live imaging reveals that the *tlr2* and *myd88* mutations affect distant macrophage migration speed and directional persistence upon tail wounding

To study the effect of the *tlr2* and *myd88* mutations on macrophage migration upon wounding, we compared macrophage behavior with their wild type siblings. The definition of distant macrophage and local resident macrophage was shown in panel A of Figures 5.9 , 5.10, 5.11, and 5.12. Macrophages located closer than 200µm to the wound were defined as local resident macrophages and further than 200µm were defined as distant macrophages. In contrast to neutrophils, the majority of macrophages do not reach the wound within the measured time period. By measuring their distance to the wound over time, we can see a trend that distant macrophages show less chemotaxis in the *tlr2*$^{-/-}$ and *myd88*$^{-/-}$ mutant groups compared to their wild type sibling groups (Figures 5.9B,C and Figures 5.10B,C). Within 50 µm to the wound, the local resident macrophages all remained at the wound in both the *tlr2* and *myd88* mutants and their wild type sibling controls (Figures 5.11B,C and Figures 5.12B,C). Within a distance of 200 µm, but outside 50 µm to the wound, local resident macrophages tend to migrate to the wound direction (Figures 5.11B,C and Figures 5.12B,C).

To quantify differences in macrophage migration behavior between *tlr2* and *myd88* mutants and their wild type siblings, we first analyzed whether the deficiency of *tlr2* and *myd88* can affect macrophage mean migration speed upon wounding. Following tail wounding, both distant and local resident macrophages migrate more slowly in the *tlr2*$^{-/-}$ and *myd88*$^{-/-}$ mutant groups than in the wild type sibling controls (Figure 5.9D; Figure 5.10D; Figure 5.11D; Figure 5.12D).

Subsequently, we studied the directional persistence of macrophage migration upon wounding. To this end, we quantified the net displacement, meandering index and mean V_{AP} in the *tlr2* and *myd88* mutants and siblings. The net displacement of the distant macrophages (Table 5.1. Eq. 5.1) was reduced in the *tlr2*$^{-/-}$ and *myd88*$^{-/-}$ mutants compared to the controls

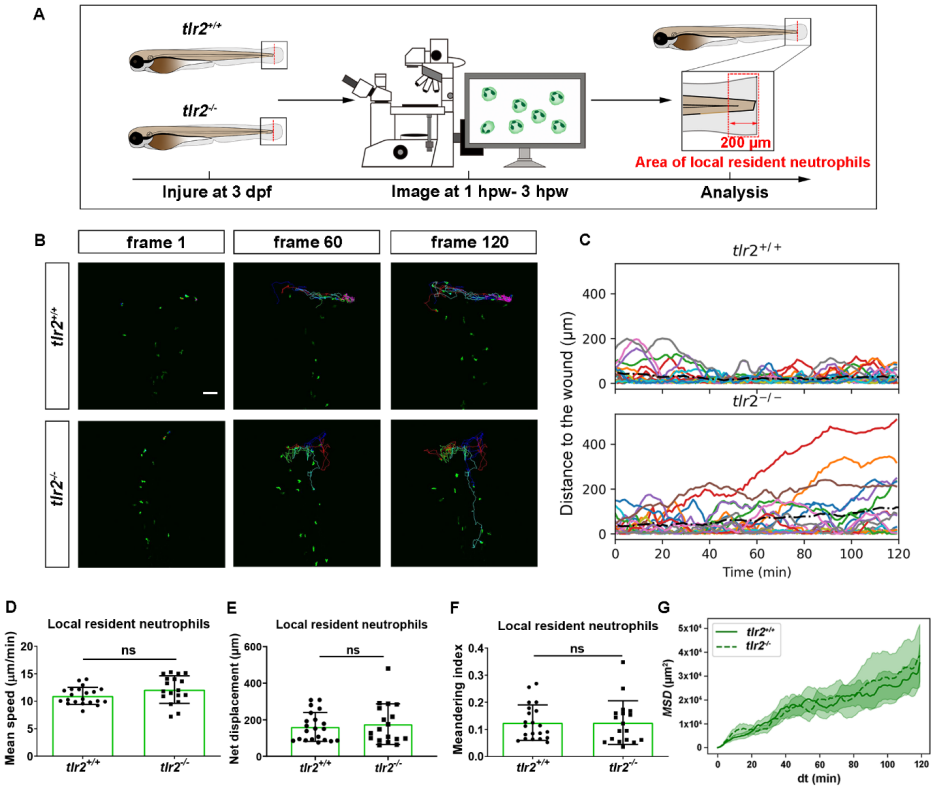


FIGURE 5.7: **Quantification of localized resident neutrophils behavior in wounded *tlr2* larvae. (A)** Experimental scheme. **(B)** Representative images of local resident neutrophils tracks in the wounded tail fin of 3 dpf *tlr2*$^{+/+}$ or *tlr2*$^{-/-}$ larvae at frame 1, frame 60 and frame 120. Cell tracking movies are shown in Supplementary Movie S17-18). Scale bar: 50 µm. **(C)** Distance to the wound. Black dash line represents average distance to the wound. Each color line represents one cell. **(D-I)** Quantification of local resident neutrophil tracks, mean speed (D); net displacement (E); Meandering index (F); MSD (G). An unpaired, two-tailed t-test was used to assess significance (ns, non-significance) and data are shown as mean± SD.

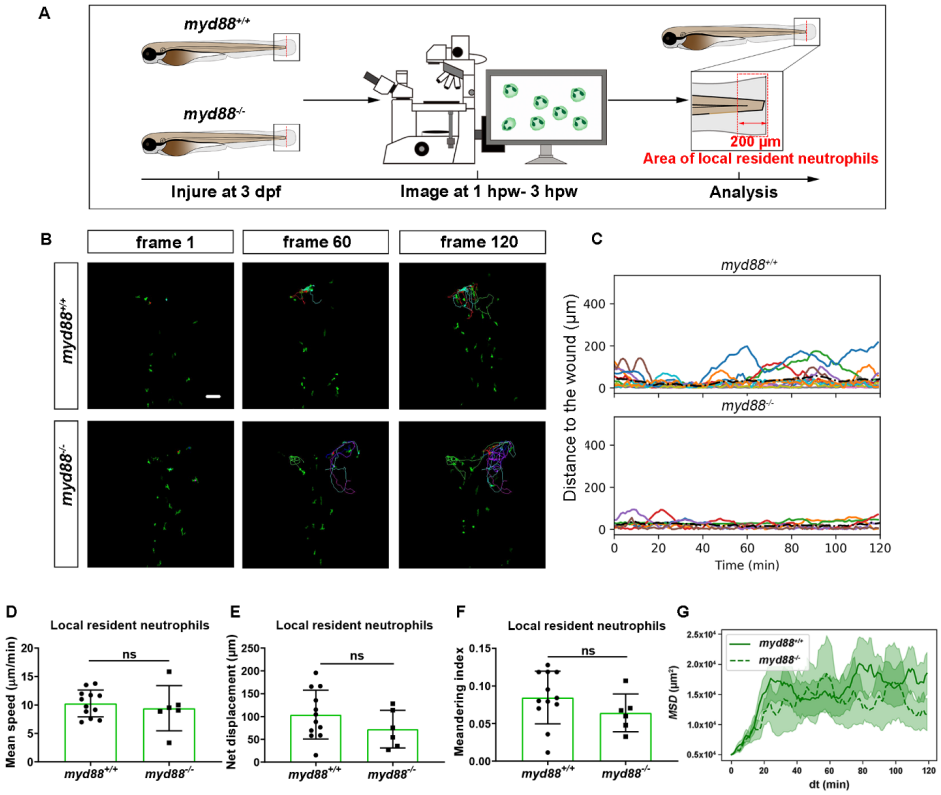
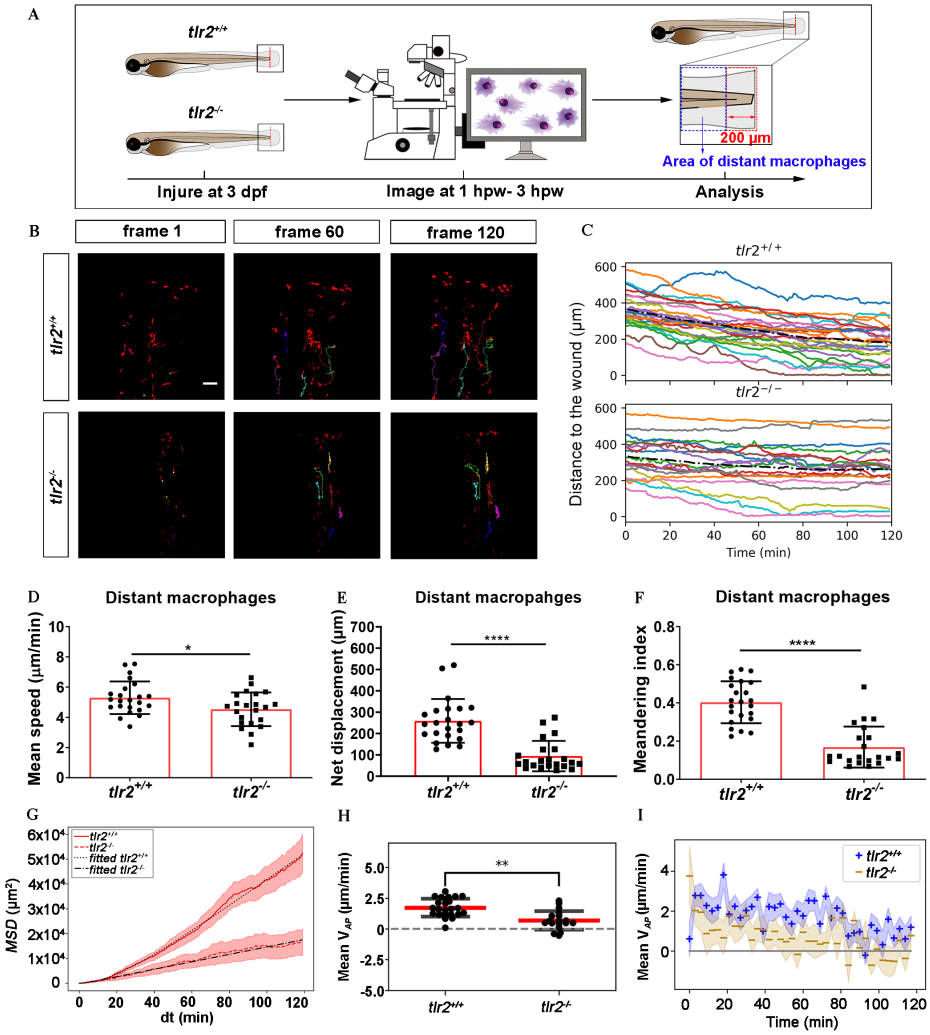


FIGURE 5.8: **Quantification of localized resident neutrophils behavior in wounded *myd88* larvae. (A)** Experimental scheme. **(B)** Representative images of local resident neutrophils tracks in the wounded tail fin of 3 dpf *myd88*$^{+/+}$ or *myd88*$^{-/-}$ larvae at frame 1, frame 60 and frame 120. Cell tracking movies are shown in Supplementary Movie S19-20). Scale bar: 50 µm. **(C)** Distance to the wound. Black dash line represents average distance to the wound. Each color line represents one cell. **(D-I)** Quantification of local resident neutrophil tracks, mean speed (D); net displacement (E); Meandering index (F); MSD (G). An unpaired, two-tailed t-test was used to assess significance (ns, non-significance) and data are shown as mean± SD.

(Figure 5.9E; Figure 5.10E). Furthermore, the meandering index (Table 5.1. Eq. 5.3) was also significantly decreased in the *tlr2*$^{-/-}$ and *myd88*$^{-/-}$ groups (Figures 5.9F,H and Figures 5.10), indicating that the decrease in net displacement was not only due to a reduced speed, but also due to loss of directionality. This is also supported by the significantly reduction in mean V_{AP} of distant macrophages in the *tlr2*$^{-/-}$ and *myd88*$^{-/-}$ groups (Figures 5.9F,H and Figures 5.10). However, no significant differences in net displacement were found in local resident *tlr2* and *myd88* macrophage groups (Supplementary Figure 5.11E and Supplementary Figure 5.12E). The trend of mean V_{AP} over time is similar to the one observed for distant neutrophils, in that *tlr2*$^{+/+}$ and *myd88*$^{+/+}$ macrophages have a higher mean V_{AP} than *tlr2*$^{-/-}$ and *myd88*$^{-/-}$ macrophages during the entire tracking period. The mean V_{AP} of macrophages is positive for a longer period of time compared to the neutrophils, as the majority of macrophages have not reached the wound site during the 2h time span.

The differences in speed and directionality also became apparent from the differences in MSD between the *tlr2*$^{+/+}$ and *myd88*$^{+/+}$ distant macrophages versus the *tlr2*$^{-/-}$ and *myd88*$^{-/-}$ distant macrophages (Figures 5.9G,5.10G).

FIGURE 5.9 *(following page)*: **Quantification of distant macrophage behavior in wounded *tlr2* mutant and sibling control larvae. (A)** Experimental scheme. *Tlr2*$^{+/+}$ and *tlr2*$^{-/-}$ larvae were wounded at 3 dpf. The red dashed line shows the site of wounding. Macrophages of wounded *tlr2* zebrafish larvae were tracked for 2 h and images were taken every 1 min by using CLSM. For cell tracking analysis, cells localized outside an area of 200 µm from the wounding edge toward the body trunk were counted as distant cells. Blue dashed box shows the area where distant macrophages were tracked. **(B)** Representative images of distant macrophage tracks in the wounded tail fin of 3 dpf *tlr2*$^{+/+}$ or *tlr2*$^{-/-}$ larvae at frame 1, frame 60 and frame 120. Time interval between two successive frames is 1 min. Each color track represents an individual macrophage. Cell tracking movies are shown in Supplementary Movie S13-14). Scale bar: 50 µm. **(C)** Distance to the wound. Black dash line represents average distance to the wound. Each color line represents one cell. **(D-I)** Quantification of distant macrophage tracks. There was a significant difference between the groups in terms of mean speed (D), net displacement (E), meandering index (F),), MSD (red) and fitted MSD (black) (G) and mean V_{AP} (H) of macrophages. The shaded regions in MSD (G) and mean V_{AP} over time (I) indicate standard error of the mean. Statistical analyses were done with 6 and 8 fish respectively for each group. An unpaired, two-tailed t-test was used to assess significance (ns, non-significance, *P < 0.05, **P < 0.01, ****P < 0.0001) and data are shown as mean± SD. Sample size (n): 23, 22 (D, E, F, H).


A
tlr2+/+
tlr2-/-
200 µm
Area of distant macrophages
Injure at 3 dpf
Image at 1 hpw- 3 hpw
Analysis
B
frame 1
frame 60
frame 120
tlr2+/+
tlr2-/-
C
tlr2+/+
tlr2−/−
Distance to the wound (µm)
Time (min)
D
Distant macrophages
Mean speed (µm/min)
E
Distant macropahges
Net displacement (µm)
F
Distant macrophages
Meandering index
G
MSD (µm²)
dt (min)
H
Mean VAP (µm/min)
I
Mean VAP (µm/min)
Time (min)

The MSD (Table 5.1. Eq. 5.5) is lower for the *tlr2*$^{-/-}$ and *myd88*$^{-/-}$ macrophages, which can reflect a speed reduction and/or a lowered directional persistence. A decreased directional persistence can also be seen through the shape of the MSD curve. For *tlr2*$^{+/+}$ and *myd88*$^{+/+}$ distant macrophages, the MSD curve, especially at short time intervals dt, has a parabolic shape, indicating straight cell trajectories. For *tlr2*$^{-/-}$, however, the MSD curve has a more linear shape, indicating random and non-persistent cell motility. The persistence time obtained from fitting Eq. 5.6, confirms this observation with an order of magnitude lower persistence time for the *tlr2*$^{-/-}$ macrophages compared to their wild type siblings group (Table 5.3). For the *myd88* groups, the persistence time is of the same order but larger in *myd88*$^{-/-}$ macrophages, which in turn have a lower intrinsic cell velocity v (Table 5.3). Consequently, the cell diffusivity D is also decreased in the *tlr2*$^{-/-}$ and *myd88*$^{-/-}$ macrophage groups compared to the *tlr2*$^{+/+}$ and *myd88*$^{+/+}$ macrophage groups. In summary, the data show that both *tlr2* and *myd88* mutations affect distant macrophage migration speed and directional persistence upon tail wounding.

FIGURE 5.10 *(following page)*: **Quantification of distant macrophages behavior in wounded *myd88* mutant and sibling control larvae. (A)** Experimental scheme. *Myd88*$^{+/+}$ and *myd88*$^{-/-}$ larvae were wounded at 3 dpf. The red dashed line shows the site of wounding. Macrophages of wounded zebrafish larvae were tracked for 2 h and images were taken every 1 min by using CLSM. For cell tracking analysis, cells localized outside an area of 200 µm from the wounding edge toward the body trunk were counted as distant cells. Blue dashed box shows the area where distant macrophages were tracked. **(B)** Representative images of distant macrophage tracks in the wounded tail fin of 3 dpf *myd88*$^{+/+}$ or *myd88*$^{-/-}$ larvae at frame 1, frame 60 and frame 120. Time interval between two successive frames is 1 min. Each color track represents an individual macrophage. Cell tracking movies are shown in Supplementary Movie S15-16). Scale bar: 50 µm. **(C)** Distance to the wound. Black dash line represents average distance to the wound. Each color line represents one cell. **(D-I)** Quantification of distant macrophage tracks. There was a significant difference between the groups in terms of mean speed (D), net displacement (E), meandering index (F), MSD (red) and fitted MSD (black) (G) and mean V_{AP} (H) of macrophages. Statistical analyses were done with 6 and 5 fish respectively for each group. The shaded regions in MSD (G) and mean V_{AP} over time (I) indicate standard error of the mean. An unpaired, two-tailed t-test was used to assess significance (ns, non-significance, *P < 0.05, **P < 0.01, ****P < 0.0001) and data are shown as mean± SD. Sample size (n): 32, 26 (D, E, F, H).

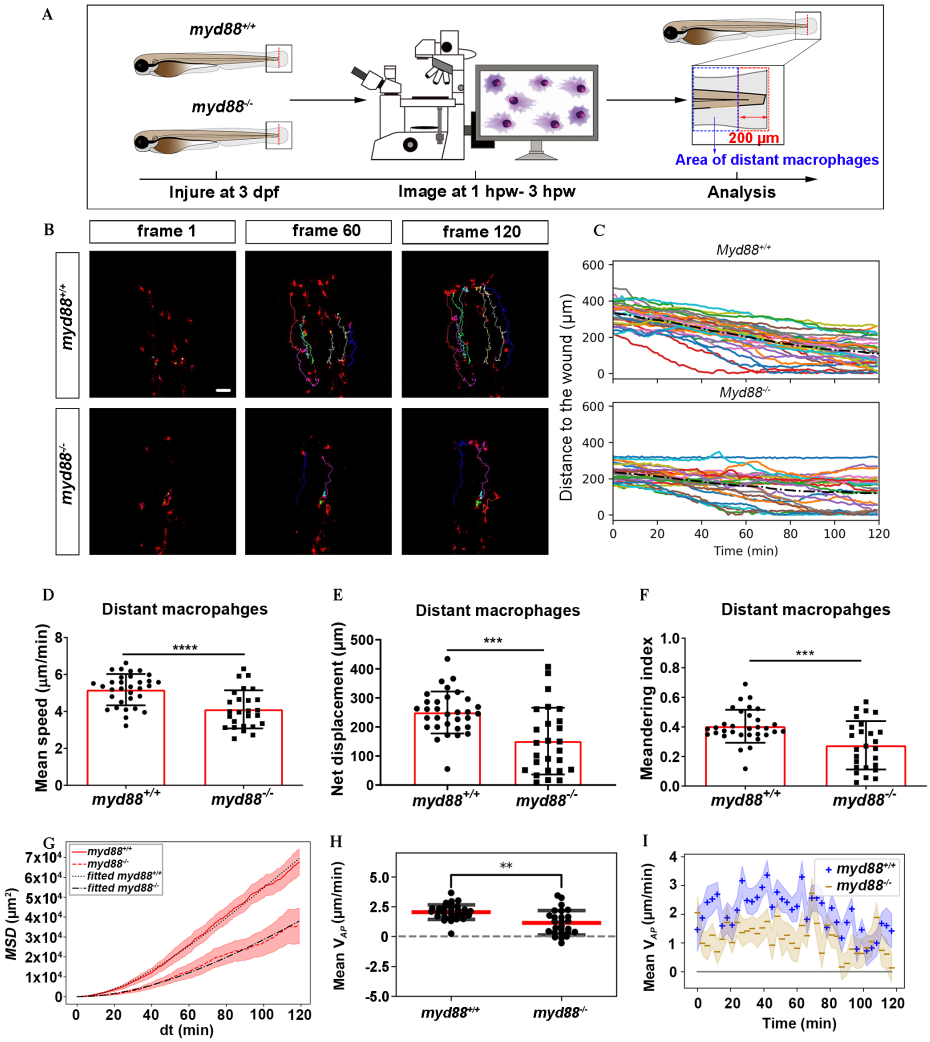


TABLE 5.3: Persistence time, intrinsic cell velocity and diffusivity obtained by fitting Eq. 5.6 to the MSD curves of distance macrophages.

Macrophages	τ (min)	v (µm/min)	D (µm^2/min)	Fitted time-frame
tlr2$^{+/+}$	20.83 ± 0.87	3.56 ± 0.04	132	$t \leq 120$
tlr2$^{-/-}$	2.47 ± 0.23	5.51 ± 0.23	38	
myd88$^{+/+}$	61.38 ± 3.81	3.04 ± 0.03	284	
myd88$^{-/-}$	104.36 ± 14.35	2.06 ± 0.03	221	

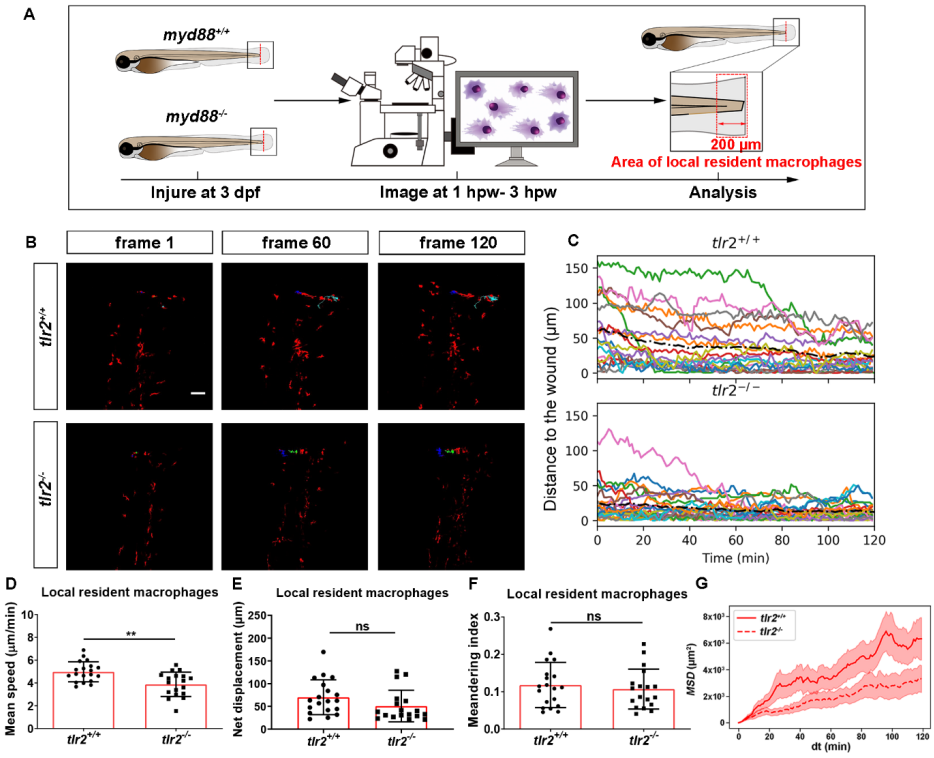


FIGURE 5.11: **Quantification of localized resident macrophages behavior in wounded *tlr2* larvae. (A)** Experimental scheme. **(B)** Representative images of local resident macrophages tracks in the wounded tail fin of 3 dpf *tlr2*$^{+/+}$ or *tlr2*$^{-/-}$ larvae at frame 1, frame 60 and frame 120. Cell tracking movies are shown in Supplementary Movie S21-22). Scale bar: 50 µm. **(C)** Distance to the wound. Black dash line represents average distance to the wound. Each color line represents one cell. **(D-I)** Quantification of local resident macrophage tracks, mean speed (D); net displacement (E); Meandering index (F); MSD (G). An unpaired, two-tailed t-test was used to assess significance (ns, non-significance) and data are shown as mean± SD.

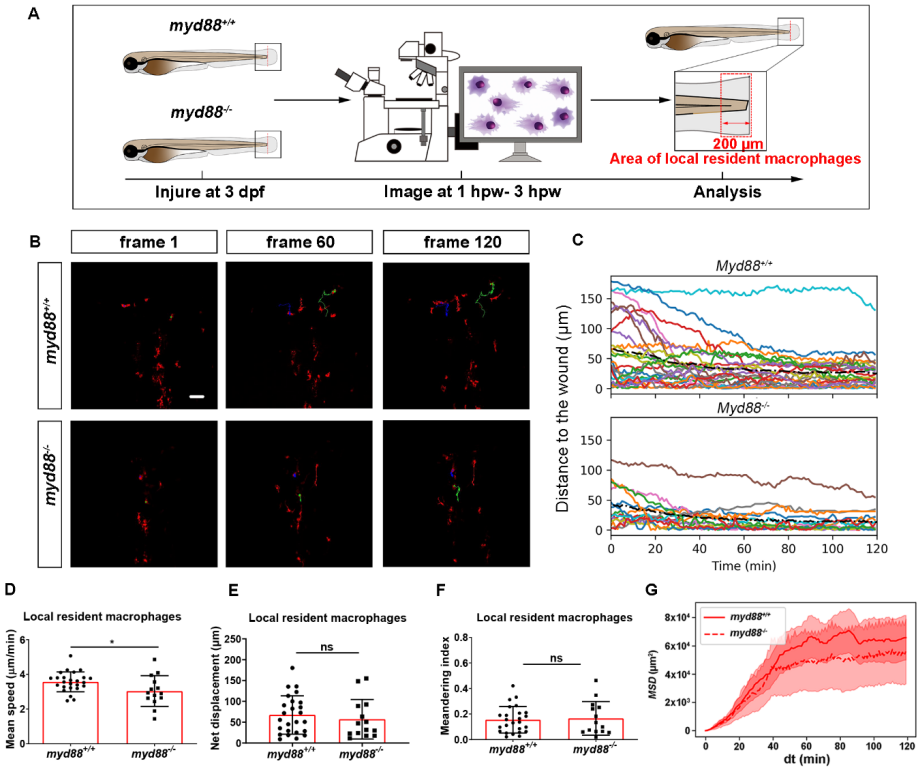


FIGURE 5.12: **Quantification of localized resident macrophages behavior in wounded *myd88* larvae. (A)** Experimental scheme. **(B)** Representative images of local resident macrophages tracks in the wounded tail fin of 3 dpf *myd88*$^{+/+}$ or *myd88*$^{-/-}$ larvae at frame 1, frame 60 and frame 120. Scale bar: 50 µm. **(C)** Distance to the wound. Black dash line represents average distance to the wound. Each color line represents one cell. Cell tracking movies are shown in Supplementary Movie S23-24). **(D-I)** Quantification of local resident macrophage tracks, mean speed (D); net displacement (E); Meandering index (F); MSD (G). An unpaired, two-tailed t-test was used to assess significance (ns, non-significance) and data are shown as mean± SD.

5.4 Discussion

In this study we visualized cell migration in *tlr2* and *myd88* mutants using live-imaging in a zebrafish tail wounding model. Thereby we demonstrated that these genes play a crucial role to control the migration of both neutrophils and macrophages upon tissue wounding. Like in mammals, neutrophils and macrophages play a dominant role in the wounding response during the first several hours after zebrafish tail fin wounding [267, 215, 252]. In mice, it has been shown previously that TLR signaling plays a role in controlling infiltration of neutrophils and macrophages into injured tissue [233, 234, 235, 236]. The function of TLR signaling in migration to epithelial wounds has only been studied so far in zebrafish larvae [242]. This study found that knock-down of *myd88* by morpholinos impairs the infiltration of neutrophils into the wound area, but the mechanisms underlying such reduced wound infiltration remained unknown. By using double transgenic lines, here we show that *tlr2* and *myd88* are both essential for directed migration of distant neutrophils and macrophages to the wounded tissue. The meandering index (Figure 5.1 and Table 5.1. Eq. 5.3) of distant neutrophils and macrophages was significantly decreased in *tlr2* and *myd88* mutant larvae compared with wild type sibling control groups (Figures 5.5F, 5.6F, 5.9F and 5.10F). Moreover, the migration speed of distant and local resident macrophages was decreased upon wounding in the *tlr2* and *myd88* mutants (Figures 5.9D and 5.10D; Supplementary Figures 5.115D and 5.126D), but not in unchallenged larvae. Taken together, these data suggest that TLR signaling is implicated in the sensitivity to signaling molecules secreted by the wound, explaining why less infiltration of neutrophils and macrophages is observed in tail wounds of the *tlr2* and *myd88* mutants (Figure 5.13).

The difference in directional persistence of the distant neutrophils and macrophages in the mutant shows already within 3 hours post wounding, suggesting that TLR signaling is involved in direct sensing of signals from the wound at the post-transcriptional level. However, since TLRs have not been implied in sensing meandering gradients, we assume that this function involves other receptors. Tlr2 has been shown to be essential for the regulation of cytokines and chemokines expression in both mice and zebrafish [236, 258]. For instance, we have shown that *tlr2* mutant shows a significant lower expression of *cxcl11aa* and also of a related chemokine, *cxcl11ac*, during mycobacterial infection. The CXCR3-CXCL11 chemokine-signaling axis has been demonstrated to play an essential role not only in infection process and but also in inflammation process by regulating leukocyte trafficking [264, 255]. It is possible that an insufficient level of

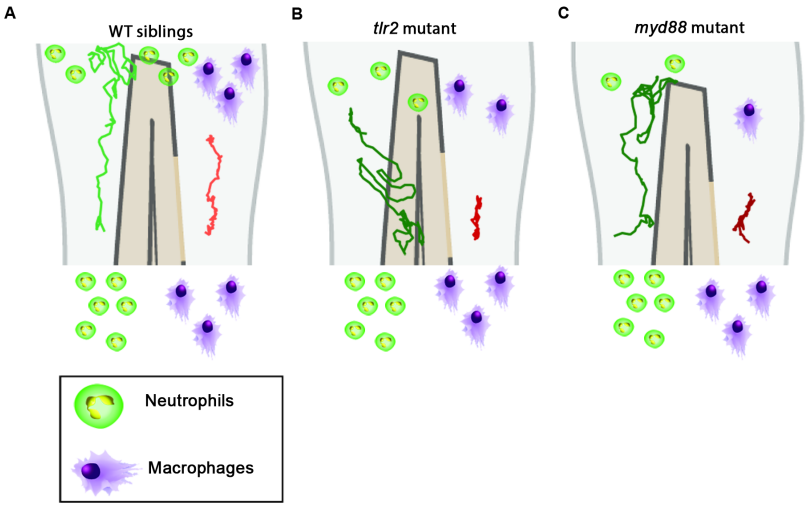


FIGURE 5.13: **Graphic summary of the data of cell migration behavior in the *tlr2* and *myd88* mutants and wild type siblings. (A)** Cell migration behavior in the wild type siblings. **(B)** Cell migration behavior in the *tlr2* mutant. **(C)** Cell migration behavior in the *myd88* mutant. In all cases, the green and red tracks are representative for the medians of the measured total displacements and net displacements in the anteroposterior axis of distant neutrophils and macrophages, respectively. The number of drawn leukocytes at the wound are only representing estimates of the relative numbers in the different genotypes. For the wild type sibling the *tlr2*$^{+/+}$ sibling was used as an example (A).

basal transcripts for chemokines at the time of wounding is responsible for the observed defects in leukocyte migration behavior. It is also possible that DAMPs released by dead cells around the wound do not lead to secretion of chemokines in the absence of TLR signaling. DAMPs are well known for activating PRRs and then activating downstream chemokines and cytokines secretion [224]. Molecules that can function as DAMPs and associated recognition factors during tissue injury such as hyaluronic acid and HMGB1, have been shown to be directly recognized by TLRs in tissues [268, 229, 269]. Chemokines can be produced by leukocytes which are exposed to reactive oxygen species (ROS) produced by injury [270, 213]. Moreover, previous studies have demonstrated that ROS are required for leukocyte recruitment upon wounding in the zebrafish larval model showing its function in long range chemotaxis to arachidonic acid [251, 254]. It has been demonstrated that the generation of ROS is related to TLR signaling in inflammation and tissue injury [271]. For example, Shishido et al. found that TLR2 mediates the generation of ROS after vascular injury ([272]. Thus, it is possible that the generation of ROS may be decreased in *tlr2* and *myd88* mutant zebrafish larvae.

To study the mechanistic basis of the differences in cell migratory behavior, mathematical and computational models can also provide insights. Chemokine and ROS gradients can easily be modelled by partial differential equations (PDEs). These can also be incorporated into a cell chemotaxis models, such as random walk models, phase field models, or the Cellular Potts model, with varying degrees of cell resolution, to study the chemotaxis of leukocytes. Such models could provide quantitative insights into how chemokine and ROS gradients affect the migration behavior of the leukocytes, and how the cells change these gradients by binding or secretion of chemokines or absorption and metabolizing ROS [273] which is known to affect the robustness of chemotaxis [274]. Using Bayesian inference on tracking data, one can infer a number of chemotaxis parameters, such as the flow rate, diffusion coefficient and production time of the chemoattractant [275]. Furthermore, simulated tracks can be compared to experimentally derived tracks. Altogether, such quantitative approaches in close interaction with new experiments could help demonstrate that the chemokine or ROS gradients are affected by the *tlr2* and *myd88* mutations. For such experiments we will need larger data sets than were currently obtained. This was partially due to the limitations of manual cell tracking. Therefore, in follow-up experiments with larger datasets, the tracking needs to be automated by using automatic tracking algorithms [276, 277, 278]. Currently, the Viterbi algorithm [278] cannot fully track the complex leukocyte cell behaviors: as cell are disappearing and appearing

in the tracking method, this leads to gaps in the time series images. Further optimization of this algorithm can result in the desired quantification of larger data sets.

Better theoretical cell migration analysis methods will also be useful for studying subsequent phases of the inflammatory response after wounding [213]. This can assist us in future studies focused on examining the involvement of the TLR signaling in neutrophil reverse migration and in the repair of wounded tissue. Previously we have reported that *myd88* mutant larvae that were raised under germ-free conditions show increased macrophage and decreased neutrophil numbers in the gut [279]. This indicates that the function of TLR signaling in leukocyte migration is dependent on the gut microbiota. It will be highly interesting to test whether the response of leukocytes to tail wounding is also dependent on the microbiome.

Data Availability Statement

The datasets in present study are available on request to the authors

Ethics Statement

No adult animals were used for experimentation. Larvae for experiments were obtained from zebrafish lines that were handled in compliance with the local animal welfare regulations and maintained according to standard protocols (zfin.org). The breeding of adult fish was approved by the local animal welfare committee (DEC) of the University of Leiden. All protocols adhered to the international guidelines specified by the EU Animal Protection Directive 2010/63/EU for which larvae under the age of 5 days post fertilization are not considered test animals.

Author Contributions

WH performed all biological experiments and manual cell tracking analyses and wrote the first version of the manuscript, LvS performed analyses of cell migration behavior and assisted with statistical analyses, RMHM, FJV and HPS supervised the study, HPS initiated the study and has the final responsibility of the manuscript. All authors delivered input for the final version of the manuscript and agreed with its contents. Chen Li and Lu Cao developed a script and performed automated cell tracking analyses for the published version.

Funding

W.H was supported by grants from the China Scholarship Council (CSC). R.M.H.M was supported by NWO Vici 865.17.004. The funding bodies had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Acknowledgments

We acknowledge Ulrike Nehrdich and Guus van der Velden for the assistance in adult zebrafish care.

Supplementary Material

The supplementary material can be accessed at the site of the publisher.

Chapter 6

Discussion

In this chapter, we will first summarize the results described in the previous chapters of this thesis. After that, we discuss future directions in the field of cell migration modeling.

6.1 Summarizing discussion

In Chapter 2, we presented ZebraGEM 2.0, an improved whole-genome scale metabolic reconstruction for zebrafish. Compared to the previous version of ZebraGEM [34], ZebraGEM 2.0 has been extended with the oxidative phosphorylation pathway, Gene-Protein-Reaction assocations and a more realistic biomass function. Due to the Gene-Protein-Reaction associations, it can now be used for knock-out studies, respiration experiments, and prediction of minimal feed, of which we have shown several examples. Furthermore, we analysed changes in metabolism upon *Mycobacterium marinum* infection by integrating gene expression data of control and *Mycobacterium marinum* infected zebrafish larvae. The model predicts a lowered growth rate and reduced histidine metabolism for the infected larvae. The biosignature of reduced histidine metabolism is also seen in other studies on human patients, mice and zebrafish [280, 281]. Overall, this improved model can be used to predict changes in zebrafish metabolism in other experimental conditions based on expression data, which can point out specific pathways, reactions or metabolites to further study experimentally. Moreover, the model can be used as a reference framework for interpreting omics data, for example, by showing RNAseq data on the metabolic network structure [282].

In Chapter 3, we studied the influence of cell-matrix adhesions on lymphocyte cell motility type. The type of motility is important for lymphocytes, as motility types differ in their effectiveness as immunosurveillance behaviour. In this chapter, we proposed an extension of the Act model [73] with cell-substrate adhesions to model lymphocyte motility. The model includes the adhesions between cell and extracellular matrix and contains four processes of adhesion dynamics: 1) *de novo* adhesion formation at the

actin-rich leading edge of the cell, 2) expansion of already existing adhesion patches, 3) spontaneous detachment of adhesions, and 4) adhesions breaking as the cell retracts. By increasing *de novo* adhesion formation, as well as the energy required to break adhesions, cell speed and cell persistence drop, and further increase results in pivoting behaviour, which is also observed in B-lymphocytes with sustained attachment areas [143]. However, the addition of these four processes fails to explain floating cells with no or few adhesions that are unable to migrate efficiently. Hence, we extended the model by including an extra feedback from the total adhesion area to the effectiveness of propulsion. Including this feedback, the model can display low motility at low total adhesion areas as well as stop-and-go motility types at sligthly higher *de novo* adhesion formation rates. Finally, we also saw that the ratio between the *de novo* adhesion formation and expansion of adhesion patches influences the spatial distribution of adhesions and the persistence of migration: cells with mainly small adhesion at the leading edge are more persistent than cells with a single or few larger adhesion patches near the cell center or rear. All in all, the behaviour captured by this model is very rich and is comparable to behaviours seen in different types of lymphocytes. Furthermore, the model show that parameter values regarding *de novo* adhesion formation, adhesion patch expansion and strength of the adhesion affect motility type. The molecular processes that underlie these parameters, such as which integrins are expressed and where they localize, or the strength of the matrix-integrin bond, could be studied experimentally to see if they result in the motility types predicted by the model. Such studies could deepen our understanding of how molecular defects in the interactions between immune cells and the ECM, e.g., in multiple myeloma [143] or inherited immune disease, eventually lead to altered immune cell migration and immunity defects.

In Chapter 4, we used the Cellular Potts model (CPM) to study why cells are efficient at performing topotaxis. Previous work on active Brownian particles (ABPs) has shown that part of the topotaxis effect can be explained by reduced cell persistence in denser pillar grids [75]. However, the active Brownian particle model cannot explain the extent to which *Dictyostelium discoideum* cells perform topotaxis. Using two different methods to model persistent cell motion, we fill in the gap between active Brownian particles and *Dictyostelium discoideum*. One method implements a persistent random walker model into the CPM and can be viewed as an ABP model with deformable volume. The other method, the Act model, phenomenologically models actin polymerization [73]. Both methods resulted in more efficient topotaxis than ABPs, so deformable volume makes cells more efficient at topotaxis. Furthermore, the actin-based method showed inhibition

of locomotion upon colliding with pillars, and reoriented in a different direction than the ABP-based method. This active reorientation leads to even more efficient topotaxis. We conclude that, for biological cells, cell volume and active reorientation enhance the persistence driven topotaxis already predicted by the ABP model. We can further use this model to explore more cell-steering grid properties, such as alternative pillar shapes or pillar adhesivity patterns, for applications in tissue engineering. Moreover, studying this model setup provides us with insight in cell motility in environment crowded by ECM and other cells, and, together with Chapter 3, allows us to explore how the cellular microenvironment can influence the direction and type of cell motility.

After mathematical of modelling cell motility, we turn to analyzing cell tracks of leukocytes in *in vivo* zebrafish in Chapter 5. Here, we studied the role of TLR2 and MYD88, both part of the TLR-signalling cascade, on leukocyte migration upon tail wounding in zebrafish larvae. Neutrophils and macrophages of both *tlr2*$^{-/-}$ larvae and *myd88*$^{-/-}$ larvae were compared with those of wildtype siblings. There was no difference in number of leukocytes and leukocyte basal migration between unchallenged mutant and wildtype larvae. However, upon tail-wounding, both *tlr2*$^{-/-}$ and *myd88*$^{-/-}$ larvae showed less recruitment of neutrophils and macrophages at 2 to 6 hours post wounding than their wildtype siblings. We further analysed cell track data of cells distant from the wound to study how leukocyte migration is changed in the mutants. Besides analyzing the speed, net displacement and meandering index, we also analyzed the mean velocity towards the tail end and the mean squared displacement from which we derived persistence times. For distant neutrophils there was no difference in speed, but the directional movement toward the wound and persistence of motility were reduced in the mutants compared to wild type neutrophils. For macrophages, there was a similar difference in directionality, but on top of that, the *tlr2*$^{-/-}$ and *myd88*$^{-/-}$ macrophages had lowered speed. From this extensive cell track analysis, we conclude that TLR2 and MYD88 play a role in the directionality of leukocyte migration upon wounding.

6.2 Future work

6.2.1 Combining signals in cell migration

In this thesis, we came across a number of cues that guide cells or influence cell motility, as seen in Chapters 3 and 4. Cells *in vivo* encounter many of these cues at the same time. There have been some efforts in understanding combined cues. Li et al. used a 3D model of breast cancer cell migration

with interstitial fluid flow, autochemotaxis and ECM fibres to study how these cues are combined in cell motion [283]. They showed that the flow of self-secreted chemoattractant and the alignment of the ECM fibres with the fluid flow resulted in synergistic motility: cell displacement was higher when fluid flow and ECM alignment with fluid flow were increased together than the sum of cell displacement when only one of the two effects (increased fluid flow; increased ECM alignment with fluid flow) were applied. In an experimental topotaxis setup, the effects of chemotaxis and topotaxis on directed migration of *Dictyostelium discoideum* were studied [74]. Here, the sum of drift in the case of aligned chemotaxis and topotaxis and in the case of opposed chemotaxis and topotaxis is equal to twice the drift of topotaxis. Thus, the chemotactic and topotactic drifts can be added up vectorially. We still need more insight in how cues work together to be able to grasp cell motility *in vivo*. Model studies can aid in this endeavour by explicitly integrating multiple cues at the same time. An interesting aspect to study here is the interplay of the molecular machinery of cell locomotion and the different signalling cues.

6.2.2 A stroll in the cellular landscape

Aside from the chemical signaling in chemotaxis and haptotaxis, more studies now also focus on the structure of the environment. While cells move through a tissue, they encounter non-motile cells, ECM, interstitial fluid, which vary through different tissues; bone tissue is structured differently from lymph nodes. In seperate models from Hecht et al. and Tweedy et al. [284, 285] chemotaxing cells within a maze were studied and both models showed that the interplay between chemotaxis and the environmental structure can result in directional cues. In the model from Hecht et al., cells could get stuck in a dead end of a maze which was permeable for a chemoattractant. Secretion of a chemorepellant which could not penetrate through the maze walls could resolve this. In the model of Tweedy et al., cells rapidly consumed the chemoattractant, guiding the cells away from chemoattractant depleted dead ends. Both studies contribute to the questions that arise: to what extent is cell motility shaped by the structural organization of the environment? And how can we find out about that? In the example of synergy between autochemotaxis, fluid flow and aligned ECM fibres [283], the model was used to test the different combinations of fibre and flow alignment and showed that they have a synergistic effect on directed cell motility. To extract this information from *in vivo* or *in vitro* experiments is a lot harder. The lack of knowledge on the exact shape of the environment could then skew the perceived effectiveness of chemotaxis.

Another example of where knowledge of the shape of the environment plays a role in interpretation of cell motility is a recent study on T cell motility in liver [286]. Cell tracking data showed that these T cells display superdiffusive behaviour, which is often associated with Lévy walks and optimal fouraging. However, the data lacked the infrequent large displacements, which are a key characteristic of Lévy walks. They hypothesized that T cells performed Brownian random walks, but that the channeled structure of the liver shaped their motion such that it became superdiffusive. To test this, they extracted the liver structure from imaging data and modeled the motion of Brownian walkers inside this structure. This sufficed to reproduce the superdiffusive behaviour, confirming their hypothesis that liver structure shapes cell motility.

A study combining modeling and imaging showed that the crowded environment of the lymph node plays a dominant role in T cell motion [72]. Both liver and lymph node show limited space for the cell to move in. This brings us to the point of the dimensionality of 3D environments. Obstacles in the form of cells, cell layers and ECM fibres can reduce the 3D space to 2D or 1D space for cells to move in. Some of this structure is immutable, such as mineralized ECM in bones, but cells can also alter parts of this structure; they can degrade or rearrange the ECM. Understanding the interplay between immutable and mutable structures in cell migration is useful for further understanding of immune cell patrolling as well as immune cell penetration in tissues such as tumours and granulomas.

6.2.3 Exchange between *in vivo* and *in vitro* motility parameters

The previous paragraphs point to a discrepancy between observed motility and the inherent motility. This also makes it hard to directly translate motility parameters between *in vivo* and *in vitro* data. Currently, Bayesian inference methods are being used to extract data on chemotaxis fields from cell track data from the lymph node [275]. However, the underlying model does not take the spatial structure of the lymph node into account. This could result in wrong estimates of the inferred chemotaxis fields or motility parameters.

Incorrect estimates often bring novel insights. In Chapter 4, we matched the CPM parameters such that our modeled cells had the same motility as *Dictyostelium* cells on a 2D substrate. The subsequent topotaxis experiments showed a discrepancy in speed between the *Dictoystelium* cells and the simulated cells. This teaches us that our model still lacks some elements that do play a role in the experiment. This points to further research on what

those elements are and, in the long run, contributes to our understanding of cell motion.

In Chapter 4, we used a simple hill-climbing algorithm to optimize two parameters to obtain experimental cell motion. In this case it was a straightforward answer to a straightforward question. However, the role of machine learning in this field is currently growing [287, 288]. When we want to increase the accuracy of inference methods, we must turn to more detailed models, which inevitably come with more parameters. Machine learning can aid in exploring the right areas of parameter space. Vice versa, computational modeling can also aid in quantifying the uncertainties in parameter estimations from *in vivo* or *in silico* data, by generating training and test data sets [288].

6.2.4 Patroling and more: other tasks of immune cells

So far, we have mainly discussed cell motility. However, cell migration is of course only a small aspect of the complex behaviour of cells. Leukocytes, such as neutrophils and macrophages play a role in clearing out pathogens, and hence, must also direct part of their energy to digestion of pathogens. Furthermore, they also relay and amplify their own recruitment by producing cytokine and chemokines, which also requires a portion of their energy. Hence, we can view the different tasks they have to fulfill from a metabolic viewpoint.

When we want to understand the possible trade-offs in immune cell motility, pathogen clearance and cell signaling, we can of course make use of multiscale models. Integrating metabolism in motility models such as the CPM or other agent-based modeling frameworks allows us test hypotheses on infection clearance on a tissue scale. Recent work by Graudenzi et al. combined the CPM and FBA framework by computing the growth rate for indivual cells in the CPM using FBA [289]. We can also think of combining constraint-based metabolic models with the CPM through the Hamiltonian of the CPM. We can compute the energy available for movement depending on the leukocyte state (migrating, phagocytizing, or signalling) and use that energy budget for energetically unfavourable moves, instead of the Boltzmann probability.

Advancements in single cell experimental techniques help in this approach. Sequencing data of isolated cells can function as a basis for the metabolic component of these models. Data from cells known to be in different states, such as infection state, or recruited/recruiting, would be of great value here.

Overall, this thesis presents a number of models that can be used as building blocks for multiscale modeling, where combining metabolism and cell migration models can give us further insight in how immune cells fight infections.

Bibliography

[1] N. Prajitha, S. S. Athira, and P. V. Mohanan. "Pyrogens, a polypeptide produces fever by metabolic changes in hypothalamus: Mechanisms and detections". *Immunology Letters* 204 (2018), 38–46.

[2] Bridgette M. Cumming, Hayden T. Pacl, and Adrie J.C. Steyn. "Relevance of the Warburg effect in tuberculosis for host-directed therapy". *Frontiers in Cellular and Infection Microbiology* 10 (2020), 506.

[3] Jae Sung Kim, Ye Ram Kim, and Chul Su Yang. "Host-directed therapy in tuberculosis: Targeting host metabolism". *Frontiers in Immunology* 11 (2020), 1790.

[4] Juho Vuononvirta, Federica M. Marelli-Berg, and Thanushiyan Poobalasingam. "Metabolic regulation of T lymphocyte motility and migration". *Molecular Aspects of Medicine* 77 (2020), 100888.

[5] Anusha Angajala et al. "Diverse roles of mitochondria in immune responses: Novel insights into immuno-metabolism". *Frontiers in Immunology* 9 (2018), 1605.

[6] Matthew F. Krummel, Frederic Bartumeus, and Audrey Gérard. "T cell migration, search strategies and mechanisms". *Nature Reviews Immunology* 16.3 (2016), 193–201.

[7] Hélène D. Moreau et al. "Integrating Physical and Molecular Insights on Immune Cell Migration". *Trends in Immunology* 39.8 (2018), 632–643.

[8] Alison Gaylo et al. "T cell interstitial migration: Motility cues from the inflamed tissue for micro- and macro-positioning". *Frontiers in Immunology* 7 (2016), 428.

[9] Nikolay Martyushenko and Eivind Almaas. "ModelExplorer - software for visual inspection and inconsistency correction of genome-scale metabolic reconstructions". *BMC Bioinformatics* (2019).

[10] Chiara Damiani et al. "Systems metabolomics: from metabolomic snapshots to design principles". *Current Opinion in Biotechnology* 63 (2020), 190–199.

[11] Eduard Sabidó et al. "Targeted proteomics of the eicosanoid biosynthetic pathway completes an integrated genomics-proteomics-metabolomics picture of cellular metabolism". *Molecular and Cellular Proteomics* 11.7 (2012), M111.014746–1 –M111.014746–9.

[12] Uwe Sauer. "Metabolic networks in motion: 13C-based flux analysis". *Molecular systems biology* 2.1 (2006), 62.

[13] Yujue Wang et al. "Utilizing tandem mass spectrometry for metabolic flux analysis". *Laboratory Investigation* (2020), 1–7.

[14] Nicolas Le Novère et al. "BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems." *Nucleic acids research* 34.Database issue (2006), D689–D691.

[15] Kieran Smallbone et al. "Towards a genome-scale kinetic model of cellular metabolism". *BMC Systems Biology* (2010).

[16] Charles J. Foster et al. "From Escherichia coli mutant 13C labeling data to a core kinetic model: A kinetic model parameterization pipeline". *PLoS Computational Biology* 15.9 (2019), e1007319.

[17] Kieran Smallbone et al. "Something from nothing - Bridging the gap between constraint-based and kinetic modelling". *FEBS Journal* 274.21 (2007), 5576–5585.

[18] Jeffrey D Orth, Ines Thiele, and Bernhard Ø Palsson. "What is flux balance analysis?" *Nat Biotechnol* 28.3 (2010), 245–248.

[19] Radhakrishnan Mahadevan, Jeremy S. Edwards, and Francis J. Doyle. "Dynamic Flux Balance Analysis of diauxic growth in Escherichia coli". *Biophysical Journal* 83.3 (2002), 1331–1340.

[20] Radhakrishnan Mahadevan and Chrisophe H. Schilling. "The effects of alternate optimal solutions in constraint-based genome-scale metabolic models". *Metabolic Engineering* 5.4 (2003), 264–276.

[21] Shiri Stempler, Keren Yizhak, and Eytan Ruppin. "Integrating transcriptomics with metabolic modeling predicts biomarkers and drug targets for Alzheimer's disease". *PLoS ONE* 9.8 (2014), e105383.

[22] Sabrina Kleessen et al. "Integration of transcriptomics and metabolomics data specifies the metabolic response of Chlamydomonas to rapamycin treatment". *Plant Journal* 81.5 (2015), 822–835.

[23] Keren Yizhak et al. "Integrating quantitative proteomics and metabolomics with a genome-scale metabolic network model". *Bioinformatics* 26.12 (2010), 255–260.

[24] Francesca Patella et al. "Proteomics-based metabolic modeling reveals that fatty acid oxidation (FAO) controls endothelial cell (EC) permeability." *Molecular & cellular proteomics : MCP* 14.3 (2015), 621–634.

[25] Brian J Schmidt et al. "GIM3E: Condition-specific models of cellular metabolism developed from metabolomics and expression data". *Bioinformatics* 29.22 (2013), 2900–2908.

[26] Svetlana Volkova et al. "Metabolic modelling as a framework for metabolomics data integration and analysis". *Metabolites* 10.8 (2020), 303.

[27] Jennifer L. Reed et al. "An expanded genome-scale model of Escherichia coli K-12 (iJR904 GSM/GPR)". *Genome biology* 4.9 (2003), R54.

[28] Jeremy S Edwards and Bernhard O Palsson. "The Escherichia coli MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities." *Proceedings of the National Academy of Sciences of the United States of America* 97.10 (2000), 5528–33.

[29] Natalie C Duarte, Markus J Herrgård, and Bernhard Ø Palsson. "Reconstruction and validation of Saccharomyces cerevisiae iND750, a fully compartmentalized genome-scale metabolic model." *Genome research* 14.7 (2004), 1298–309.

[30] Ines Thiele et al. "A community-driven global reconstruction of human metabolism". *Nature Biotechnology* 31.5 (2013), 419–425.

[31] Martin I Sigurdsson et al. "A detailed genome-wide reconstruction of mouse metabolism based on human Recon 1." *BMC systems biology* 4 (2010), 140.

[32] Hooman Hefzi et al. "A consensus genome-scale reconstruction of Chinese hamster ovary cell metabolism". *Cell Systems* 3.5 (2016), 434–443.

[33] Edik M. Blais et al. "Reconciled rat and human metabolic networks for comparative toxicogenomics and biomarker predictions". *Nature Communications* 8 (2017), 14250.

[34] Michaël Bekaert. "Reconstruction of Danio rerio Metabolic Model Accounting for Subcellular Compartmentalisation". *PLoS ONE* 7.11 (2012), e49903.

[35] Leonie van Steijn et al. "Predicting Metabolism from Gene Expression in an Improved Whole-Genome Metabolic Network Model of Danio rerio". *Zebrafish* 16.4 (2019), 348–362.

[36] Eileen Marie Hanna et al. "ReCodLiver0.9: Overcoming challenges in genome-scale metabolic reconstruction of a non-model species". *Frontiers in Molecular Biosciences* 7 (2020), 345.

[37] Svetlana Kalujnaia, Neil Hazon, and Gordon Cramb. "Myo-inositol phosphate synthase expression in the European eel (*Anguilla anguilla*) and Nile tilapia (*Oreochromis niloticus*): effect of seawater acclimation". *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 311.2 (2016), R287–R298.

[38] Matthew J. Footer et al. "Direct measurement of force generation by actin filament polymerization using an optical trap". *Proceedings of the National Academy of Sciences of the United States of America* 104.7 (2007), 2181–2186.

[39] Kinneret Keren et al. "Mechanism of shape determination in motile cells". *Nature* 453.7194 (2008), 475–480.

[40] Tim Lämmermann and Michael Sixt. "Mechanical modes of 'amoeboid' cell migration". *Current Opinion in Cell Biology* 21.5 (2009), 636–644.

[41] Verena Kölsch, Pascale G Charest, and Richard A Firtel. "The regulation of cell motility and chemotaxis by phospholipid signaling." *Journal of cell science* 121.Pt 5 (2008), 551–559.

[42] Scott A. Weed et al. "Cortactin localization to sites of actin assembly in lamellipodia requires interactions with F-actin and the Arp2/3 complex". *Journal of Cell Biology* 151.1 (2000), 29–40.

[43] Kate M. Byrne et al. "Bistability in the Rac1, PAK, and RhoA Signaling Network Drives Actin Cytoskeleton Dynamics and Cell Motility Switches". *Cell Systems* 2.1 (2016), 38–48.

[44] Sam Walcott and Sean X. Sun. "A mechanical model of actin stress fiber formation and substrate elasticity sensing in adherent cells". *Proceedings of the National Academy of Sciences of the United States of America* 107.17 (2010), 7757–7762.

[45] Albert Einstein. "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen [AdP 17, 549 (1905)]". *Annalen der Physik* 17 (1905), 549–560.

[46] William Sutherland. "A dynamical theory of diffusion for non-electrolytes and the molecular mass of albumin". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 9.54 (1905), 781–785.

[47] Reinhold Fürth. "Die Brownsche Bewegung bei Berücksichtigung einer Persistenz der Bewegungsrichtung. Mit Anwendungen auf die Bewegung lebender Infusorien". *Zeitschrift für Physik* 2.3 (1920), 244–256.

[48] V. Zaburdaev, S. Denisov, and J. Klafter. "Lévy walks". *Reviews of Modern Physics* 87.2 (2015), 483–530.

[49] Catherine Beauchemin, Narendra M. Dixit, and Alan S. Perelson. "Characterizing T Cell Movement within Lymph Nodes in the Absence of Antigen". *The Journal of Immunology* 178.9 (2007), 5505–5512.

[50] Roumen Tsekov and Marga C. Lensen. "Brownian Motion and the Temperament of Living Cells". *Chinese Physics Letters* 30.7 (2013), 070501.

[51] Hui Li et al. "Zigzag Generalized Lévy Walk: the In Vivo Search Strategy of Immunocytes." *Theranostics* 5.11 (2015), 1275–90.

[52] Elizabeth R. Jerison and Stephen R. Quake. "Heterogeneous T cell motility behaviors emerge from a coupling between speed and turning in vivo". *eLife* 9 (2020), e53933.

[53] Mark J Miller et al. "Autonomous T cell trafficking examined in vivo with intravital two-photon microscopy." *Proceedings of the National Academy of Sciences of the United States of America* 100.5 (2003), 2604–9.

[54] Tim Worbs et al. "CCR7 ligands stimulate the intranodal motility of T lymphocytes in vivo". *Journal of Experimental Medicine* 204.3 (2007), 489–495.

[55] Johannes Textor et al. "Defining the quantitative limits of intravital two-photon lymphocyte tracking." *Proceedings of the National Academy of Sciences of the United States of America* 108.30 (2011), 12401–6.

[56] Tajie H. Harris et al. "Generalized Lévy walks and the role of chemokines in migration of effector CD8+ T cells". *Nature* 486.7404 (2012), 545–548.

[57] Gabriel Espinosa-Carrasco et al. "Integrin $\beta 1$ optimizes diabetogenic T cell migration and function in the pancreas". *Frontiers in Immunology* 9.MAY (2018), 1156.

[58] Athanasius F.M. Marée et al. "Polarization and movement of keratocytes: A multiscale modelling approach". *Bulletin of Mathematical Biology* 68.5 (2006), 1169–1211.

[59] Athanasius F.M. Marée, Verônica A. Grieneisen, and Leah Edelstein-Keshet. "How cells integrate complex stimuli: The effect of feedback from phosphoinositides and cell shape on cell polarization and motility". *PLoS Computational Biology* 8.3 (2012), e1002402.

[60] Erin M. Craig et al. "Model for adhesion clutch explains biphasic relationship between actin flow and traction at the cell leading edge". *Physical Biology* 12.3 (2015), 035002.

[61] Erin L. Barnhart et al. "Adhesion-Dependent Wave Generation in Crawling Cells". *Current Biology* 27.1 (2017), 27–38.

[62] Elisabeth G. Rens and Leah Edelstein-Keshet. "From energy to cellular forces in the Cellular Potts Model: An algorithmic approach". *PLoS Computational Biology* 15.12 (2019), e1007459.

[63] Gaudenz Danuser, Jun Allard, and Alex Mogilner. "Mathematical modeling of eukaryotic cell migration: Insights beyond experiments". *Annual Review of Cell and Developmental Biology* 29 (2013), 501–528.

[64] Louis S. Prahl and David J. Odde. "Modeling cell migration mechanics". In: *Advances in Experimental Medicine and Biology*. Vol. 1092. 2018, 159–187.

[65] Falko Ziebert and Igor S. Aranson. "Effects of Adhesion Dynamics and Substrate Compliance on the Shape and Motility of Crawling Cells". *PLoS ONE* 8.5 (2013), e64511.

[66] Danying Shao, Herbert Levine, and Wouter Jan Rappel. "Coupling actin flow, adhesion, and morphology in a computational cell motility model". *Proceedings of the National Academy of Sciences of the United States of America* 109.18 (2012), 6851–6856.

[67] Jakob Löber, Falko Ziebert, and Igor S. Aranson. "Modeling crawling cell movement on soft engineered substrates". *Soft Matter* 10.9 (2014), 1365–1373.

[68] François Graner and James A. Glazier. "Simulation of biological cell sorting using a two-dimensional extended Potts model". *Physical Review Letters* 69.13 (1992), 2013–2016.

[69] Koo Chul Lee. "Rejection-free Monte Carlo technique". *Journal of Physics A: Mathematical and General* 28.17 (1995), 4835–4842.

[70] Noriyuki Bob Ouchi et al. "Improving the realism of the cellular Potts model in simulations of biological cells". *Physica A: Statistical Mechanics and its Applications* 329.3-4 (2003), 451–458.

[71] Nicholas J. Savill and Paulien Hogeweg. "Modelling Morphogenesis: From Single Cells to Crawling Slugs". *Journal of Theoretical Biology* 184.3 (1997), 229–235.

[72] Joost B. Beltman et al. "Lymph node topology dictates T cell migration behavior". *Journal of Experimental Medicine* 204.4 (2007), 771–780.

[73] Ioana Niculescu, Johannes Textor, and Rob J. de Boer. "Crawling and Gliding: A Computational Model for Shape-Driven Cell Migration". *PLoS Computational Biology* 11.10 (2015), e1004280.

[74] Joeri A J Wondergem et al. "Chemotaxis and topotaxis add vectorially for amoeboid cell migration". *bioRxiv* (2019).

[75] Koen Schakenraad et al. "Topotaxis of active Brownian particles". *Physical Review E* 101.3 (2020), 1–11.

[76] Asha Seth, Derek L. Stemple, and Inês Barroso. "The emerging use of zebrafish to model metabolic disease". *Disease Models & Mechanisms* 6.5 (2013), 1080–1088.

[77] Massimo M. Santoro. "Zebrafish as a model to explore cell metabolism". *Trends in Endocrinology and Metabolism* 25.10 (2014), 546–554.

[78] Kerstin Howe et al. "The zebrafish reference genome sequence and its relationship to the human genome". *Nature* 496.7446 (2013), 498–503.

[79] Nannan Chang et al. "Genome editing with RNA-guided Cas9 nuclease in Zebrafish embryos". *Cell Research* 23.4 (2013), 465–472.

[80] Rob C. van Wijk et al. "Systems pharmacology of hepatic metabolism in zebrafish larvae". *Drug Discovery Today: Disease Models* 22 (2016), 27–34.

[81] Jeremy S Edwards and Bernhard O Palsson. "Systems properties of the Haemophilus influenzae Rd metabolic genotype." *The Journal of biological chemistry* 274.25 (1999), 17410–6.

[82] Daniel Segrè, Dennis Vitkup, and George M Church. "Analysis of optimality in natural and perturbed metabolic networks." *Proceedings of the National Academy of Sciences of the United States of America* 99.23 (2002), 15112–7.

[83] Caroline Colijn et al. "Interpreting Expression Data with Metabolic Flux Models: Predicting Mycobacterium tuberculosis Mycolic Acid Production". *PLoS Computational Biology* 5.8 (2009), e1000489.

[84] Mats Åkesson, Jochen Förster, and Jens Nielsen. "Integration of gene expression data into genome-scale metabolic models". *Metabolic Engineering* 6.4 (2004), 285–293.

[85] Jochen Förster et al. "Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network." *Genome research* 13.2 (2003), 244–53.

[86] Markus J Herrgård et al. "A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology". *Nature Biotechnology* 26.10 (2008), 1155–1160.

[87] Ines Thiele et al. "A community effort towards a knowledge-base and mathematical model of the human pathogen Salmonella Typhimurium LT2". *BMC Systems Biology* 5.1 (2011), 8.

[88] Neema Jamshidi and Bernhard Ø Palsson. "Investigating the metabolic capabilities of Mycobacterium tuberculosis H37Rv using the in silico strain iNJ661 and proposing alternative drug targets". *BMC Systems Biology* 1.1 (2007), 26.

[89] Shuzhao Li et al. "Constructing a fish metabolic network model". *Genome Biology* 11.11 (2010), R115.

[90] Aarash Bordbar, Neema Jamshidi, and Bernhard O Palsson. "iAB-RBC-283: A proteomically derived knowledge-base of erythrocyte metabolism that can be used to simulate its physiological and patho-physiological states". *BMC Systems Biology* 5.1 (2011), 110.

[91] Aarash Bordbar et al. "A multi-tissue type genome-scale metabolic network for analysis of whole-body systems physiology." En. *BMC systems biology* 5.1 (2011), 180.

[92] Herman P Spaink, Hans J Jansen, and Ron P Dirks. "Advances in genomics of bony fish". *Briefings in Functional Genomics* 13.2 (2014), 144–156.

[93] Yen-Chia Chang et al. "Taurine homeostasis requires de novo synthesis via cysteine sulfinic acid decarboxylase during zebrafish early embryogenesis". *Amino acids* 44.2 (2013), 615–629.

[94] Zachary A. King et al. "BiGG Models: A platform for integrating, standardizing and sharing genome-scale models". *Nucleic Acids Research* 44.D1 (2016), D515–D522.

[95] Nathan E. Lewis et al. "Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models". *Molecular Systems Biology* 6 (2010), 390.

[96] Ali Ebrahim et al. "COBRApy: COnstraints-Based Reconstruction and Analysis for Python". *BMC Systems Biology* 7.1 (2013), 74.

[97] Daniel Machado and Markus Herrgård. "Systematic Evaluation of Methods for Integration of Transcriptomic Data into Constraint-Based Models of Metabolism". *PLoS Computational Biology* 10.4 (2014), e1003580.

[98] Rajib Saha, Anupam Chowdhury, and Costas D Maranas. "Recent advances in the reconstruction of metabolic models and integration of omics data". *Current Opinion in Biotechnology* 29 (2014), 39–45.

[99] Ehsan Motamedian et al. "TRFBA: an algorithm to integrate genome-scale metabolic and transcriptional regulatory networks with incorporation of expression data". *Bioinformatics* 33.7 (2016), 1057–1063.

[100] Nick Fyson et al. "Gene-centric constraint of metabolic models". *bioRxiv* (2017), 116558.

[101] Daniel Machado, Markus J. Herrgård, and Isabel Rocha. "Stoichiometric Representation of Gene–Protein–Reaction Associations Leverages Constraint-Based Analysis from Reaction to Gene-Level Phenotype Prediction". *PLoS Computational Biology* 12.10 (2016), 1–24.

[102] Young-Mo Kim et al. "Salmonella modulates metabolism during growth under conditions that induce expression of virulence genes". *Molecular BioSystems* 9.6 (2013), 1522.

[103] Michael Hucka et al. "The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models." *Bioinformatics (Oxford, England)* 19.4 (2003), 524–31.

[104] Brett G. Olivier and Frank T. Bergmann. "The Systems Biology Markup Language (SBML) Level 3 Package: Flux Balance Constraints". *Journal of Integrative Bioinformatics* 12.2 (2017), 660–690.

[105] Eric W. Sayers et al. "Database resources of the National Center for Biotechnology Information". *Nucleic Acids Research* 47.D1 (2019), D23–D28.

[106] Hiroyuki Ogata et al. "KEGG: Kyoto encyclopedia of genes and genomes". *Nucleic Acids Research* 27.1 (1999), 29–34.

[107] Minoru Kanehisa et al. "KEGG as a reference resource for gene and protein annotation". *Nucleic Acids Research* 44.D1 (2016), D457–D462.

[108] Minoru Kanehisa et al. "KEGG: New perspectives on genomes, pathways, diseases and drugs". *Nucleic Acids Research* 45.D1 (2017), D353–D361.

[109] Elizabeth Brunk et al. "Recon3D enables a three-dimensional view of gene variation in human metabolism". *Nature Biotechnology* 36.3 (2018), 272–281.

[110] Stefanía Magnúsdóttir et al. "Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota". *Nature Biotechnology* 35.1 (2016), 81–89.

[111] Scott A. Becker and Bernhard O. Palsson. "Context-specific metabolic networks are consistent with experiments". *PLoS Computational Biology* 4.5 (2008), e1000082.

[112] Ed Reznik, Pankaj Mehta, and Daniel Segrè. "Flux Imbalance Analysis and the Sensitivity of Cellular Growth to Changes in Metabolite Pools". *PLoS Computational Biology* 9.8 (2013), e1003195.

[113] Adam M. Feist and Bernhard O. Palsson. "The biomass objective function". *Current Opinion in Microbiology* 13.3 (2010), 344–349.

[114] Chih Hung Chou et al. "FMM: A web server for metabolic pathway reconstruction and comparative analysis". *Nucleic Acids Research* 37.SUPPL. 2 (2009), W129–W134.

[115] Wei-Dan Jiang et al. "Growth, digestive capacity and intestinal microflora of juvenile Jian carp (Cyprinus carpio var. Jian) fed graded levels of dietary inositol". *Aquaculture Research* 40.8 (2009), 955–962.

[116] Sanaz Khosravi et al. "Dietary myo-inositol requirement of parrot fish, Oplegnathus fasciatus". *Aquaculture* 436 (2015), 1–7.

[117] Shi-Yen Shiau and Shu-Lin Su. "Juvenile tilapia (Oreochromis niloticus × Oreochromis aureus) requires dietary myo-inositol for maximal growth". *Aquaculture* 243.1-4 (2005), 273–277.

[118] Yann Gibert, Sean L McGee, and Alister C Ward. "Metabolic profile analysis of zebrafish embryos." *Journal of visualized experiments : JoVE* 14.71 (2013), e4300.

[119] Adam Amsterdam et al. "Identification of 315 genes essential for early zebrafish development." *Proceedings of the National Academy of Sciences of the United States of America* 101.35 (2004), 12792–12797.

[120] Gregory Golling et al. "Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development". *Nature Genetics* 31.2 (2002), 135–140.

[121] Adam Amsterdam et al. "A large-scale insertional mutagenesis screen in zebrafish". *Genes and Development* 13.20 (1999), 2713–2724.

[122] W. Driever et al. "A genetic screen for mutations affecting embryogenesis in zebrafish". *Development* 123.1 (1996), 37–46.

[123] Akira Muto et al. "Forward Genetic Analysis of Visual Behavior in Zebrafish". *PLoS Genetics* 1.5 (2005), e66.

[124] Laurence D. Covassin et al. "A genetic screen for vascular mutants in zebrafish reveals dynamic roles for Vegf/Plcg1 signaling during artery development". *Developmental Biology* 329.2 (2009), 212–226.

[125] Norimasa Iwanami et al. "Forward Genetic Screens in Zebrafish Identify Pre-mRNA-Processing Pathways Regulating Early T Cell Development". *Cell Reports* 17.9 (2016), 2259–2270.

[126] Susan E Brockerhoff et al. "A behavioral screen for isolating zebrafish mutants with visual system defects". *Proceedings of the National Academy of Sciences* 92.23 (1995), 10545–10549.

[127] Richard N Smith et al. "InterMine: a flexible data warehouse system for the integration and analysis of heterogeneous biological data". *Bioinformatics* 28.23 (2012), 3163–3165.

[128] Wouter J Veneman et al. "Analysis of RNAseq datasets from a comparative infectious disease zebrafish model using GeneTiles bioinformatics". *Immunogenetics* 67.3 (2014), 135–147.

[129] Vinay Satish Kumar and Costas D. Maranas. "GrowMatch: An Automated Method for Reconciling In Silico/In Vivo Growth Predictions". *PLoS Computational Biology* 5.3 (2009), e1000308.

[130] Jennifer L. Reed et al. "Systems approach to refining genome annotation". *Proceedings of the National Academy of Sciences of the United States of America* 103.46 (2006), 17480–17484.

[131] Markus J Herrgård, Stephen S Fong, and Bernhard Ø Palsson. "Identification of genome-scale metabolic network models using experimentally measured flux profiles". *PLoS computational biology* 2.7 (2006), e72.

[132] Sylvain Prigent et al. "Meneco, a Topology-Based Gap-Filling Tool Applicable to Degraded Genome-Wide Metabolic Networks". *PLoS Computational Biology* 13.1 (2017), e1005276.

[133] Alex Thomas et al. "Network reconstruction of platelet metabolism identifies metabolic signature for aspirin resistance". *Scientific Reports* 4.1 (2015), 3925.

[134] Daniel Fraher et al. "Zebrafish embryonic lipidomic analysis reveals that the yolk cell is metabolically active in processing lipid". *Cell Reports* 14.6 (2016), 1317–1329.

[135] Douglas R Tocher et al. "Nutritional regulation of hepatocyte fatty acid desaturation and polyunsaturated fatty acid composition in zebrafish (Danio rerio) and tilapia (Oreochromis niloticus)". *Fish Physiology and Biochemistry* 24.4 (2001), 309–320.

[136] Daniel A. Beard, Shou-dan Liang, and Hong Qian. "Energy balance for analysis of complex metabolic networks". *Biophysical Journal* 83.1 (2002), 79–86.

[137] Giorgio Volpe and Giovanni Volpe. "The topography of the environment alters the optimal search strategy for active particles". *Proceedings of the National Academy of Sciences of the United States of America* 114.43 (2017), 11350–11355.

[138] Vincent Tejedor, Raphael Voituriez, and Olivier Bénichou. "Optimizing persistent random searches". *Physical Review Letters* 108.8 (2012), 088103.

[139] G. M. Viswanathan et al. "Optimizing the success of random searches". *Nature* 401.6756 (1999), 911–914.

[140] F. Bartumeus et al. "Optimizing the Encounter Rate in Biological Interactions: Lévy versus Brownian Strategies". *Physical Review Letters* 88.9 (2002), 097901.

[141] Paulus Mrass et al. "Cell-autonomous and environmental contributions to the interstitial migration of T cells". *Seminars in Immunopathology* 32.3 (2010), 257–274.

[142] Matthew F. Krummel, Frederic Bartumeus, and Audrey Gérard. "T cell migration, search strategies and mechanisms". *Nature Reviews Immunology* 16.3 (2016), 193–201.

[143] Javier Rey-Barroso et al. "Switching between individual and collective motility in B lymphocytes is controlled by cell-matrix adhesion and intercellular interactions". *Scientific Reports* 8.1 (2018), 5800.

[144] Jordan Jacobelli et al. "Myosin-IIA and ICAM-1 regulate the interchange between two distinct modes of T cell migration." *Journal of immunology (Baltimore, Md. : 1950)* 182.4 (2009), 2041–50.

[145] Calina A. Copos et al. "Mechanosensitive Adhesion Explains Stepping Motility in Amoeboid Cells". *Biophysical Journal* 112.12 (2017), 2672–2682.

[146] Guangyuan Yu et al. "Phenomenological modeling of durotaxis". *Physical Review E* 96.1 (2017), 010402.

[147] Inge M. N. Wortel et al. "Actin-inspired feedback couples speed and persistence in a Cellular Potts Model of cell migration". *bioRxiv* (2020).

[148] James A. Glazier and François Graner. "Simulation of the differential adhesion driven rearrangement of biological cells". *Physical Review E* 47.3 (1993), 2128–2154.

[149] Mark S. Alber et al. "On Cellular Automaton Approaches to Modeling Biological Cells". In: *Mathematical systems theory in biology, communications, computation, and finance*. 2003, 1–39.

[150] Raïssa Houmadi et al. "The Wiskott-Aldrich Syndrome Protein Contributes to the Assembly of the LFA-1 Nanocluster Belt at the Lytic Synapse". *Cell Reports* 22.4 (2018), 979–991.

[151] Kris A DeMali, Christy A Barlow, and Keith Burridge. "Recruitment of the Arp2/3 complex to vinculin: coupling membrane protrusion to matrix adhesion." *The Journal of cell biology* 159.5 (2002), 881–91.

[152] Grégory Giannone et al. "Lamellipodial Actin Mechanically Links Myosin Activity with Adhesion-Site Formation". *Cell* 128.3 (2007), 561–575.

[153] Paula Stanley et al. "Intermediate-affinity LFA-1 binds α-actinin-1 to control migration at the leading edge of the T cell". *EMBO Journal* 27.1 (2008), 62–75.

[154] Ana-Sunčana Smith and Udo Seifert. "Effective adhesion strength of specifically bound vesicles". *Physical Review E* 71.6 (2005), 061902.

[155] R Bruinsma, M Goulian, and P Pincus. "Self-assembly of membrane junctions." *Biophysical Journal* 67.2 (1994), 746.

[156] Erich Sackmann and Robijn F. Bruinsma. "Cell Adhesion as Wetting Transition?" *ChemPhysChem* 3.3 (2002), 262.

[157] Murray Eden. "A two-dimensional growth process". *Dynamics of fractal surfaces* 4 (1961), 223–239.

[158] David Selmeczi et al. "Cell motility as persistent random motion: Theories from experiments". *Biophysical Journal* 89.2 (2005), 912–931.

[159] Maria Zeitz, Katrin Wolff, and Holger Stark. "Active Brownian particles moving in a random Lorentz gas". *The European Physical Journal E* 40.2 (2017), 23.

[160] Peter Dieterich et al. "Anomalous dynamics of cell migration." *Proceedings of the National Academy of Sciences of the United States of America* 105.2 (2008), 459–63.

[161] Clarence E. Chan and David J. Odde. "Traction dynamics of filopodia on compliant substrates". *Science* 322.5908 (2008), 1687–1691.

[162] Ke Hu et al. "Differential transmission of actin motion within focal adhesions". *Science* 315.5808 (2007), 111–115.

[163] Carlos Jurado, John R. Haserick, and Juliet Lee. "Slipping or Gripping? Fluorescent Speckle Microscopy in Fish Keratocytes Reveals Two Different Mechanisms for Generating a Retrograde Flow of Actin". *Molecular Biology of the Cell* 16.2 (2005), 507–518.

[164] Azadeh Samadani, Jerome Mettetal, and Alexander Van Oudenaarden. "Cellular asymmetry and individuality in directional sensing". *Proceedings of the National Academy of Sciences of the United States of America* 103.31 (2006), 11549–11554.

[165] Edward J. Banigan et al. "Heterogeneous CD8+ T Cell Migration in the Lymph Node in the Absence of Inflammation Revealed by Quantitative Migration Analysis". *PLoS Computational Biology* 11.2 (2015), e1004058.

[166] K. Scott Weber, Mark J. Miller, and Paul M. Allen. "Th17 Cells Exhibit a Distinct Calcium Profile from Th1 and Th2 Cells and Have Th1-Like Motility and NF-AT Nuclear Localization". *The Journal of Immunology* 180.3 (2008), 1442–1450.

[167] Alison Gaylo-Moynihan et al. "Programming of Distinct Chemokine-Dependent and -Independent Search Strategies for Th1 and Th2 Cells Optimizes Function at Inflamed Sites". *Immunity* 51.2 (2019), 298–309.e6.

[168] Paolo Maiuri et al. "Actin flows mediate a universal coupling between cell speed and cell persistence". *Cell* 161.2 (2015), 374–386.

[169] Begoña Álvarez-González et al. "Three-dimensional balance of cortical tension and axial contractility enables fast amoeboid migration". *Biophysical Journal* 108.4 (2015), 821–832.

[170] Herbert B. Schiller et al. "β 1 - And α v -class integrins cooperate to regulate myosin II during rigidity sensing of fibronectin-based microenvironments". *Nature Cell Biology* 15.6 (2013), 625–636.

[171] René F.M. van Oers et al. "Mechanical Cell-Matrix Feedback Explains Pairwise and Collective Endothelial Cell Behavior In Vitro". *PLoS Computational Biology* 10.8 (2014), e1003774.

[172] Elisabeth G. Rens and Roeland M.H. Merks. "Cell Contractility Facilitates Alignment of Cells and Tissues to Static Uniaxial Stretch". *Biophysical Journal* 112.4 (2017), 755–766.

[173] Elisabeth G. Rens and Roeland M. H. Merks. "Cell Shape and Durotaxis Explained from Cell-Extracellular Matrix Forces and Focal Adhesion Dynamics". *iScience* 23.9 (2020), 101488.

[174] Jordan Jacobelli et al. "Confinement-optimized three-dimensional T cell amoeboid motility is modulated via myosin IIA-regulated adhesions". *Nature Immunology* 11.10 (2010), 953–961.

[175] Philip Vitorino et al. "MAP4K4 regulates integrin-FERM binding to control endothelial cell motility". *Nature* 519.7544 (2015), 425–430.

[176] Jorn Starruß et al. "Morpheus: A user-friendly modeling environment for multiscale and multicellular systems biology". *Bioinformatics* 30.9 (2014), 1331–1332.

[177] Pierre-François Lenne and Alice Nicolas. "Physics puzzles on membrane domains posed by cell biology". *Soft Matter* 5.15 (2009), 2841.

[178] Rudolf Merkel et al. "Energy landscapes of receptor–ligand bonds explored with dynamic force spectroscopy". *Nature* 397.6714 (1999), 50–53.

[179] Bryan T. Marshall et al. "Direct observation of catch bonds involving cell-adhesion molecules". *Nature* 423.6936 (2003), 190–193.

[180] Amit Pathak and Sanjay Kumar. "Independent regulation of tumor cell migration by matrix stiffness and confinement". *Proceedings of the National Academy of Sciences of the United States of America* 109.26 (2012), 10334–10339.

[181] Anthony Han et al. "Crowding tunes 3D collagen fibrils and reveals matrix regulation of cancer cell morphogenesis". *bioRxiv* (2018).

[182] Matthew F Krummel, Rachel S Friedman, and Jordan Jacobelli. "Modes and mechanisms of T cell motility: roles for confinement and Myosin-IIA". *Current Opinion in Cell Biology* 30 (2014), 9–16.

[183] Jennet Toyjanova et al. "Matrix confinement plays a pivotal role in regulating neutrophil-generated tractions, speed, and integrin utilization". *Journal of Biological Chemistry* 290.6 (2015), 3752–3763.

[184] Janina K. Hellmann et al. "Environmental constraints guide migration of malaria parasites during transmission". *PLoS Pathogens* 7.6 (2011), e1002080.

[185] Anna Battista, Friedrich Frischknecht, and Ulrich S. Schwarz. "Geometrical model for malaria parasite migration in structured environments". *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 90.4 (2014), 042720.

[186] Jin Seok Park et al. "Directed migration of cancer cells guided by the graded texture of the underlying matrix". *Nature Materials* 15.7 (2016), 792–801.

[187] Chaejeong Heo et al. "Cellular behavior controlled by bio-inspired and geometry-tunable nanohairs". *Nanoscale* 9.45 (2017), 17743–17751.

[188] Goher Mahmud et al. "Directing cell motions on micropatterned ratchets". *Nature Physics* 5.8 (2009), 606–612.

[189] David Caballero et al. "Ratchetaxis: Long-Range Directed Cell Migration by Local Cues." *Trends in cell biology* 25.12 (2015), 815–827.

[190] Delphine Arcizet et al. "Contact-controlled amoeboid motility induces dynamic cell trapping in 3D-microstructured surfaces". *Soft Matter* 8.5 (2012), 1473.

[191] Mari Gorelashvili et al. "Amoeboid migration mode adaption in quasi-3D spatial density gradients of varying lattice geometry". *New Journal of Physics* 16 (2014), 1–22.

[192] Liang Li, Edward C Cox, and Henrik Flyvbjerg. "'Dicty dynamics': Dictyostelium motility as persistent random motion." *Physical biology* 8.4 (2011), 046006.

[193] Elizaveta A. Novikova et al. "Persistence-Driven Durotaxis: Generic, Directed Motility in Rigidity Gradients". *Physical Review Letters* 118.7 (2017), 078103.

[194] Howard C. Berg and Douglas A. Brown. "Chemotaxis in Escherichia coli analysed by three-dimensional tracking". *Nature* 239.5374 (1972), 500–504.

[195] Ka Ming Pang, Eunkyung Lee, and David A. Knecht. "Use of a fusion protein between GFP and an actin-binding domain to visualize transient filamentous-actin structures". *Current Biology* 8.7 (1998), 405–408.

[196] Till Bretschneider et al. "Dynamic Actin Patterns and Arp2/3 Assembly at the Substrate-Attached Surface of Motile Cells". *Current Biology* 14.1 (2004), 1–10.

[197] Pascal Vallotton et al. "Tracking retrograde flow in keratocytes: News from the front". *Molecular Biology of the Cell* 16.3 (2005), 1223–1231.

[198] Shin I. Nishimura, Masahiro Ueda, and Masaki Sasai. "Non-Brownian dynamics and strategy of amoeboid cell locomotion". *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 85.4 (2012), 041909.

[199] Eric J. Campbell and Prosenjit Bagchi. "A computational model of amoeboid cell motility in the presence of obstacles". *Soft Matter* 14.28 (2018), 5741–5763.

[200] Andrew D. Doyle et al. "One-dimensional topography underlies three-dimensional fibrillar cell migration". *The Journal of Cell Biology* 184.4 (2009), 481–490.

[201] Laurent Pieuchot et al. "Curvotaxis directs cell migration through cell-scale curvature landscapes". *Nature Communications* 9.1 (2018), 1–13.

[202] Kwang Hoon Song et al. "Sinusoidal wavy surfaces for curvature-guided migration of Tlymphocytes". *Biomaterials* 51 (2015), 151–160.

[203] Jennifer A. Mitchel and Diane Hoffman-Kim. "Cellular scale anisotropic topography guides Schwann cell motility". *PLoS ONE* 6.9 (2011), e24316.

[204] Noritaka Masaki et al. "Robustness of self-organizing chemoattractant field arising from precise pulse induction of its breakdown enzyme: A single-cell level analysis of pde expression in dictyostelium". *Biophysical Journal* 104.5 (2013), 1191–1202.

[205] Daniele Proverbio and Marco Maggiora. "Dynamical strategies for obstacle avoidance during Dictyostelium discoideum aggregation: A Multi-agent system model". *arXiv* (2019).

[206] Ramon Grima. "Directed cell migration in the presence of obstacles". *Theoretical Biology and Medical Modelling* 4.1 (2007), 1–12.

[207] Cally Scherber et al. "Epithelial cell guidance by self-generated EGF gradients". *Integrative Biology* 4.3 (2012), 259.

[208] Elliot J. Marsden et al. "Chemotactic clusters in confined run-and-tumble bacteria: a numerical investigation". *Soft Matter* 10.1 (2014), 157–165.

[209] Rachelle N. Palchesko et al. "Development of Polydimethylsiloxane Substrates with Tunable Elastic Modulus to Study Cell Mechanobiology in Muscle and Nerve". *PLoS ONE* 7.12 (2012), e51499.

[210] Simon Youssef, Sebastian Gude, and Joachim O. Rädler. "Automated tracking in live-cell time-lapse movies". *Integrative Biology* 3.11 (2011), 1095–1101.

[211] Wanbin Hu et al. "A novel function of TLR2 and MyD88 in the regulation of leukocyte cell migration behavior during wounding in zebrafish larvae". *Frontiers in Cell and Developmental Biology* 9 (2021), 210.

[212] Graham J. Lieschke et al. "Morphologic and functional characterization of granulocytes and macrophages in embryonic and adult zebrafish". *Blood* 98.10 (2001), 3087–3096.

[213] Oliver Soehnlein and Lennart Lindbom. "Phagocyte partnership during the onset and resolution of inflammation". *Nature Reviews Immunology* 10.6 (2010), 427–439.

[214] Charles N. Serhan et al. "Resolution of in flammation: state of the art, definitions and terms". *The FASEB Journal* 21.2 (2007), 325–332.

[215] Li Li et al. "Live imaging reveals differing roles of macrophages and neutrophils during zebrafish tail fin regeneration". *Journal of Biological Chemistry* 287.30 (2012), 25353–25360.

[216] Carl Nathan. "Neutrophils and immunity: Challenges and opportunities". *Nature Reviews Immunology* 6.3 (2006), 173–182.

[217] Franklin H. Epstein and Stephen J. Weiss. "Tissue Destruction by Neutrophils". *New England Journal of Medicine* 320.6 (1989), 365–376.

[218] Jennifer C. Brazil, Nancy A. Louis, and Charles A. Parkos. "The role of polymorphonuclear leukocyte trafficking in the perpetuation of inflammation during inflammatory bowel disease". *Inflammatory Bowel Diseases* 19.7 (2013), 1556–1565.

[219] Paul Martin and S. Joseph Leibovich. "Inflammatory cells during wound repair: The good, the bad and the ugly". *Trends in Cell Biology* 15.11 (2005), 599–607.

[220] Anthony L. Mescher. "Macrophages and fibroblasts during inflammation and tissue repair in models of organ regeneration". *Regeneration* 4.2 (2017), 39–53.

[221] Sophie J. Hopkin et al. "Triggering the Resolution of Immune Mediated Inflammatory Diseases: Can Targeting Leukocyte Migration Be the Answer?" *Frontiers in Pharmacology* 10 (2019), 184.

[222] Takashi Hato and Pierre C. Dagher. "How the innate immune system senses trouble and causes trouble". *Clinical Journal of the American Society of Nephrology* 10.8 (2015), 1459–1469.

[223] Charles A. Janeway and Ruslan Medzhitov. "Innate immune recognition". *Annual Review of Immunology* 20 (2002), 197–216.

[224] Philipp Niethammer. "The early wound signals". *Current Opinion in Genetics and Development* 40 (2016), 17–22.

[225] Li Yu, Liantang Wang, and Shangwu Chen. "Endogenous toll-like receptor ligands and their biological significance". *Journal of Cellular and Molecular Medicine* 14.11 (2010), 2592–2603.

[226] Kumar Vijay. "Toll-like receptors in immunity and inflammatory diseases: Past, present, and future". *International Immunopharmacology* 59 (2018), 391–412.

[227] Valerie J. Quesniaux et al. "Toll-Like Receptor 2 (TLR2)-Dependent-Positive and TLR2-Independent-Negative Regulation of Proinflammatory Cytokines by Mycobacterial Lipomannans". *The Journal of Immunology* 172.7 (2004), 4425–4434.

[228] Alexander Poltorak et al. "Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: Mutations in Tlr4 gene". *Science* 282.5396 (1998), 2085–2088.

[229] Marco E. Bianchi. "HMGB1 loves company". *Journal of Leukocyte Biology* 86.3 (2009), 573–576.

[230] Hideyuki Yanai et al. "HMGB proteins function as universal sentinels for nucleic-acid-mediated innate immune responses". *Nature* 462.7269 (2009), 99–103.

[231] Hellen S. Teixeira et al. "TLR3-Dependent Activation of TLR2 Endogenous Ligands via the MyD88 Signaling Pathway Augments the Innate Immune Response". *Cells* 9.8 (2020).

[232] Laura Oliveira-Nascimento, Paola Massari, and Lee M. Wetzler. "The role of TLR2 ininfection and immunity". *Frontiers in Immunology* 3.APR (2012), 79.

[233] Jürgen Schauber et al. "Injury enhances TLR2 function and antimicrobial peptide expression through a vitamin D-dependent mechanism". *Journal of Clinical Investigation* 117.3 (2007), 803–811.

[234] Angela Castoldi et al. "TLR2, TLR4 and the Myd88 signaling pathway are crucial for neutrophil migration in acute kidney injury induced by sepsis". *PLoS ONE* 7.5 (2012), e37584.

[235] Yifei Xu et al. "Toll-like receptor 2 in promoting angiogenesis after acute ischemic injury". *International Journal of Molecular Medicine* 31.3 (2013), 555–560.

[236] Anna Moles et al. "A TLR2/S100A9/CXCL-2 signaling network is necessary for neutrophil recruitment in acute and chronic liver injury in the mouse". *Journal of Hepatology* 60.4 (2014), 782–791.

[237] Kamalika Mojumdar et al. "Divergent impact of Toll-like receptor 2 deficiency on repair mechanisms in healthy muscle versus Duchenne muscular dystrophy". *Journal of Pathology* 239.1 (2016), 10–22.

[238] Donghoon Kim et al. "Toll-like receptor 2 contributes to chemokine gene expression and macrophage infiltration in the dorsal root ganglia after peripheral nerve injury". *Molecular Pain* 7 (2011), 74.

[239] Ekihiro Seki, Eekjoong Park, and Jiro Fujimoto. "Toll-like receptor signaling in liver regeneration, fibrosis and carcinogenesis". *Hepatology Research* 41.7 (2011), 597–610.

[240] Kouichi Miura et al. "Toll-like receptor 2 and palmitic acid cooperatively contribute to the development of nonalcoholic steatohepatitis through inflammasome activation in mice". *Hepatology* 57.2 (2013), 577–589.

[241] Lingling Ji et al. "Toll like receptor 2 knock-out attenuates carbon tetrachloride (CCl 4)-induced liver fibrosis by downregulating MAPK and NF-κB signaling pathways". *FEBS Letters* 588.12 (2014), 2095–2100.

[242] Qing Deng, Elizabeth A. Harvie, and Anna Huttenlocher. "Distinct signalling mechanisms mediate neutrophil attraction to bacterial infection and tissue injury". *Cellular Microbiology* 14.4 (2012), 517–528.

[243] Kiyoshi Takeda and Shizuo Akira. "Microbial recognition by Toll-like receptors". *Journal of Dermatological Science* 34.2 (2004), 73–82.

[244] Lingfeng Chen et al. "Myeloid Differentiation Primary Response Protein 88 (MyD88): The Central Hub of TLR/IL-1R Signaling". *Journal of Medicinal Chemistry* 63.22 (2020), 13316–13329.

[245] Natalie Wagner et al. "Microglia Activation in Retinal Ischemia Triggers Cytokine and Toll-Like Receptor Response". *Journal of Molecular Neuroscience* (2020).

[246] Mohan R. Dasu et al. "TLR2 expression and signaling-dependent inflammation impair wound healing in diabetic mice". *Laboratory Investigation* 90.11 (2010), 1628–1636.

[247] Lisa Macedo et al. "Wound healing is impaired in MyD88-deficient mice: A role for MyD88 in the regulation of wound healing by adenosine A2A receptors". *American Journal of Pathology* 171.6 (2007), 1774–1788.

[248] Ruth A. Houseright et al. "Cell type specific gene expression profiling reveals a role for complement component C3 in neutrophil responses to tissue damage". *Scientific Reports* 10.1 (2020), 15716.

[249] Annemarie H. Meijer and Herman P. Spaink. "Host-Pathogen Interactions Made Transparent with the Zebrafish Model". *Current Drug Targets* 12.7 (2011), 1000–1017.

[250] Stephen A. Renshaw et al. "Atransgenic zebrafish model of neutrophilic inflammation". *Blood* 108.13 (2006), 3976–3978.

[251] Philipp Niethammer et al. "A tissue-scale gradient of hydrogen peroxide mediates rapid wound detection in zebrafish". *Nature* 459.7249 (2009), 996–999.

[252] Yufei Xie et al. "Glucocorticoids inhibit macrophage differentiation towards a pro-inflammatory phenotype upon wounding without affecting their migration". *DMM Disease Models and Mechanisms* 12.5 (2019).

[253] Audrey Bernut et al. "Deletion of cftr Leads to an Excessive Neutrophilic Response and Defective Tissue Repair in a Zebrafish Model of Sterile Inflammation". *Frontiers in Immunology* 11 (2020), 1733.

[254] Anushka Katikaneni et al. "Lipid peroxidation regulates long-range wound detection through 5-lipoxygenase in zebrafish". *Nature Cell Biology* 22.9 (2020), 1049–1055.

[255] Frida Sommer et al. "Frontline Science: Antagonism between regular and atypical Cxcr3 receptors regulates macrophage migration during infection and injury in zebrafish". *Journal of Leukocyte Biology* 107.2 (2020), 185–203.

[256] Katherine M. Henry et al. "Zebrafish as a model for the study of neutrophil biology". *Journal of Leukocyte Biology* 94.4 (2013), 633–642.

[257] Michiel van der Vaart et al. "Functional analysis of a zebrafish myd88 mutant identifies key transcriptional components of the innate immune system". *DMM Disease Models and Mechanisms* 6.3 (2013), 841–854.

[258] Wanbin Hu et al. "Infection and RNA-seq analysis of a zebrafish tlr2 mutant shows a broad function of this toll-like receptor in transcriptional and metabolic control and defense to Mycobacterium marinum infection". *BMC Genomics* 20.1 (2019), 878.

[259] Annemarie H. Meijer et al. "Expression analysis of the Toll-like receptor and TIR domain adaptor families of zebrafish". *Molecular Immunology* 40.11 (2004), 773–783.

[260] Shuxin Yang et al. "Common and specific downstream signaling targets controlled by Tlr2 and Tlr5 innate immune signaling in zebrafish". *BMC Genomics* 16.1 (2015), 547.

[261] Audrey Bernut et al. "Mycobacterium abscessus cording prevents phagocytosis and promotes abscess formation". *Proceedings of the National Academy of Sciences of the United States of America* 111.10 (2014), 943–52.

[262] Antonia Chatzopoulou et al. "Glucocorticoid-induced attenuation of the inflammatory response in zebrafish". *Endocrinology* 157.7 (2016), 2772–2784.

[263] Erik Meijering, Oleh Dzyubachyk, and Ihor Smal. "Methods for cell and particle tracking". *Methods in Enzymology* 504 (2012), 183–200.

[264] Vincenzo Torraca et al. "The CXCR3-CXCL11 signaling axis mediates macrophage recruitment and dissemination of mycobacterial infection." *Disease models & mechanisms* 8.3 (2015), 253–69.

[265] Cynthia L. Stokes, Douglas A. Lauffenburger, and Stuart K. Williams. "Migration of individual microvessel endothelial cells: Stochastic model and parameter measurement". *Journal of Cell Science* 99.2 (1991), 419–430.

[266] Harriet B. Taylor et al. "P38 and JNK have opposing effects on persistence of in vivo leukocyte migration in zebrafish". *Immunology and Cell Biology* 91.1 (2013), 60–69.

[267] Caroline Gray et al. "Simultaneous intravital imaging of macrophage and neutrophil behaviour during inflammation using a novel transgenic zebrafish". *Thrombosis and Haemostasis* 105.5 (2011), 811–819.

[268] Dianhua Jiang et al. "Regulation of lung injury and repair by Toll-like receptors and hyaluronan". *Nature Medicine* 11.11 (2005), 1173–1179.

[269] Kyoko Komai et al. "Role of scavenger receptors as damage-associated molecular pattern receptors in Toll-like receptor activation". *International Immunology* 29.2 (2017), 59–70.

[270] Shinichiro Yamamoto et al. "TRPM2-mediated Ca2+ influx induces chemokine production in monocytes that aggravates inflammatory neutrophil infiltration". *Nature Medicine* 14.7 (2008), 738–747.

[271] Manish Mittal et al. "Reactive oxygen species in inflammation and tissue injury". *Antioxidants and Redox Signaling* 20.7 (2014), 1126–1167.

[272] Tetsuro Shishido et al. "Central role of endogenous Toll-like receptor-2 activation in regulating inflammation, reactive oxygen species production, and subsequent neointimal formation after vascular injury". *Biochemical and Biophysical Research Communications* 345.4 (2006), 1446–1453.

[273] Erika Donà et al. "Directional tissue migration through a self-generated chemokine gradient". *Nature* 503.7475 (2013), 285–289.

[274] Luke Tweedy et al. "Self-Generated Chemoattractant Gradients: Attractant Depletion Extends the Range and Robustness of Chemotaxis". *PLOS Biology* 14.3 (2016), e1002404.

[275] Ioanna Manolopoulou et al. "Bayesian spatio-dynamic modeling in cell motility studies: Learning nonlinear taxic fields guiding the immune response". *Journal of the American Statistical Association* 107.499 (2012), 855–865.

[276] Jean Yves Tinevez et al. "TrackMate: An open and extensible platform for single-particle tracking". *Methods* 115 (2017), 80–90.

[277] Vladimír Ulman et al. "An objective comparison of cell-tracking algorithms". *Nature Methods* 14.12 (2017), 1141–1152.

[278] Klas E.G. Magnusson et al. "Global linking of cell tracks using the viterbi algorithm". *IEEE Transactions on Medical Imaging* 34.4 (2015), 911–929.

[279] Bjørn E.V. Koch et al. "Correction to: Intestinal microbiome adjusts the innate immune setpoint during colonization through negative regulation of MyD88 (Nature Communications, (2018), 9, 1, (4099), 10.1038/s41467-018-06658-4)". *Nature Communications* 10.1 (2019), 526.

[280] January Weiner et al. "Biomarkers of inflammation, immunosuppression and stress with active disease are revealed by metabolomic profiling of tuberculosis patients". *PLoS ONE* 7.7 (2012), e40221.

[281] Yi Ding et al. "Tuberculosis causes highly conserved metabolic changes in human patients, mycobacteria-infected mice and zebrafish larvae". *Scientific Reports* 10.1 (2020), 11635.

[282] Thomas Pfau et al. "IDARE2—Simultaneous Visualisation of Multiomics Data in Cytoscape". *Metabolites* 11.5 (2021).

[283] Ang Li et al. "Are the Effects of Independent Biophysical Factors Linearly Additive? A 3D Tumor Migration Model". *Biophysical Journal* 117.9 (2019), 1702–1713.

[284] Inbal Hecht et al. ""Self-Assisted" Amoeboid Navigation in Complex Environments". *PLoS ONE* 6.8 (2011), e21955.

[285] Luke Tweedy et al. "Seeing around corners: Cells solve mazes and respond at a distance using attractant breakdown". *Science* 369.6507 (2020).

[286] Harshana Rajakaruna et al. "Environment-imposed constraints make Brownian walkers efficient searchers". *bioRxiv* (2020).

[287] Ben Lambert et al. "Bayesian inference of agent-based models: a tool for studying kidney branching morphogenesis". *Journal of Mathematical Biology* 76.7 (2018), 1673–1697.

[288] Jochen Kursawe, Ruth E. Baker, and Alexander G. Fletcher. "Approximate Bayesian computation reveals the importance of repeated measurements for parameterising cell-based models of growing tissues". *Journal of Theoretical Biology* 443 (2018), 66–81.

[289] Alex Graudenzi, Davide Maspero, and Chiara Damiani. "Modeling Spatio-Temporal Dynamics of Metabolic Networks with Cellular Automata and Constraint-Based Methods". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 11115 LNCS. 2018, 16–29.

Samenvatting

Om infecties te bestrijden, gaat het lichaam hard aan het werk: het maakt ontstekingstoffen aan, verhoogt de lichaamstemperatuur, en de immuuncellen sporen de infectie op. In dit proefschrift onderzoeken we metabolisme tijdens een infectie, en de beweging en navigatie van immuuncellen. Hiervoor ontwikkelen we wiskundige en computationele modellen.

In hoofdstuk 2 bestuderen we het metabolisme van zebravissen die geïnfecteerd zijn met *Mycobacterium marinum*, verwant aan de tuberculosebacterie. Hiervoor verbeteren we een gepubliceerde netwerkreconstructie op genoomschaal van zebravismetabolisme, genaamd ZebraGEM. Naast verbeteringen in de pathways in het netwerk is een belangrijke verbetering in ZebraGEM 2.0 de toevoeging van Gen-Proteïne-Reactie-associaties (GPRs). Deze GPRs koppelen de enzymatische reacties in het netwerk met de genen die coderen voor dat enzym. Dit maakt het mogelijk te voorspellen hoe de metabole routes veranderen onder invloed van gen-knockouts. In 9% van de gen-knockouts voorspelt het model een afwijkende groei. Van deze knock-outs heeft 15% ook een afwijkend fenotype in experimentele studies. Bovendien hebben alle suboptimaal groeiende genknockouts een menselijke homoloog gelinkt aan een metabole ziekte. De GPRs stellen ons ook in staat om kwantitatieve genexpressiedata in het model te integreren door reacties met lage expressie van enzymgenen te begrenzen. Hiermee voorspellen we wat het zebravismetabolisme doet onder controle-omstandigheden en bij infectie met *Mycobacterium marinum*. Het model voorspelt dat geïnfecteerde larven langzamer groeien. Daarnaast wordt voor geïnfecteerde larven van vier dagen oud ook nog een verlaagd histidinemetabolisme voorspeld, wat in overeenstemt met experimentele observaties in zebravis.

Het uiteindelijke opruimen van een infectie gebeurt door immuuncellen. De volgende hoofdstukken gaan over de bewegingen en navigatie van immuuncellen. In hoofdstuk 3 onderzoeken we hoe de extracellulaire matrix, het netwerk van structurele eiwitten die veel van onze cellen omringen, de beweging van de immuuncellen beïnvloedt. Hiervoor breiden we een bestaande model voor persistente celbeweging uit met hechting van de cellen aan de matrix. Het model kan meerdere soorten beweging vertonen, grofweg in te delen in een toevalsbeweging met korte of lange persistentietijden (hoelang een cel dezelfde richting aanhoudt), of een beweging van vastgeplakte cellen die om hun aanhectingsplek pivoteren en uiteindelijk maar een klein gebiedje verkennen. Op korte tijdschalen zien we ook nog spurts waar de cellen voldoende grip op de ondergrond hebben afgewisseld met het verlies van grip en snelheid. Al met al wijst het model

erop dat de dynamiek van adhesie tussen de cellen en de omgeving de snelheid en persistentie van de celbeweging reguleert.

In hoofdstuk 4 onderzoeken we het fenomeen topotaxis, waar de vorm van de omgeving de cellen stuurt. Een zo'n sturende omgevingvorm is de dichtheid van obstakels, zoals andere cellen of extracellulaire matrix. Een modelomgeving hiervoor is een pilarenrooster waartussen cellen bewegen. Om het effect van obstakeldichtheid op celbeweging te bestuderen, laten we de dichtheid van de pilaren afnemen van links naar rechts. Als model voor immuuncellen gebruiken we *Dictyostelium discoideum* cellen, een amoebe die op vergelijkbare wijze beweegt als immuuncellen, maar eenvoudiger is om te kweken. Eerdere studies naar *D. discoideum*-cellen in zo'n rooster wezen uit dat de cellen gemiddeld gezien naar het minder dichte deel van het rooster bewegen, oftewel, de cellen vertonen topotaxis. *D. discoideum*-cellen bewegen persistent: ze blijven een tijd lang in dezelfde richting bewegen. Een eerder model van persistent bewegende deeltjes wees uit dat het verlies van persistentie bij een botsing een deel van de topotaxis kan verklaren. Echter, deze deeltjes hadden wel zwakkere topotaxis dan de cellen. Het Cellular Potts model (CPM) is een realistischer model voor *D. discoideum*-cellen dan het deeltjesmodel omdat het rekening houdt met de vervormbaarheid van de cellen. We verkrijgen op twee verschillende wijzen persistente cellen: de een op dezelfde wijze als het model van persistent bewegende deeltjes, de ander een variant van het CPM die de amoeboide beweging van immuuncellen en *Dictyostelium* cellen goed blijkt te benaderen. Eerst fitten we de modelparameters om celbeweging met dezelfde snelheid en persistentie als gemeten in *D. discoideum*-cellen te kunnen simuleren. Beide persistentiemodellen vertonen topotaxis, het eerste model 1-2% van de celbeweging. Het tweede model vertoont efficiëntere topotaxis met 3-6% van de celbeweging en benadert de experimenteel data van *D. discoideum*-cellen van 2,5-4% beter. Uit verdere analyse blijkt die efficiëntere topotaxis te komen doordat gesimuleerde cellen met het tweede model bij een botsing een groter verlies van persistentie kennen en zich sneller heroriënteren.

Na infectie of verwonding worden macrofagen en neutrofielen aangetrokken door de geïnfecteerde of verwonde plek via signaalmoleculen. In hoofdstuk 5 bestuderen we de paden van macrofagen en neutrofielen in zebravislarven na het amputeren van een deel van de staartvin. We kijken hierbij naar twee mutanten in de TLR2-signaleringspathway die een rol speelt in de herkenning van lichaamsvreemde stoffen: *tlr2* en *myd88*. De neutrofielen van wildtypelarven en mutanten verschillen niet in hun bewegingssnelheid of afgelegde afstand, maar de neutrofielen van mutanten bewegen minder gericht naar de wond toe. Voor macrofagen is er ook een

afname in snelheid en daarbovenop ook nog een verminderde gerichtheid naar de wond voor de macrofagen van mutanten ten opzichte van wildtype. De TLR2-signaleringspathway speelt dus een rol in de gerichte beweging in het wildtype tijdens het rekruteren van immuuncellen tijdens wondheling.

Het proefschrift eindigt met een samenvattende discussie en suggesties voor toekomstig werk. Hier bespreken we onder andere hoe cellen omgaan met meerdere omgevingssignalen, zoals chemische en topografische signalen, het verdere samenspel tussen celbeweging en de celomgeving, de verschillen en gelijkenissen tussen motiliteitsparameters van cellen in *in vitro* experimenten en cellen in levende organismen en het combineren van metabolismemodellen met het CPM.

Summary

The body is very active during an infection: it releases inflammatory compounds, raises the body temperature, and prompts immune cells to detect the infection. In this thesis, we study metabolism during infection, as well as the motion and navigation of immune cells. To this end, we develop mathematical and computational models.

In chapter 2, we study the metabolism of zebrafish larvae infected with *Mycobacterium marinum*, a relative of the bacterium that causes tuberculosis. We improve a published whole-genome metabolic network reconstruction of zebrafish metabolism, called ZebraGEM. Besides improvements on the network pathways, the major improvement in ZebraGEM 2.0 is the addition of Gene-Protein-Reaction associations (GPRs). These GPRs links the enzymatic reactions in the network to the genes coding for those enzymes. This makes it possible to predict changes in metabolic routes under gene knockouts. The model predicts reduced growth in 9% of the knockouts. From these knockouts, 15% also shows an aberrant phenotype in experimental studies. Morever, all growth reducing knockouts have a human homolog associated with metabolic diseases. The GPRs also facilitate integrating quantitative gene expression data into the model by limiting reactions with lowly expressed genes. We use this to predict zebrafish metabolism in control larvae and *Mycobacterium marinum*-infected larvae. The model predicts that infected larvae grow more slowly than uninfected larvae. Furthermore, four day old infected larvae have reduced histidine metabolism, in agreement with experimental observations.

Infections are cleared by immune cells. The next chapters study the motility and navigation of immune cells. In chapter 3, we study how the extracellular matrix, the protein network that surrounds most of our cells, influences the motility of immune cells. To this end, we extend a published model for persistent cell motility with the adhesion of cells to the matrix. The model can display multiple types of motility, which can be roughly divided into random walks with short or long persistence times (the time a cell will continue moving in the same direction), or a motility type where cells are stuck to matrix and pivot around their adhesive patch, and where the explored area remains small. On short time scales, we also observe accelerations when the cells have sufficient grip on the substrate alternated with the loss of grip and speed. Altogether, the model shows that the dynamics of adhesion between the cell and its surroundings regulate the speed and persistence of cell motility.

In chapter 4, we study a phenomenon called topotaxis: the shape of the environment guides the cells. One such environmental cue is the

density of obstacles such as other cells and extracellular matrix. As a model for obstacles within a tissue, we use a silicon pillar grid on which cells move in between the pillars. To study the effect of obstacle density, we change pillar density from left to right. As a model for motile immune cells, we study *Dictyostelium discoideum* cells, an amoeba that moves similar to immune cells, but can be cultured more easily. Previous studies on *D. discoideum* cells in such grids showed that cells drifted, on average, to the less dense area of the grid, i.e.: the cells perform topotaxis. A previous model of persistently moving particles showed that the loss of persistence upon collision can explain topotaxis in part. However, the particles drifted slower towards the less dense area than observed in cells. The Cellular Potts model (CPM) is a more realistic model for *D. discoideum* cells than the particle model, because it takes the malleability of cells into account. We model persistent cell motion in two ways: one is similar to the persistently moving particle model, the other is a variation of the CPM and closely resembles the amoeboid motion of immune cells and *D.discoideum*. Both persistence models perform topotaxis, the first model shows 1-2% of the cell speed. The second model shows more efficient topotaxis of 3-6% of cell speed and better resembles the experimental data of *D. discoideum* cells of 2.5-4% of cell speed. Further analysis shows that the more efficient topotaxis can be explained by the cells from the second model losing more of their persistence upon collision and sooner reorient themselves after that.

Upon infection or wounding, macrophages and neutrophils are recruited to the infected or wounded site by signalling molecules. In chapter 5, we study the trajectories of macrophages and neutrophils in zebrafish larvae upon a tailfin cut. We specifically study two mutants of the TLR2-signalling pathway which plays a role in recognizing microbial molecules: *tlr2* en *myd88*. Neutrophils of wild type larvae and mutant have similar speeds and displacements, but the mutant neutrophils show a less directed motion towards the wound. Mutant macrophages showed a lower speed beside a less directed motion towards the wound compared to wild type macrophages. We conclude that the TLR2 signalling pathway plays a role in the directed motion in wild type during the recruitment of immune cells upon wound healing.

The thesis ends with a summarizing discussion and suggestions for future work. We discuss how cells integrate multiple environmental cues, such as chemical and topographical signal, the interplay between cell motion and cell environment, the similarities and differences of motility parameters between cells of *in vitro* experiments and cells in live organisms, and combining metabolic models with the CPM.

Dankwoord

Tenslotte rest mij nog de personen te bedanken die mij tijdens dit promotieonderzoek hebben geholpen en ondersteund. Allereerst mijn eerste promotor Roeland Merks, wie ik het geluk had te ontmoeten tijdens mijn master. Zonder jouw enthousiasme was ik nooit aan een promotieonderzoek begonnen, en had ik het waarschijnlijk ook nooit afgemaakt. Je vertrouwen, zeker bij het starten met het onderwerp celbeweging, heeft regelmatig mijn zelfvertrouwen overstegen, maar het daardoor ook doen groeien.

Mijn andere twee promotores, Herman Spaink en Fons Verbeek, wil ik ook bedanken. Het was in het begin wel even wennen aan de "professortaal" van jullie, maar langzaam begon ik te snappen dat het gaat om zoveel mogelijke verbindingen tussen allerlei zaken waar ik op dat moment nog geen weet van had. Herman, bedankt voor het noemen van alle biologische verbanden die in onze gesprekken opborrelden. Door jou heb ik geleerd dat sommige, mij onbenullig ogende dingen biologisch heel relevant kunnen zijn. Fons, bedankt voor de bredere blik die jij en je groep hebben laten zien van wat je allemaal met een computer en biologie kan doen.

I also want to thank Guy Theraulaz, Clément Sire and Loïc Dupré. Your inquisitive remarks have elevated the work beyond what I could have achieved on my own. Also, thanks for the Toulousian restaurant recommendations, they were delicious.

Ook wil ik mijn PhD-collega's bedanken met wie ik heb samengewerkt. Joeri en Koen, met jullie besmettelijke nieuwsgierigheid werd het werken aan het pilarenproject een bron van nieuwe spannende onderzoeksvragen. Wanbin, thank you for our joyful conversations of what we want to learn from the cell track data that you so diligently obtained by hand.

Daarnaast wil ik ook de schare aan kantoorgenoten bedanken die ik onderweg in de zes verschillende kamers waar ik gedurende mijn promotieonderzoek heb gezeten heb leren kennen. Dus, Esmée, Koen, Claudiu, Christian, Robbin, Olfa, Sandro, Roy, David, bedankt voor het oplossen van bugs, printproblemen en de hulp bij allerhande kleine en grote zaken. Lief, leed en lunch delen doet een mens goed, en ik heb jullie het afgelopen jaar dan ook zeer gemist.

En tenslotte wil ik mijn vriend en thuiskantoorgenoot Remie bedanken. Doordat jij ook aan het promoveren was, heb ik altijd kunnen rekenen op begrip voor de rare periodes die een promotieonderzoek soms kent. Bedankt voor het lezen van elke tekst die ik je heb voorgelegd en voor het zijn van oefenpubliek voor belangrijke presentaties. Zonder jou was het een minder interessante show geweest.

Curriculum Vitae

Leonie van Steijn werd geboren op 26 januari 1991 te Noordwijkerhout. In 2009 behaalde ze het gymnasiumdiploma met het profiel "Natuur en Techniek" en "Natuur en Gezondheid" aan het Fioretti College te Lisse. In 2009 was ze onderdeel van de Nederlandse delegatie voor de Internationale Biologie Olympiade te Japan, waar ze een bronzen medaille behaalde. Ze volgde de bacheloropleiding Wiskunde met minor Biologie aan de Universiteit Utrecht en behaalde hiervoor in 2013 haar diploma. Vervolgens behaalde ze in 2016 een masterdiploma van de opleiding Molecular and Cellular Life Sciences, eveneens aan de Universiteit Utrecht. Voor deze master deed ze twee onderzoeksstages: één onder begeleiding van dr. Rutger Hermsen aan de Universiteit Utrecht over de fixatiekans van een mutatie binnen een populatie, en één onder begeleiding van prof. dr. Roeland Merks aan het Centrum Wiskunde & Informatica over de effecten van antibiotica op de darmflora.

In 2016 begon ze aan promotieonderzoek aan het Mathematisch Instituut van de Universiteit Leiden onder begeleiding van prof. dr. Roeland Merks, prof. dr. Herman Spaink en prof. dr. Fons Verbeek. Hier werkte ze eerst aan het modelleren van zebravismetabolisme, en daarna aan het modelleren van celbeweging. Eind 2018 bracht ze drie maanden door aan de Université Paul Sabatier te Toulouse met steun van een Bourse d'Excellence Descartes van de Franse Ambassade om te werken aan een model voor lymfocytbeweging. Naast onderzoek heeft ze gedurende haar PhD ook vakken gedoceerd en geassisteerd, studenten begeleid en haar werk gepresenteerd op nationale en internationale conferenties.

In de zomer van 2021 begint ze als docent voor de opleiding Molecular and Biophysical Life Sciences aan de Universiteit Utrecht